

SocialRank: An Ego- and Time-Centric Workflow for Social Relationship Identification

Chris Diehl

October 15, 2008



Collaborators:

*Jaime Montemayor, Mike Pekala (JHU/APL)
Lise Getoor, Galileo Namata (UMCP)*

USMA-ARO Network Science Workshop

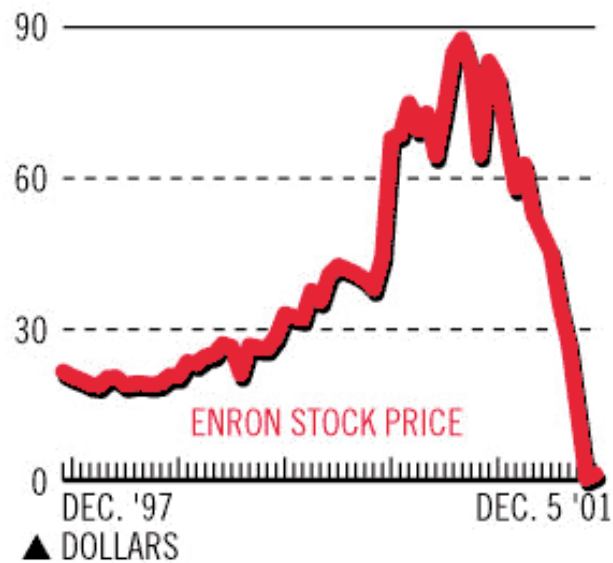
APL

The Johns Hopkins University
APPLIED PHYSICS LABORATORY

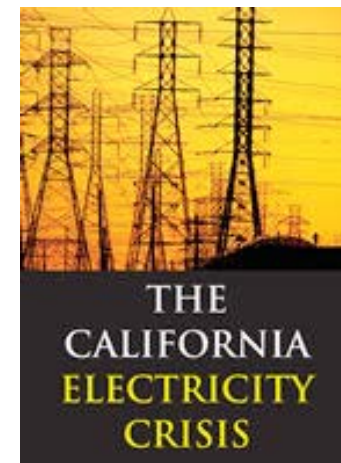
What Happened?



STUNNING COLLAPSE



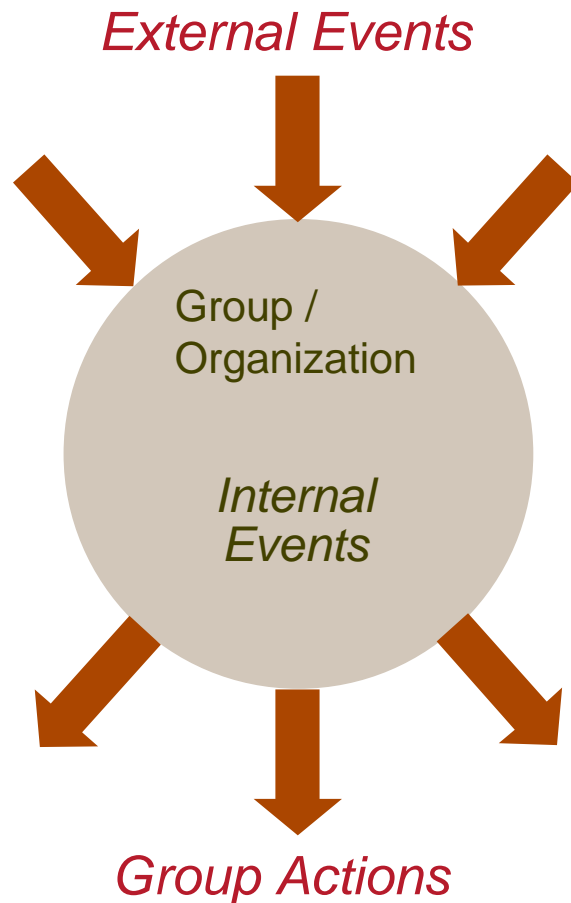
Data: Bloomberg Financial Markets



What Happened?

