The morphology of a breakdown

How the semantics and mechanics of communication networks from an organization in crises relate

Jana Diesner
Kathleen Carley
Harald Katzmair

FAS.research

CASOS

The Enron Case and Data: Recap

- Founded in 1985 by Kenneth Lay in Houston, Texas
- Business: Gas supplier, Energy broker, Global commodity trading, and things fare beyond
- $$ Success $$!
  - 2001: 7th-largest business organization (by revenue) in the USA, 21,000 employees in more than 40 countries
  - Stock market's darling

Smartest guys in the room

The Enron Case and Data: Recap

- 12/2001: Bankruptcy
- Charged with Illicit accounting and business practices
- Involved Auditor: Arthur Andersen
- Inquiries by SEC (Securities and Exchange Commission) and FERC (Federal Energy Regulatory Commission)
- 2002: FERC releases 620,000 emails from 158 people
    - Real communication from a real organization 10/1998 to 07/2002
    - Rare glimpse into organizational processes, culture, crisis

Fusaro, P.C. & Miller, R.M. (2002). What Went Wrong at Enron. Wiley & Sons: Hoboken, N.J.
Fox, L. (2003). Enron. The Raise and Fall. Wiley & Sons: Hoboken, N.J.

Our Question and Hypotheses

- Question:
  - How do the social structure and the semantics of communication networks change during an organizational crisis?
- Hypotheses:
  - The network segmentation and cohesion of network clusters increase – possibly because people engage in strategic alliances and small groups with trusted others.
  - The semantic entropy of communication networks decreases – possibly because the discourse drifts towards polarized ends of themes and issues.
- Definitions:
  - Social structure: as observed from email headers (relates to 1st hypothesis)
  - Semantics: respective email bodies (relates to 2nd hypothesis)

Why does it matter?

- We assume that communication networks are the place where organizational culture and identity are created through discourse and the circulation of stories.

- We furthermore assume that the semantic and structural mechanism of this process change during crises. But how?

What did People already learn from SNA of the Enron Email Data?

- Patterns of intra-organizational communication changed during crises:
- Interpersonal communication becomes intensified, diversified, and tends to by-pass formal chains or hierarchies of communication more strongly
- However, the connection between the semantics and the morphology of communication networks from organizations in crises is not well understood yet

Workshop on Link Analysis, Counterterrorism and Security, SIAM International Conference on Data Mining 2005. Newport Beach, CA, April 21-23, 2005.

Hard Work First: Data

- Issue: Nodes are e-mail addresses, not people
- We changed that:
  - Identify and add data employee's names, email addresses, positions, business units, supervisors
  - Career history for 676 people, 110 job titles, 13 positions, 8 ranks
- Identified 535 employees (16 to 412 per month) with full name, email addresses (1 to 17 per person, average: 2.2) and career history
- This covers 39.5% of email instances (797,569 out of 2,019,847)
- Effect: Personalization of data set – We now deal with people

Diesner, J., Frantz, T., & Carley, K.M. (2005). Communication Networks from the Enron Email Corpus. Journal of Computational and Mathematical Organization Theory 11, 201-228.

Data: Time Chunks

- From the Enron history selected crucial positive (6 POS), neutral (4 NEU) and negative (10 NEG) events in 2001 that we assume might have impacted Enron's employees

- For each event, extract data for that day plus the 2 following days (results in 20 time chunks that represent 60 days)

