Analysis of Latent Relationships in Semantic Graphs using DEDICOM

Brett Bader*, Richard Harshman** & Tamara Kolda*
  *Sandia National Laboratories
  **University of Western Ontario

Workshop for Algorithms on Modern Massive Data Sets
June 24, 2006
Common Graph Analysis Technique

Best rank-$k$ matrix filters out noise and captures “latent” information, which improves certain data mining tasks.

For example:
Web search - HITS (Kleinberg, 1998)

\[ A_k = U_k \Sigma_k V_k^T = \sum_{i=1}^{k} \sigma_i u_i v_i^T \]

But we may have ignored critical information by not considering edge metadata!
Semantic Graphs

- Different types of edges

- Examples
  - WWW (anchor text)
  - Subway map [thanks Orly!]
  - Email communications (time stamp, to/cc)
New Paradigm: “Multidimensional Data Mining”

Third dimension offers more explanatory power: uncovers new latent information and reveals subtle relationships.
Objective

Use DEDICOM to analyze a semantic graph of email communications changing over time

3-way DEDICOM
DEDICOM

- DEcomposition into DIrectional COMponents

- Introduced in 1978 by Harshman

- Past applications
  - Study asymmetries in telephone calls among cities
  - Marketing research
    - car switching: car owners and what they buy next
    - free associations of words
      - words to describe hair in advertising shampoo:
        "body" evokes "fullness" more often than "fullness"
        evokes "body"
      - Asymmetric measures of world trade (import/export)

- Variations
  - Three-way DEDICOM
  - Constrained DEDICOM
DEDICOM Models & Algorithms

\[ X = A R A^T \]

- Generalized Takane method (Takane, 1985; Kiers et al., 1990)
- New algorithm

\[ X = A R A^T \]

- Kiers’ method (Kiers, 1993)
- New algorithm

All are “alternating” algorithms
Mathematical Notation

- Scalars $a$
- Vectors $\mathbf{a}$
- Matrices $\mathbf{A}$
- Tensors (3-way array): $\mathcal{D}$, $\mathcal{X}$
  - Frontal slices of $\mathcal{X}$: $X_i$

- Special symbols
  - Kronecker product
    \[
    \mathbf{A} \otimes \mathbf{B} = \begin{bmatrix}
    a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\
    \vdots & \ddots & \vdots \\
    a_{m1}\mathbf{B} & \cdots & a_{mn}\mathbf{B}
    \end{bmatrix}
    \]
  - Hadamard product (elementwise)
    \[
    \mathbf{A} \ast \mathbf{B} = \begin{bmatrix}
    a_{11}b_{11} & \cdots & a_{1n}b_{1n} \\
    \vdots & \ddots & \vdots \\
    a_{m1}b_{m1} & \cdots & a_{mn}b_{mn}
    \end{bmatrix}
    \]
Two-way DEDICOM

\[ X = A R A^T + E \]
\[ X \approx A R A^T \]

\[ \min_{A,R} \left\| X - A R A^T \right\|_F^2 \]
\[ \text{s.t. } A \text{ orthogonal} \]

- \( A \) \((n \times p)\) is an orthogonal matrix of loadings or weights
- \( R \) \((p \times p)\) is a dense matrix that captures asymmetric relationships

- Decomposition is not unique
  - \( A \) can be transformed with no loss of fit to the data
  - Nonsingular transformation \( Q \):
    \[ A R A^T = (AQ)(Q^{-1}RQ^{-T})(AQ)^T \]
  - Usually “fix” \( A \) with some standard rotation (e.g., VARIMAX)
New Algorithm

Solving for $A$:

Stack data and model “side by side” in a single equation

$\begin{pmatrix} X & X^T \end{pmatrix} = \begin{pmatrix} ARA^T & AR^TA^T \end{pmatrix}$