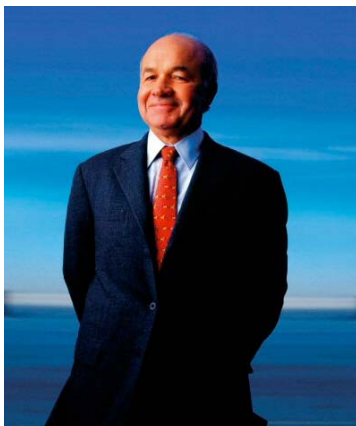


Retrospective Analysis

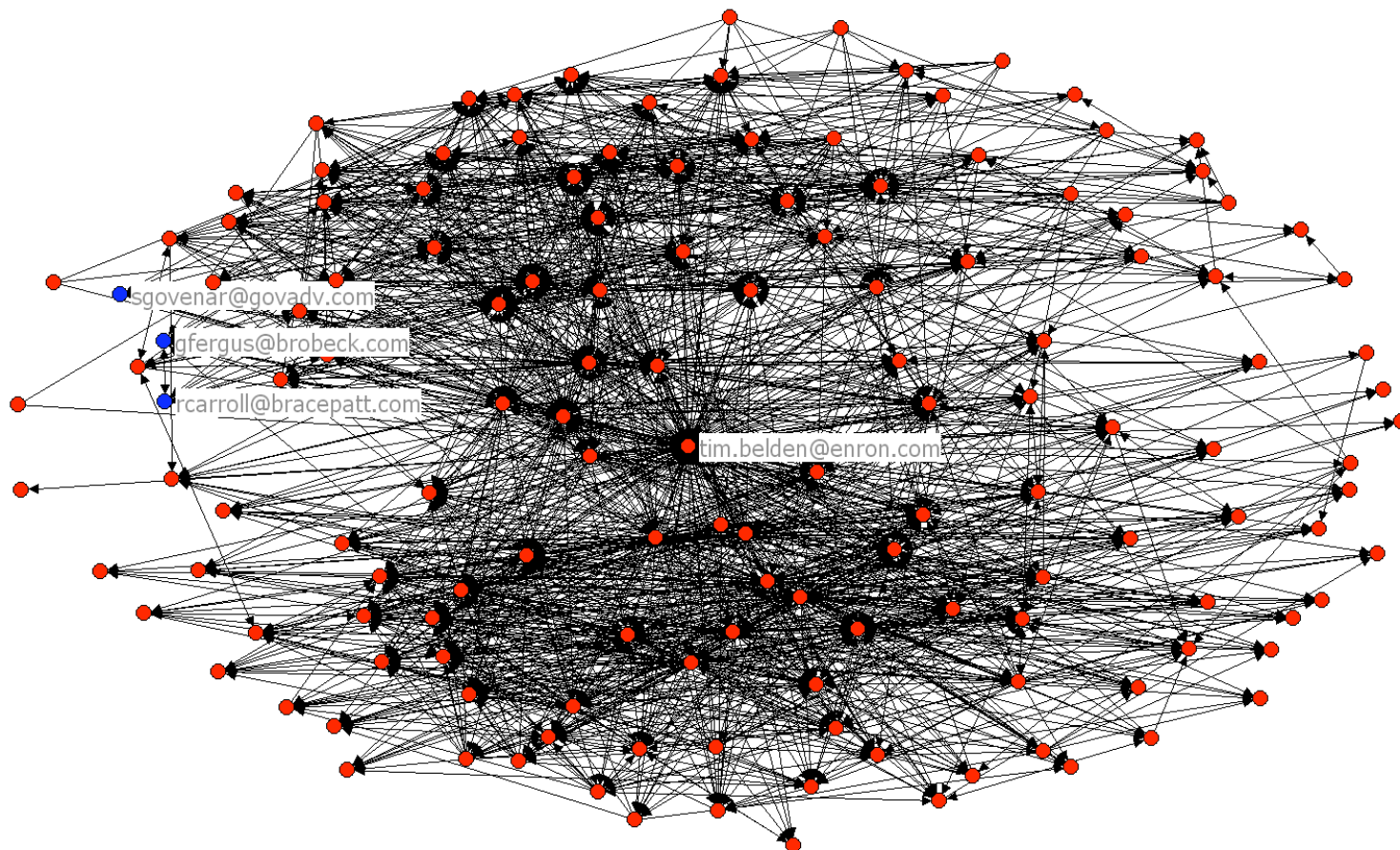


- **Group Structure and Context**
 - What defines the group?
 - What relationship structure exists?
 - What is its purpose?
- **Behavior - Trends and Shocks**
 - What factors are shaping the group's actions?
 - What external events have impacted the group's behavior?

The Enron Scandal: Key Players



Tim Belden's Ego Network With Alter-Alter Relationships



January 2000 - November 2001

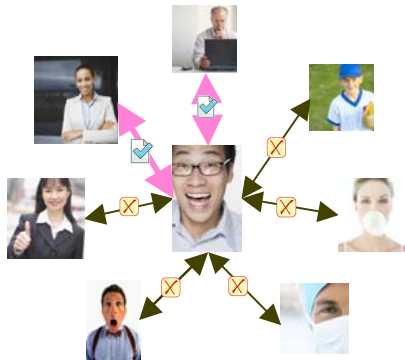
Validated Organizational Structure: Three Days Later