Data: Time Chunks Details

| Valence	Time Chunk	Event
	01/22_01/24_POS	Enron reports recurring Annual Earnings
	02/06_02/08_POS	Enron named "most innovative company in America" for 6th consecutive year by Fortune Magazine
	02/12_02/14_NEU	Skilling is named CEO
	03/05_03/07_NEG	Fortune Magazine article "Is ENRON Overpriced?"
	03/22_03/24_POS	Enron reaffirms positive outlook and 2001 EPS targets
	04/17_04/19_POS	Enron announces first quarter profit of $536 million
	06/19_06/21_POS	Enron reiterates confidence in operations and earnings outlook
	07/12_07/14_POS	Enron reports 2nd quarter good earnings, confirms 2001 EPS estimate, announces 2002 target
	08/14_08/16_NEU	Skilling resigns as CEO. Lay replaces him. Sherron Watkins letter to Lay. Lay discusses Skilling's departure with employees.
	08/21_08/23_NEU	Lay emails employees: "one of my highest priorities is to restore investor confidence in Enron. This should result in a significantly higher stock price."
	09/26_09/28_NEU	Lay tells employees: Enron's accounting practices "legal and totally appropriate," Enron stock is "an incredible bargain," "the third quarter is looking great"
	10/12_10/14_NEG	Arthur Andersen lawyer Nancy Temple emails an Andersen partner reminding him of the Andersen document retention and destruction policy.
	10/16_10/18_NEG	Enron announces a third quarter loss of $618 million. The Enron 401(k) retirement plan is frozen for administrative changes.
	10/22_10/24_NEG	Enron announces SEC inquiry into Enron. The Arthur Andersen gives order to shred documents.
	10/25_10/27_NEG	Enron tells all employees and Arthur Andersen to preserve pertinent documents. Lay phone call to Alan Greenspan (chairman of the Federal Reserve) about Enron.
	10/28_10/30_NEG	Enron asks treasury Under-secretary to ask Enron's creditors to extend its credit and help with upcoming credit rating review. Unsuccessfully.
	10/31_11/02_NEG	Enron announces that the SEC inquiry is now a formal investigation.
	11/08_11/10_NEG	Enron announces it overstated profits by $586 million over five years. Arthur Andersen stops the shredding.
	11/19_11/21_NEG	Enron announces payment of a $690 million due to decrease credit rating.
	12/02_12/04_NEG	Enron files for bankruptcy.

Results: SNA measures

- Measure that changed significantly during crisis:

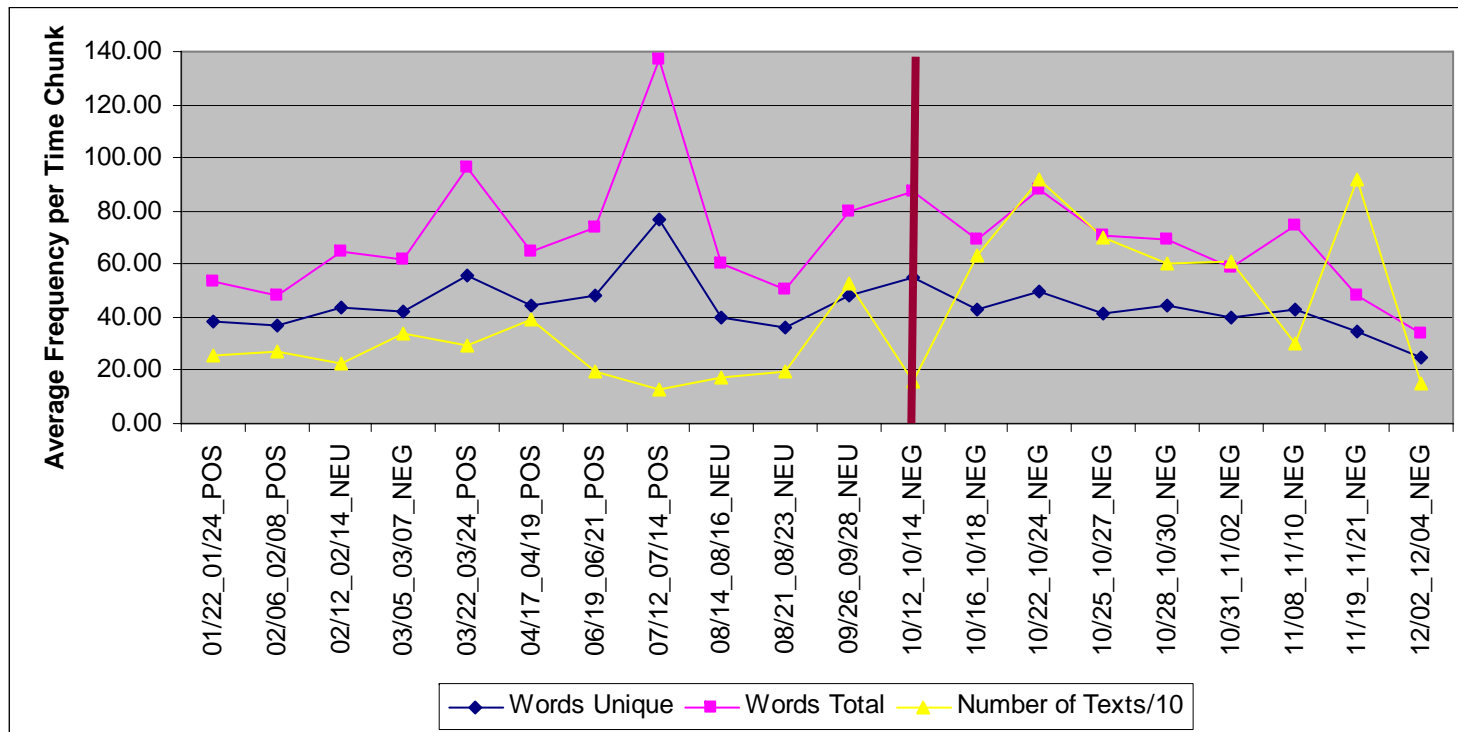| Increase	Decrease
Distance (Shortest Path), Average	Component, Count-Weak
Centrality-Bonacich Power, Average	Hierarchy
Clique Count, Average	Network Centralization-Total Degree
Connectedness	Transitivity
Number of Nodes (here, equals the diameter)	Redundancy
Edge Count	Fragmentation
Reciprocity	Density*
Upper Boundedness	Centrality*
	Inverse Closeness*