$\begin{align*}
  &= A \begin{pmatrix} R & R^T \end{pmatrix} \begin{pmatrix} A^T & 0 \\ 0 & A^T \end{pmatrix} \\
normalsize
  &= A \begin{pmatrix} R & R^T \end{pmatrix} \begin{pmatrix} A^T & 0 \\ 0 & A^T \end{pmatrix} \end{align*}$

...and solve least-squares problem:

$\min_A \left\| Y - AZ^T \right\|_F^2$

$A_{new} \leftarrow (X \hspace{1em} X^T) \begin{pmatrix} R & R^T \end{pmatrix} \begin{pmatrix} A^T & 0 \\ 0 & A^T \end{pmatrix} \dagger$

or

$A_{new} = (XAR^T + X^TAR) \left( R(A^T A)R^T + R^T(A^T A)R \right)^{-1}$.

Solving for $R$:

$R_{new} = A^\dagger X(A^T)^\dagger$
Three-way DEDICOM

\[ X_i = AD_iRD_iA^T + E_i \quad \text{for } i = 1, \ldots, m, \]

- \( A \) \((n \times p)\) is a matrix of loadings or weights (not necessarily orthogonal)
- \( R \) \((p \times p)\) is a dense matrix that captures asymmetric relationships
- \( D \) \((p \times p \times m)\) is a tensor with diagonal frontal slices giving the weights of the columns of \( A \) for each slice in third mode

- *Unique* solution with enough slices of \( X \) with sufficient variation
  - i.e., no rotation of \( A \) possible
  - greater confidence in interpretation of results
New Algorithm - Updating $A$

$$\min_{A,R,D} \sum_{i=1}^{m} \left\| X_i - AD_iRD_iA^T \right\|^2_F$$

Solving for $A$:

$$(X_1 \quad X_1^T \quad \cdots \quad X_m \quad X_m^T) = A \left( D_1RD_1 \quad D_1R^TD_1 \quad \cdots \quad D_mRD_m \quad D_mR^TD_m \right) (I_{2m} \otimes A^T)$$

$$Y = \begin{bmatrix} A \end{bmatrix} \begin{bmatrix} Z^T \end{bmatrix}$$

$$A = YZ(Z^TZ)^{-1}$$

$$A = \left[ \sum_{i=1}^{m} \left( X_iAD_iR^TD_i + X_i^TAD_iRD_i \right) \right] \left[ \sum_{i=1}^{m} (B_i + C_i) \right]^{-1}$$

where $B_i \equiv D_iRD_i(A^TA)D_iR^TD_i$,  
$C_i \equiv D_iR^TD_i(A^TA)D_iRD_i$. 

\[\text{Sandia National Laboratories}\]
New Algorithm - Updating D

\[
\min_{D_i} \left\| X_i - AD_iRD_iA^T \right\|_F^2
\]

Solving for D:

Use Newton’s method to solve the optimization problem for \( d = \text{diag}(D_i) \)

\[
d_{new} = d - H^{-1}g
\]

Gradient: \( g_k = -\sum_{i,j} \left[ 2(X - ADRDA^T) \ast (ADr_k a_k^T + a_k r_{k,:} D^T) \right]_{i,j} \)

Hessian: \( h_{st} = -2 \sum_{i,j} \left[ (X - ADRDA^T) \ast (a_s r_{st} a_t^T + a_t r_{ts} a_s^T) \right. \]

\[
- (ADr_s a_s^T + a_s r_s : D A^T) \ast (ADr_t a_t^T + a_t r_t : D A^T) \left. \right]_{i,j}
\]

Use compression

QR factorization: \( A = Q\tilde{A} \),

\[
\min_{D_i} \left\| Q^T X_i Q - \tilde{A}D_iR\tilde{D}_i\tilde{A}^T \right\|_F^2 \quad \text{Smaller problem (p x p)}
\]
Our Algorithm - Updating $R$

$$\min_R \sum_{i=1}^{m} \| X_i - A D_i R D_i A^T \|_F^2$$

Solving for $R$:

Use the approach in (Kiers, 1993)

$$\min_R f(R) = \left\| \begin{pmatrix} \text{Vec}(X_1) \\ \vdots \\ \text{Vec}(X_m) \end{pmatrix} - \begin{pmatrix} A D_1 \otimes A D_1 \\ \vdots \\ A D_m \otimes A D_m \end{pmatrix} \text{Vec}(R) \right\|$$

$$\text{Vec}(R) = \left( \sum_{i=1}^{m} (D_i A^T A D_i) \otimes (D_i A^T A D_i) \right)^{-1} \sum_{i=1}^{m} \text{Vec}(D_i A^T X_i A D_i)$$
Algorithm Costs

Updating $\mathbf{A}$ is most expensive part

Dominant costs:

$$
\begin{align*}
\mathcal{O}(p^2n) & \quad \left< \begin{array}{c}
Q^T \mathbf{X}_i \mathbf{Q} \\
\mathbf{X}_i \mathbf{A} \mathbf{R}^T \\
\mathbf{X}_i^T \mathbf{A} \mathbf{R} \\
\mathbf{A}^T \mathbf{A}
\end{array} \right> \\
& \quad \text{QR factorization of } \mathbf{A}
\end{align*}
$$

linear in nnz of $\mathbf{X}_i$
• Links consist of email communications

• What can we learn about this network strictly from their communication patterns? (Social network analysis)
Enron Corp.

- U.S. corporation involved with creating energy markets
  - 7th largest by revenue
- EnronOnline: e-trading business
  - natural gas
  - electric power

Investigations
- U.S. Federal Energy Regulatory Commission (FERC)
  - energy market manipulation
  - involved energy traders
- U.S. Securities and Exchange Commission (SEC)
  - accounting fraud
  - insider trading
Enron Email Data

- FERC collected email of ~150 employees as evidence
  - Included emails saved in inbox, sent items, deleted items, and all other folders

- Released to the public in 2002 by FERC as part of their investigation
  - To/from, date, subject, body
  - Attachments and some names/emails removed
  - Approx. 500,000 email messages
We used a smaller data set prepared by Priebe et al. 34,427 emails among 184 employees over 44 months

- Limited information on the 184 employees
- No org chart
DEDICOM Experiment

- Aggregate communications
  - Sparse matrix of size 184 x 184 (3007 nonzeros)

- Time series of communication graphs
  - Sparse tensor of size 184 x 184 x 44 (9838 nonzeros)

- Weighted adjacency matrix
  - scaling: x number of messages scaled by log(x)+1
  - other common choices give similar results

- Models:
  - SVD
  - 2-way DEDICOM
  - 3-way DEDICOM
Social Network Analysis

Communication graph among employees over all times

- Description of employees by their roles
- Aggregate communication patterns among roles

Edition of Communication (DEDICOM)

Adjacency matrix
DEDICOM Results

![Patterns of communications in R matrix diagram](image)