Relationship Discovery and Validation

Labeled Ego Networks

Training



Large-Margin Egocentric Relationship Ranking

$$h(f_r) - h(f_o) > 0$$

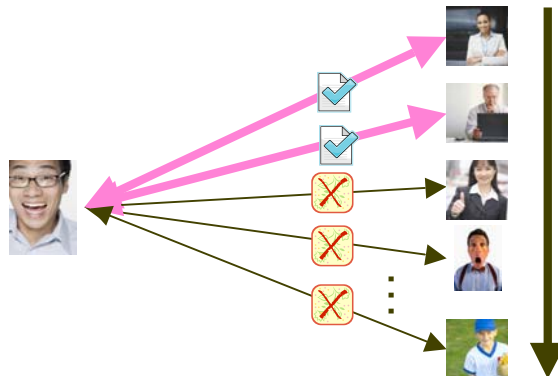
$$\forall f_r \in \mathcal{F}_r(n), f_o \in \mathcal{F}_o(n), n \in N_t$$

$$h(f) = w \cdot \Phi(f) : \mathbb{R}^p \rightarrow \mathbb{R}$$

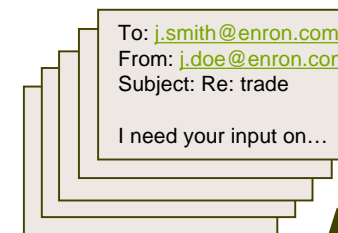
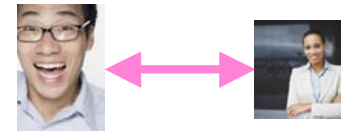
$$C(w) = \frac{1}{2} \|w\|^2 + \lambda \sum_{n \in N_t} \sum_{f_r \in \mathcal{F}_r(n)} \sum_{f_o \in \mathcal{F}_o(n)} g(m(f_r, f_o))$$