* very low values for all time chunks

- No change: Triads

SNA Results: Sampler
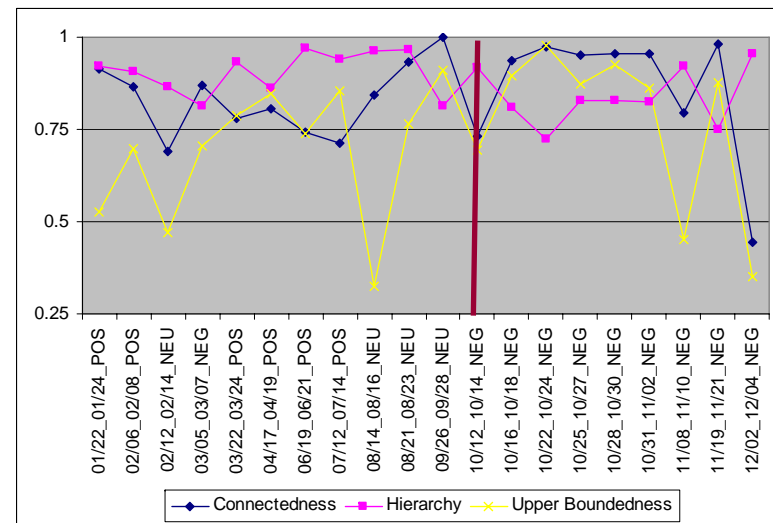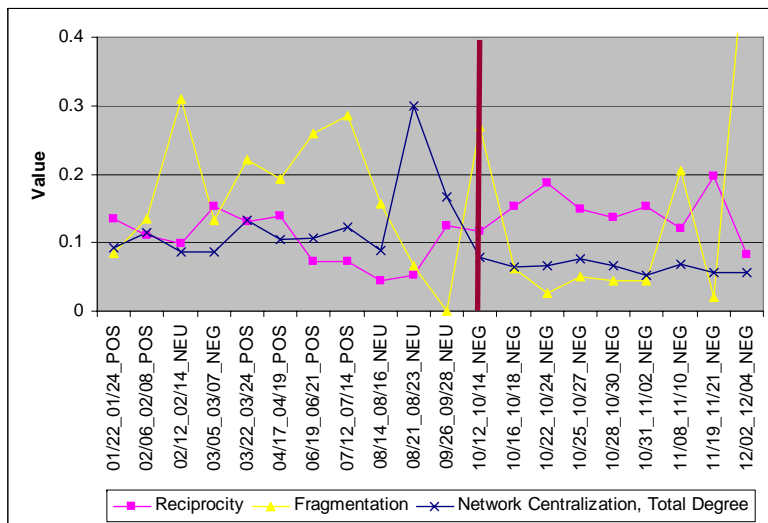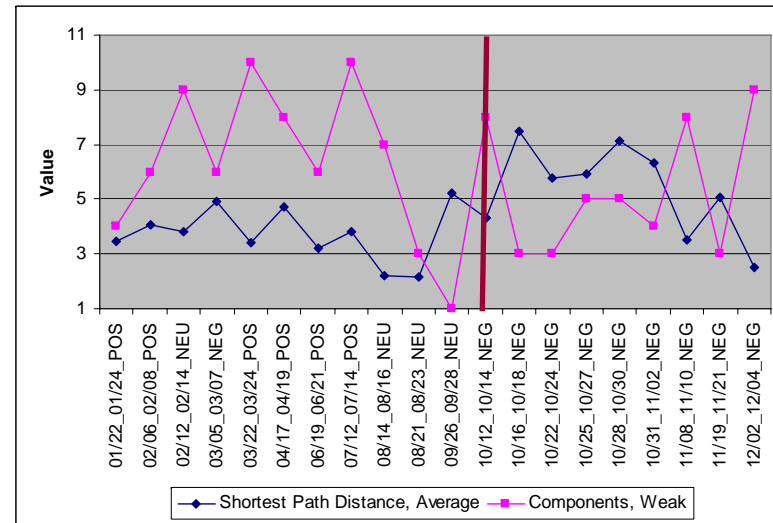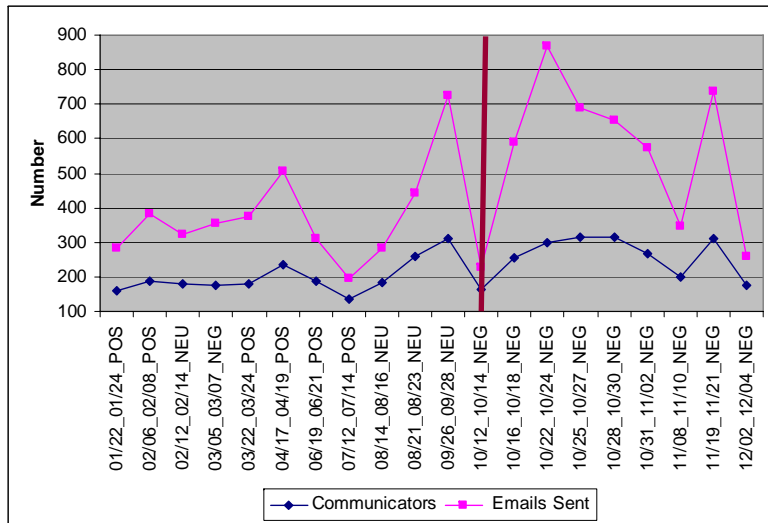
Results: Semantics