<table>
<thead>
<tr>
<th>EMPLOYEE</th>
<th>DEDICOM Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>J. Lavorato - CEO, Enron America</td>
<td>0.41</td>
</tr>
<tr>
<td>L. Kitchen - President, Enron Online</td>
<td>0.26</td>
</tr>
<tr>
<td>M. Grigsby - Director, West Desk Gas Trading</td>
<td>0.22</td>
</tr>
<tr>
<td>D. Delainey - CEO, ENA and Enron Energy Services</td>
<td>0.20</td>
</tr>
<tr>
<td>G. Whalley - President,</td>
<td>0.17</td>
</tr>
<tr>
<td>L. Taylor - Executive Assistant to Greg Whalley,</td>
<td>0.17</td>
</tr>
<tr>
<td>T. Jones - Employee, Financial Trading Group (ENA Legal)</td>
<td>-0.12</td>
</tr>
<tr>
<td>M. Taylor - Manager, Financial Trading Group ENA Legal</td>
<td>-0.10</td>
</tr>
<tr>
<td>S. Shackleton - Employee, ENA Legal</td>
<td>-0.13</td>
</tr>
<tr>
<td>S. Panus - Senior Legal Specialist, ENA Legal</td>
<td>-0.11</td>
</tr>
<tr>
<td>M. Heard - Senior Legal Specialist, ENA Legal</td>
<td>-0.10</td>
</tr>
<tr>
<td>E. Sager - VP and Asst Legal Counsel, ENA Legal</td>
<td>-0.01</td>
</tr>
<tr>
<td>S. Corman - VP, Regulatory Affairs</td>
<td>-0.04</td>
</tr>
<tr>
<td>K. Watson - Employee, Transwestern Pipeline Company (ETS)</td>
<td>-0.08</td>
</tr>
<tr>
<td>L. Donoho - Employee, Transwestern Pipeline Company (ETS)</td>
<td>-0.08</td>
</tr>
<tr>
<td>D. Fossum - VP, Transwestern Pipeline Company (ETS)?</td>
<td>-0.06</td>
</tr>
<tr>
<td>M. Lokay - Admin. Asst., Transwestern Pipeline Company (ETS)</td>
<td>-0.07</td>
</tr>
<tr>
<td>K. Hyatt - Director, Asset Development TW Pipeline Co. (ETS)</td>
<td>-0.06</td>
</tr>
<tr>
<td>R. Hayslett - VP, Also CFO and Treasurer</td>
<td>-0.04</td>
</tr>
</tbody>
</table>

Some employees have dual roles

Pattern of communications in R matrix

R matrix

<table>
<thead>
<tr>
<th></th>
<th>15</th>
<th>11</th>
<th>5</th>
<th>6.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Legal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pipeline</td>
<td>6.7</td>
<td></td>
<td>6.7</td>
<td></td>
</tr>
<tr>
<td>Execs</td>
<td>11</td>
<td>5</td>
<td>6.7</td>
<td></td>
</tr>
</tbody>
</table>