Relationship Ranking

Analysis



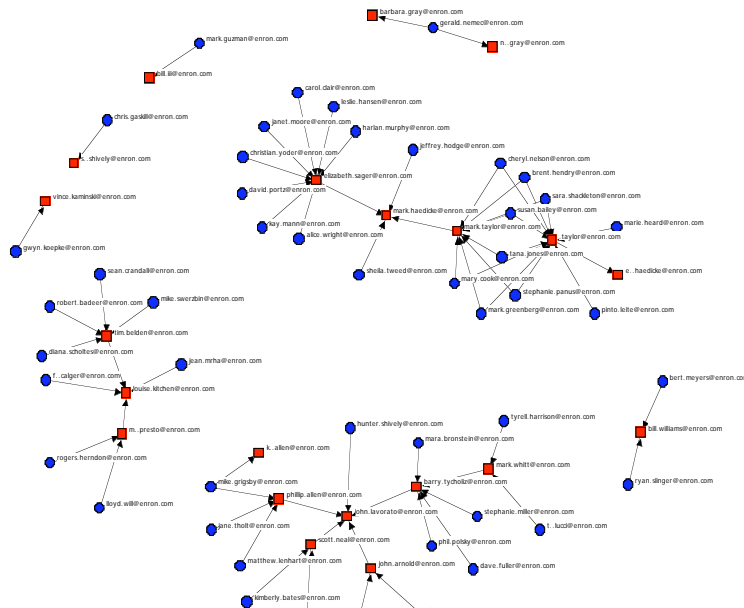
Message Ranking



APL

Evaluation - Enron Manager-Subordinate Relations

**Organizational Ground Truth
January 2000 – November 2001**



Relationship Ranking

Approach	Mean Reciprocal Rank
Content-Based with Attribute Selection	0.719
Content-Based	0.660
Traffic-Based	0.518
Random Selection	0.211
Worst Case	0.141

Evaluation - Enron Manager-Subordinate Relations

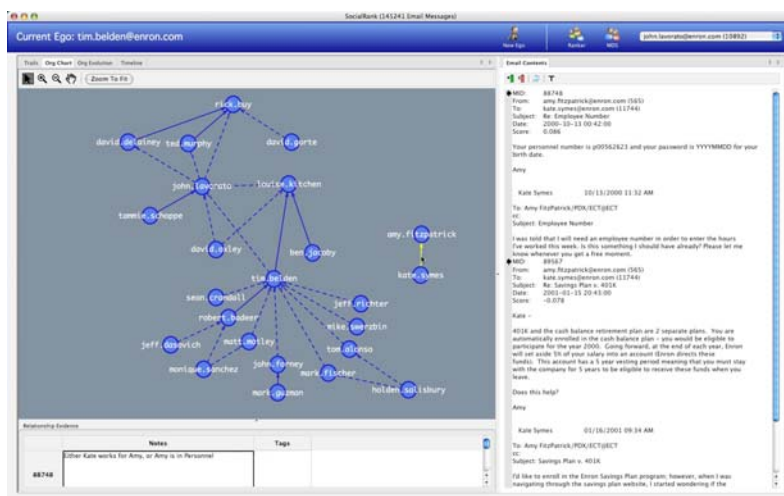
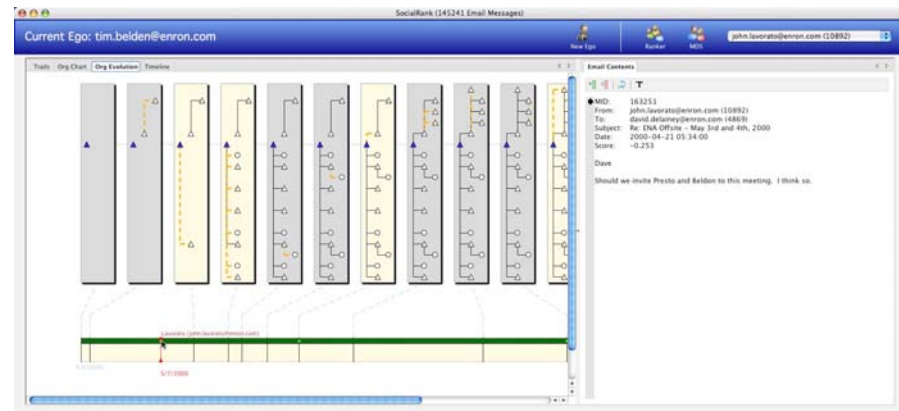
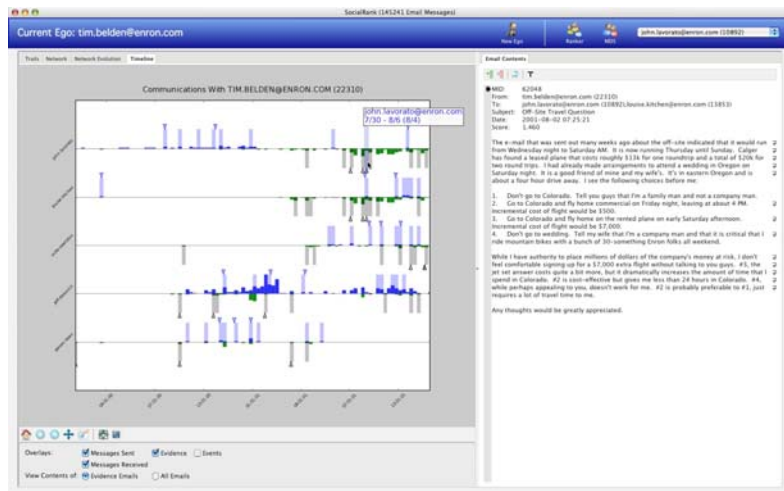
From: tim.belden@enron.com