- As the crisis emerges, on average, less words and less distinct words are used (redundancy increases) while sheer amount of data exchanged increases
- People say more with less words and less diversity

Results: Semantics

- Formalize the idea of redundancy and diversity: Information Entropy (H) of an email (X)

$$H(X) = -\sum_{x \in X} p(x) \log_2 p(x)$$

where probability p of a word x = ratio of total frequency of x to length (total number of words) of an email

W. Weaver & C. E. Shannon, The Mathematical Theory of Communication, Urbana, Illinois: University of Illinois Press, 1949
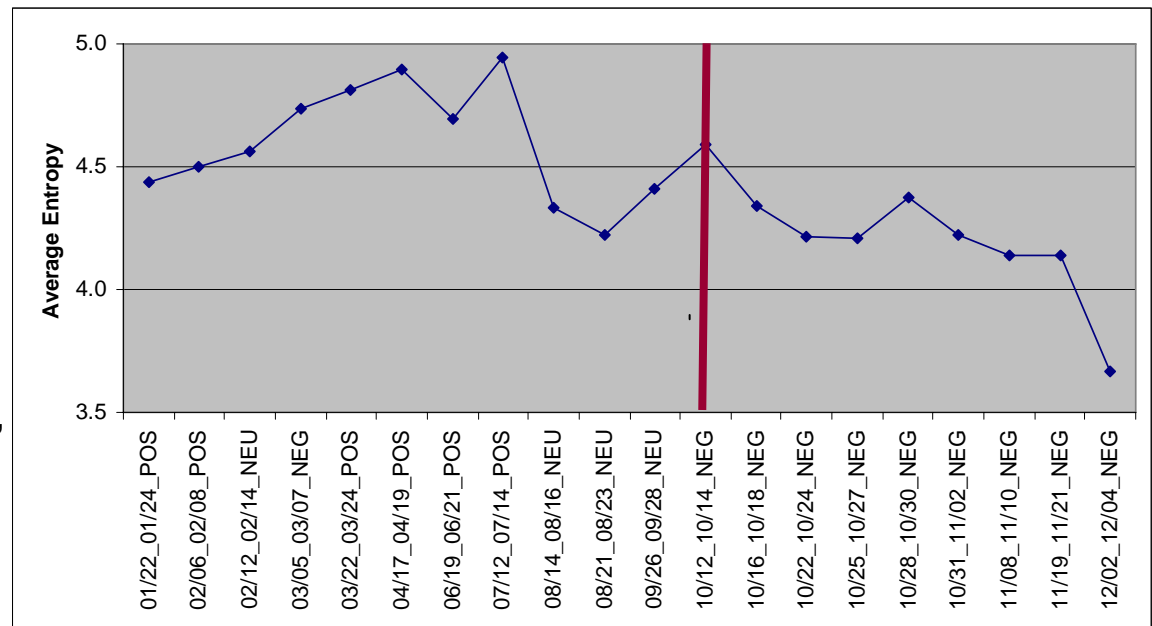
Dyad Selection

- From personalized database extract those dyads (pairs of people) that communicated in at least more than half of the time chunks
  - Why? To cancel out individual biases
  - 1 pair of agents having messages in 19 time intervals,
    4 pairs in 18 and 14 time intervals,
    5 pairs in 16 and 15 time intervals
  - Dasovich origin for 8 pairs, Steffes for 7 pairs, Mara for 2 pairs (Dasovich and Steffes) (important for the next slides)
  - 19 pairs (11 distinct people), non symmetric, but 2 symmetric pairs
  - Emails across folders can be redundant (this happens when X sends an email to Y and Z)
  - Text analysis based on individual pairs -> no redundant email per dyad