Identify shared characteristics to label group
Social Network Analysis

Communication graph among employees over all times

“Hubs” and “authorities” for different roles

SVD
## DEDICOM & SVD Results

<table>
<thead>
<tr>
<th>Employee</th>
<th>DEDICOM Solution</th>
<th>U (hubs)</th>
<th>V (authorities)</th>
</tr>
</thead>
<tbody>
<tr>
<td>J. Lavorato - CEO, Enron America</td>
<td>0.41 0.07 0.04</td>
<td>0.30 -0.07 -0.21</td>
<td>0.31 -0.09 -0.07</td>
</tr>
<tr>
<td>L. Kitchen - President, Enron Online</td>
<td>0.26 0.21 0.04</td>
<td>0.31 0.07 -0.05</td>
<td>0.29 0.02 0.04</td>
</tr>
<tr>
<td>M. Grigsby - Director, West Desk Gas Trading</td>
<td>0.22 -0.01 -0.01</td>
<td>0.16 -0.09 -0.33</td>
<td>0.14 -0.06 -0.20</td>
</tr>
<tr>
<td>D. Delainey - CEO, ENA and Enron Energy Services</td>
<td>0.20 0.06 0.06</td>
<td>0.20 -0.05 -0.00</td>
<td>0.20 -0.05 0.03</td>
</tr>
<tr>
<td>G. Whalley - President,</td>
<td>0.17 0.05 0.04</td>
<td>0.08 -0.02 -0.02</td>
<td>0.24 -0.07 0.02</td>
</tr>
<tr>
<td>L. Taylor - Executive Assistant to Greg Whalley,</td>
<td>0.17 0.06 0.03</td>
<td>0.24 -0.05 -0.08</td>
<td>0.09 -0.01 -0.02</td>
</tr>
<tr>
<td>T. Jones - Employee, Financial Trading Group (ENA Legal)</td>
<td>-0.12 0.38 -0.02</td>
<td>0.17 0.36 0.13</td>
<td>0.10 0.24 0.10</td>
</tr>
<tr>
<td>M. Taylor - Manager, Financial Trading Group ENA Legal</td>
<td>-0.10 0.35 -0.01</td>
<td>0.13 0.27 0.13</td>
<td>0.13 0.26 0.12</td>
</tr>
<tr>
<td>S. Shackleton - Employee, ENA Legal</td>
<td>-0.13 0.31 -0.02</td>
<td>0.08 0.26 0.10</td>
<td>0.08 0.26 0.10</td>
</tr>
<tr>
<td>S. Panus - Senior Legal Specialist, ENA Legal</td>
<td>-0.11 0.26 -0.02</td>
<td>0.09 0.27 0.10</td>
<td>0.05 0.20 0.08</td>
</tr>
<tr>
<td>M. Heard - Senior Legal Specialist, ENA Legal</td>
<td>-0.10 0.24 -0.02</td>
<td>0.06 0.20 0.09</td>
<td>0.08 0.22 0.09</td>
</tr>
<tr>
<td>E. Sager - VP and Asst Legal Counsel, ENA Legal</td>
<td>-0.01 0.24 0.02</td>
<td>0.12 0.13 0.10</td>
<td>0.15 0.21 0.12</td>
</tr>
<tr>
<td>S. Corman - VP, Regulatory Affairs</td>
<td>-0.04 -0.01 0.33</td>
<td>0.08 -0.18 0.22</td>
<td>0.07 -0.18 0.21</td>
</tr>
<tr>
<td>K. Watson - Employee, Transwestern Pipeline Company (ETS)</td>
<td>-0.08 -0.03 0.32</td>
<td>0.03 -0.16 0.19</td>
<td>0.04 -0.18 0.22</td>
</tr>
<tr>
<td>L. Donoho - Employee, Transwestern Pipeline Company (ETS)</td>
<td>-0.08 -0.03 0.30</td>
<td>0.03 -0.16 0.18</td>
<td>0.03 -0.17 0.20</td>
</tr>
<tr>
<td>D. Fossum - VP, Transwestern Pipeline Company (ETS)?</td>
<td>-0.06 -0.00 0.30</td>
<td>0.07 -0.18 0.23</td>
<td>0.05 -0.13 0.16</td>
</tr>
<tr>
<td>M. Lokay - Admin. Asst., Transwestern Pipeline Company (ETS)</td>
<td>-0.07 -0.02 0.28</td>
<td>0.03 -0.14 0.17</td>
<td>0.04 -0.17 0.20</td>
</tr>
<tr>
<td>K. Hyatt - Director, Asset Development TW Pipeline Co. (ETS)</td>
<td>-0.06 -0.02 0.25</td>
<td>0.03 -0.13 0.17</td>
<td>0.04 -0.14 0.17</td>
</tr>
<tr>
<td>R. Hayslett - VP, Also CFO and Treasurer</td>
<td>-0.04 -0.01 0.23</td>
<td>0.04 -0.13 0.16</td>
<td>0.05 -0.14 0.16</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>R matrix / singular values</th>
<th>70.3 11.6 6.7</th>
<th>86.3</th>
<th>86.3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>15.4 68.2 5.0</td>
<td>54.1</td>
<td>54.1</td>
</tr>
<tr>
<td></td>
<td>9.9 6.7 59.5</td>
<td>52.6</td>
<td>52.6</td>
</tr>
</tbody>
</table>

### SVD: Hubs and Authorities in U and V

Roles more difficult to identify in singular vectors

No patterns of communication
Temporal Social Network Analysis

Time series of communication graphs among employees

- Unique description of employees by their roles
- Aggregate communication patterns among roles
- Behavior over time
# Roles of Employees