To: john.lavorato@enron.com,
louise.kitchen@enron.com

Subject: Off-Site Travel Question

The e-mail that was sent out many weeks ago about the off-site indicated that it would run from Wednesday night to Saturday AM. It is now running Thursday until Sunday... I had already made arrangements to attend a wedding in Oregon on Saturday night. It is a good friend of mine and my wife's... While I have authority to place millions of dollars of the company's money at risk, I don't feel comfortable signing up for a \$7,000 extra flight without talking to you guys... Any thoughts would be greatly appreciated.

SocialRank: Interactive Social Relationship Identification



- Relationship and Message Ranking
- Network Knowledge Capture
- Dynamic Network Analysis
- Automatic Report Generation

Current Ego: None

New Ego

Ranker

MDS

Trails Network Network Evolution Timeline

Email Contents

+ - [icon] [icon] T

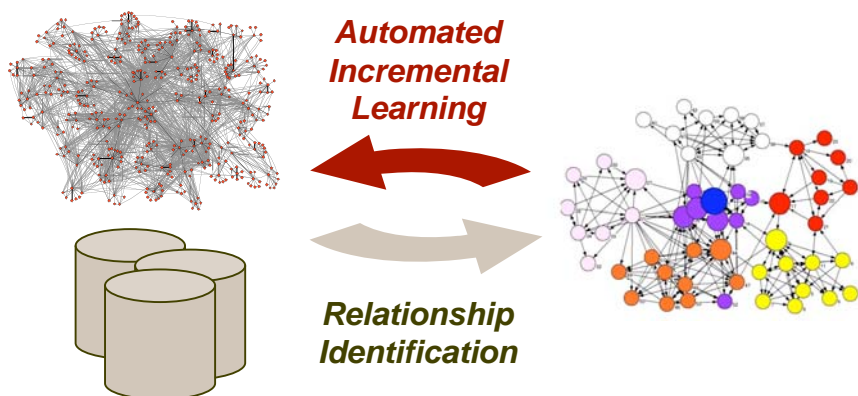


Overlays: ☒ Messages Sent ☒ Evidence ☐ Events
☒ Messages Received
View Contents of: ☒ Evidence Emails ☐ All Emails

Outcomes

- **Developed workflow to accelerate relationship identification**
 - Key phases: discovery, validation, annotation, dissemination
- **Multiple synergies provide significant leverage**
 - Structure and content
 - Supervised and unsupervised
 - Algorithms and visualization
- **Initial foray into dynamic network capture and visualization**

Ongoing Work



- **Automated Incremental Learning**
 - Automated model selection
 - Uncertainty in relationship temporal extent
 - Partially labeled data
- **Characterization and Visualization of Group Dynamics**
 - Understanding changepoints
 - Embedding relationship evolution in group context