Dyad Selection

- Alan Comnes, Timothy Belden, Jeff Dasovich, Susan Mara, Harry Kingerski: Director
- Paul Kaufman, Steven Kean: VP
- Richard Shapiro, James Steffes: VP governmental affairs
- Linda Robertson: former Clinton Treasury official, governmental affairs executives
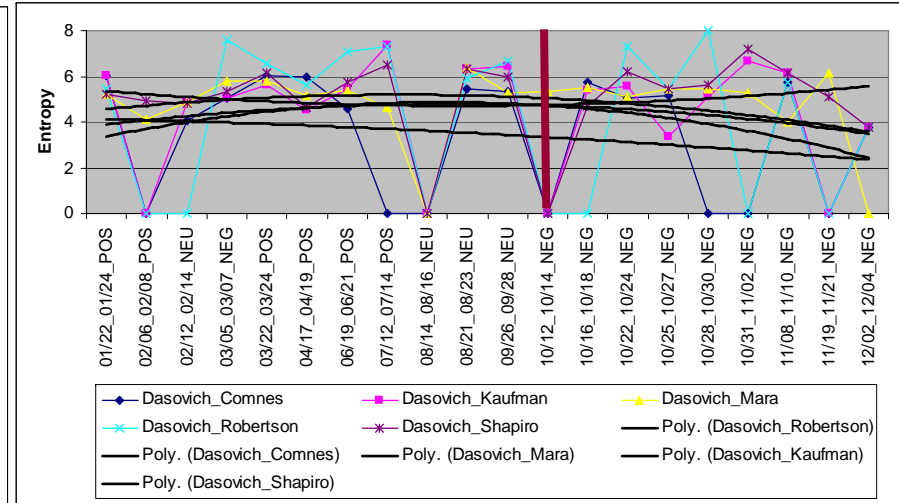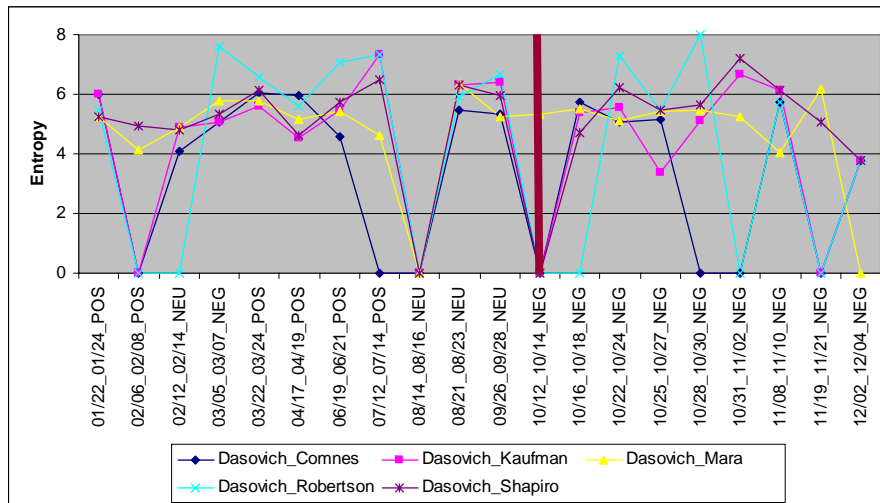- Christi Nicolay: Lawyer
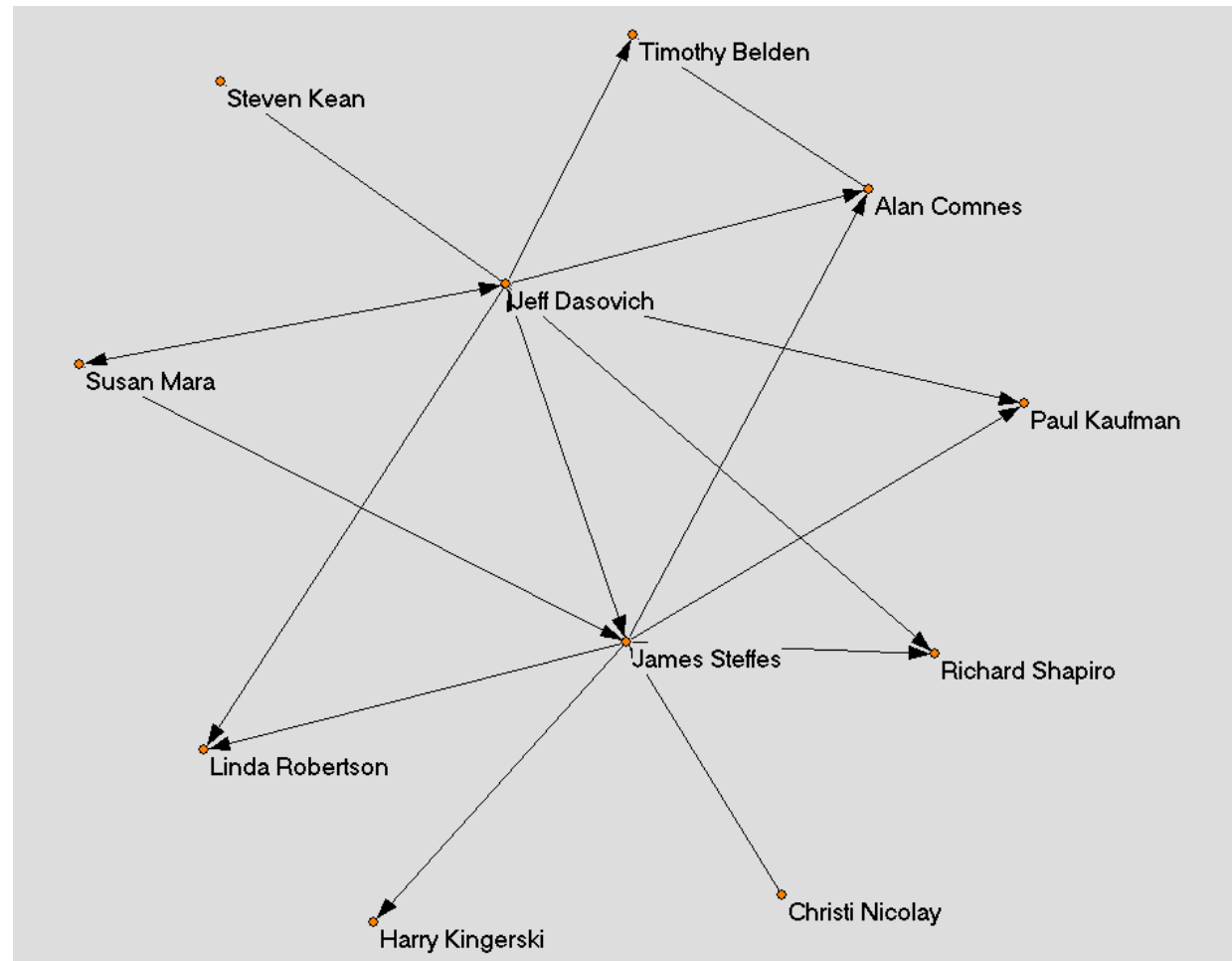
Dyads: Results
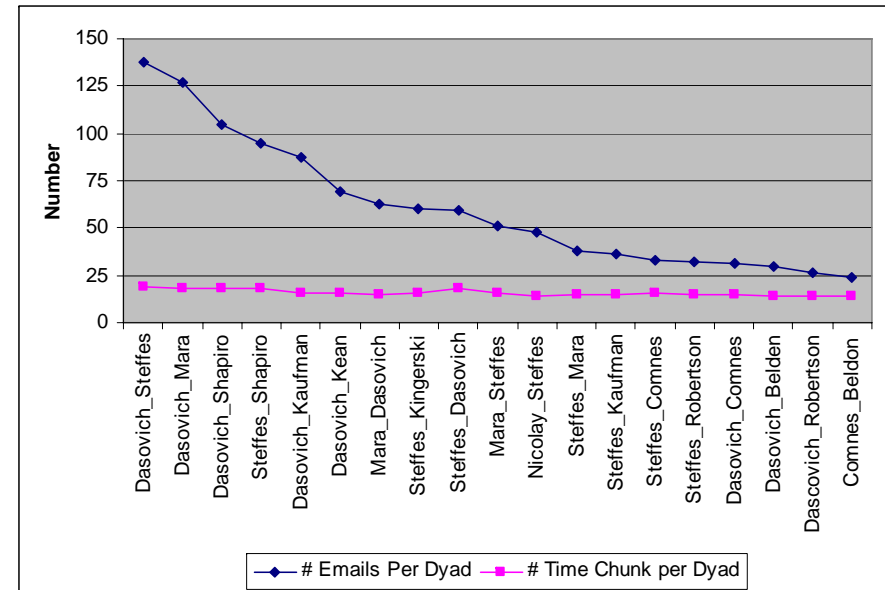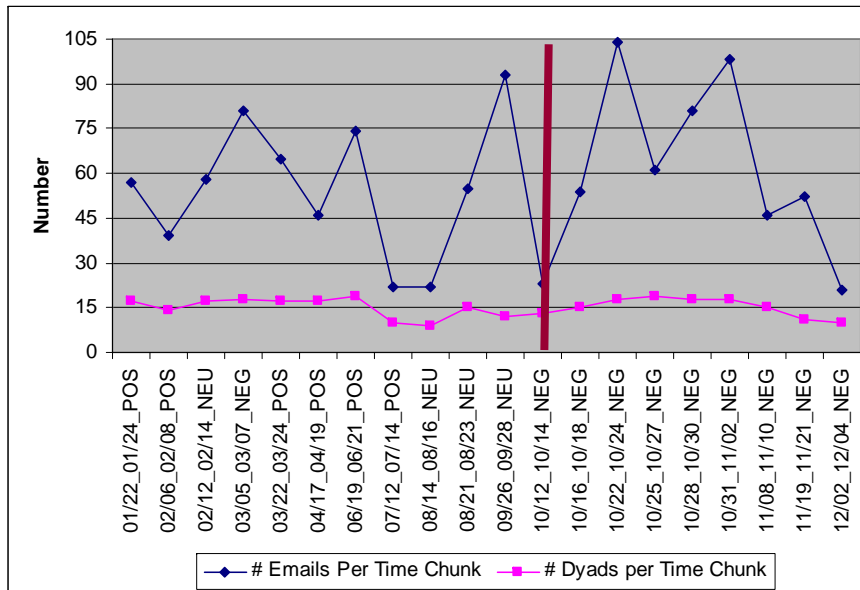
- **Trends from analysis of full network resembled (more people communicate more)**
- **Good coverage of dyads across time chunks and even more so vice versa**
- **Very different frequency distribution of emails across dyads**