<table>
<thead>
<tr>
<th>Employee</th>
<th>Legal</th>
<th>Gov’t affairs</th>
<th>Trade execs</th>
<th>Pipeline</th>
</tr>
</thead>
<tbody>
<tr>
<td>T. Jones - Employee, Financial Trading Group (ENA Legal)</td>
<td>0.64</td>
<td>-0.01</td>
<td>0.02</td>
<td>-0.00</td>
</tr>
<tr>
<td>S. Shackleton - Employee, ENA Legal</td>
<td>0.45</td>
<td>-0.00</td>
<td>-0.01</td>
<td>-0.00</td>
</tr>
<tr>
<td>M. Taylor - Manager, Financial Trading Group ENA Legal</td>
<td>0.37</td>
<td>0.01</td>
<td>0.02</td>
<td>-0.00</td>
</tr>
<tr>
<td>S. Bailey - Legal Assistant, ENA Legal</td>
<td>0.26</td>
<td>-0.00</td>
<td>-0.01</td>
<td>-0.00</td>
</tr>
<tr>
<td>S. Panus - Senior Legal Specialist, ENA Legal</td>
<td>0.26</td>
<td>-0.00</td>
<td>-0.00</td>
<td>-0.00</td>
</tr>
<tr>
<td>M. Heard - Senior Legal Specialist, ENA Legal</td>
<td>0.23</td>
<td>-0.00</td>
<td>0.00</td>
<td>-0.00</td>
</tr>
<tr>
<td>J. Hodge - Asst General Counsel, ENA Legal</td>
<td>0.13</td>
<td>0.03</td>
<td>0.01</td>
<td>-0.00</td>
</tr>
<tr>
<td>L. Kitchen - President, Enron Online</td>
<td>0.11</td>
<td>-0.09</td>
<td>0.53</td>
<td>0.00</td>
</tr>
<tr>
<td>S. Dickson - Employee, ENA Legal</td>
<td>0.09</td>
<td>-0.00</td>
<td>0.00</td>
<td>-0.00</td>
</tr>
<tr>
<td>E. Sager - VP and Asst Legal Counsel, ENA Legal</td>
<td>0.08</td>
<td>0.02</td>
<td>0.07</td>
<td>-0.00</td>
</tr>
<tr>
<td>J. Dasovich - Employee, Government Relationship Executive</td>
<td>-0.01</td>
<td>0.58</td>
<td>0.06</td>
<td>0.01</td>
</tr>
<tr>
<td>J. Steffes - VP, Government Affairs</td>
<td>0.00</td>
<td>0.53</td>
<td>-0.06</td>
<td>-0.01</td>
</tr>
<tr>
<td>R. Shapiro - VP, Regulatory Affairs</td>
<td>-0.00</td>
<td>0.40</td>
<td>0.10</td>
<td>-0.00</td>
</tr>
<tr>
<td>S. Kean - VP, Chief of Staff</td>
<td>-0.00</td>
<td>0.37</td>
<td>-0.04</td>
<td>-0.00</td>
</tr>
<tr>
<td>R. Sanders - VP, Enron Wholesale Services</td>
<td>0.03</td>
<td>0.16</td>
<td>-0.01</td>
<td>-0.00</td>
</tr>
<tr>
<td>D. Delainey - CEO, ENA and Enron Energy Services</td>
<td>0.01</td>
<td>0.09</td>
<td>0.09</td>
<td>-0.00</td>
</tr>
<tr>
<td>S. Cormen - VP, Regulatory Affairs</td>
<td>-0.00</td>
<td>0.08</td>
<td>-0.00</td>
<td>0.20</td>
</tr>
<tr>
<td>M. Carson - Employee, Corporate and Environmental Policy</td>
<td>-0.00</td>
<td>0.08</td>
<td>-0.02</td>
<td>-0.00</td>
</tr>
<tr>
<td>S. Scott - Employee, Transwestern Pipeline Company (ETS)</td>
<td>-0.00</td>
<td>0.08</td>
<td>-0.00</td>
<td>0.04</td>
</tr>
<tr>
<td>J. Lavorato - CEO, Enron America</td>
<td>0.02</td>
<td>-0.04</td>
<td>0.49</td>
<td>0.00</td>
</tr>
<tr>
<td>M. Grigsby - Director, West Desk Gas Trading</td>
<td>0.00</td>
<td>-0.03</td>
<td>0.20</td>
<td>-0.00</td>
</tr>
<tr>
<td>G. Whalley - President,</td>
<td>0.01</td>
<td>-0.01</td>
<td>0.19</td>
<td>0.00</td>
</tr>
<tr>
<td>J. Steffes - VP, Government Affairs</td>
<td>0.00</td>
<td>-0.02</td>
<td>0.18</td>