Dyads: More detailed Results

- What looks like a mess is the communication entropy from pairs where Dasovich is the origin for all of them
- Yet it is typical for what we learned from studying the individual or dyadic level, for that helped e.g.
  - Graph at the right: Trends per dyad across time chunks (polynomial (2) fitting)

Dyads: More detailed Results
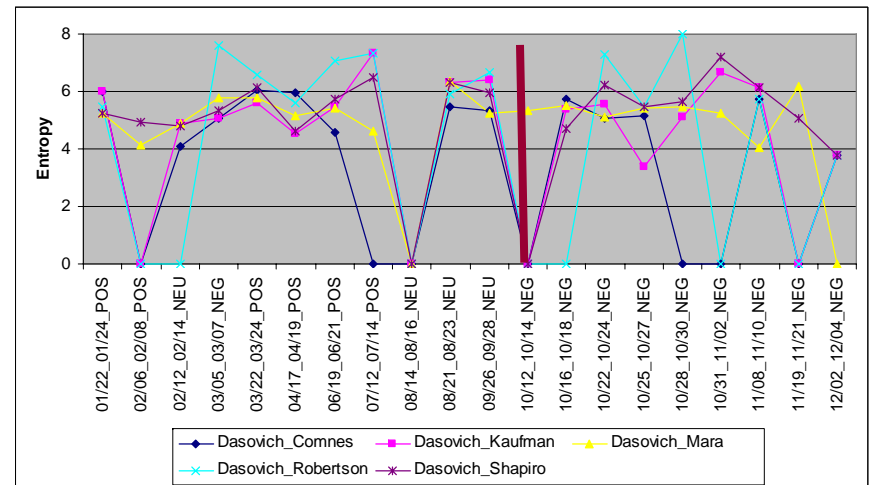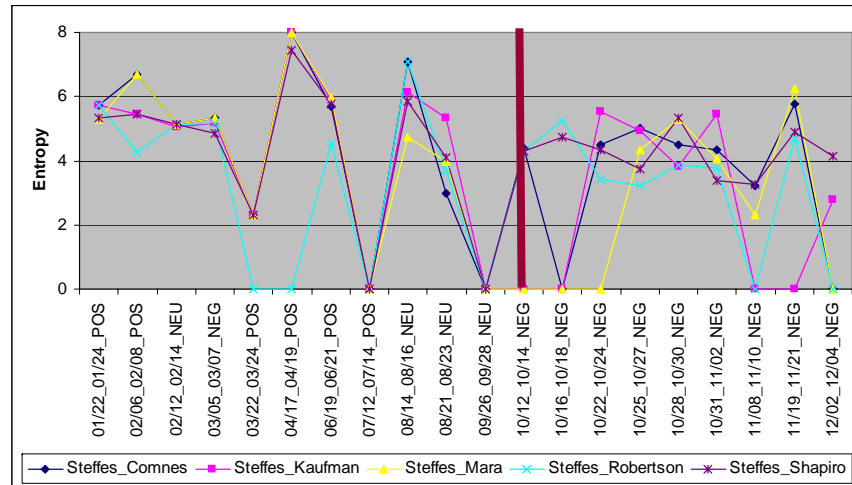
Until crisis: Sender's emails show similar patterns (ups and downs) in unique and total words, redundancy and entropy with varying frequencies UNTIL crisis

From emergence of crisis on, individuals differ their communication strongly depending on who they address (e.g. Dasovich increases entropy only to Mara and lowers it to the rest)

Dyads: More Detailed Results

- Higher variation (before, averaging smoothed that)
- Trends: mainly same as globally for MOST target people EXCEPT for a few (e.g. Dasovich to Mara)
- Previous Slide: Results indicate that in crisis people not only change their overall way of expressing themselves, but also their communicative behavior from individual to individual

Thank you.
Q&A