<td>0.00</td>
</tr>
<tr>
<td>K. Presto - VP, East Power Trading</td>
<td>0.01</td>
<td>-0.05</td>
<td>0.18</td>
<td>0.00</td>
</tr>
<tr>
<td>S. Beck - COO,</td>
<td>0.01</td>
<td>-0.03</td>
<td>0.17</td>
<td>0.00</td>
</tr>
<tr>
<td>B. Tycholiz - VP, Marketing</td>
<td>0.01</td>
<td>-0.02</td>
<td>0.16</td>
<td>0.00</td>
</tr>
<tr>
<td>J. Arnold - VP, Financial Enron Online</td>
<td>0.03</td>
<td>-0.04</td>
<td>0.16</td>
<td>-0.00</td>
</tr>
<tr>
<td>J. Williamson - Executive Assistant,</td>
<td>0.00</td>
<td>-0.02</td>
<td>0.14</td>
<td>0.01</td>
</tr>
<tr>
<td>K. Watson - Employee, Transwestern Pipeline Company (ETS)</td>
<td>-0.00</td>
<td>-0.00</td>
<td>0.01</td>
<td>0.59</td>
</tr>
<tr>
<td>M. Lokay - Admin. Asst., Transwestern Pipeline Company (ETS)</td>
<td>-0.00</td>
<td>0.01</td>
<td>0.01</td>
<td>0.42</td>
</tr>
<tr>
<td>L. Donoho - Employee, Transwestern Pipeline Company (ETS)</td>
<td>-0.00</td>
<td>0.01</td>
<td>0.01</td>
<td>0.35</td>
</tr>
<tr>
<td>M. Mcconnell - Employee, Transwestern Pipeline Company (ETS)</td>
<td>-0.00</td>
<td>-0.00</td>
<td>0.01</td>
<td>0.26</td>
</tr>
<tr>
<td>L. Blair - Employee, Northern Natural Gas Pipeline (ETS)</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.22</td>
</tr>
<tr>
<td>K. Hyatt - Director, Asset Development TW Pipeline Business (ETS)</td>
<td>-0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>0.20</td>
</tr>
<tr>
<td>D. Schoolcraft - Employee, Gas Control (ETS)</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.18</td>
</tr>
<tr>
<td>T. Geaccone - Manager, (ETS)</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.01</td>
<td>0.17</td>
</tr>
<tr>
<td>R. Hayslett - VP, Also CFO and Treasurer</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.02</td>
<td>0.16</td>
</tr>
</tbody>
</table>
Communication Patterns

- Mostly communication within roles
- Some large exchanges
- Negative values complicates interpretation
  - Non-negative factorization being investigated
Temporal Patterns

Communication patterns over time

- **Legal**
- **Government & regulatory affairs**
- **Trade executives**
- **Pipeline employee**

Normalized scale

Month

- Enron crisis breaks; investigation begins
- Filed for bankruptcy
Summary

- Improvements to DEDICOM
  - New procedure for finding $A$
  - Newton step for finding $D$

- Modifications to handle large data arrays
  - Compression

- Novel approach to social network analysis using DEDICOM
  - Roles of employees
  - Communication patterns among roles and over time

- Future research
  - Nonnegative DEDICOM
  - Constrained DEDICOM
  - PARAFAC
More Information

bwbader@sandia.gov

- DEDICOM paper on Social Network Analysis:
  - Tech report SAND2006-2161 available

- MATLAB Tensor Toolbox:
  - http://csmr.ca.sandia.gov/~tgkolda/TensorToolbox
  - Tech report SAND2004-5189 available on website
  - sparse_tensor class to be released soon