THE REASONABLE EFFECTIVENESS OF ROLES IN NETWORKS

Tina Eliassi-Rad

tina@eliassi.org  @tinaeliassi

Supported by NSF, DTRA, DARPA, IARPA, DOE/LLNL & WaPo Labs
What are Roles?

- Functions of nodes in the network
  - Similar to functional roles of species in ecosystems
- Roles are defined in terms of structural behaviors
  - What is your connectivity pattern?
  - To what kinds of individuals are you connected?
Example of Roles in an IP × IP Network

The types of neighbors that are connected to a given host are indicators of the host’s role.

Node sizes indicate communication volume relative to the central node in each frame.
Each Node has a Mixture of Roles
Research Questions

1. How are roles different from communities and from positions/equivalences (from sociology)?

2. Given a network, how can we automatically discover roles of nodes?

3. How can we make sense of these roles?

4. Are there good features that we can extract for nodes that indicate role-membership?

5. What are the applications in which these discovered roles can be effectively used?
Roles & Communities are Complementary

Roles (similar structural properties)

Communities (well-connectedness)

Roles are Similar to Positions from Sociology

• Two nodes with the same position are in an equivalence relation

• Equivalence, $Q$, is any relation that satisfies these three conditions:
  
  • Transitivity: $(a,b), (b,c) \in Q \Rightarrow (a,c) \in Q$
  
  • Symmetry: $(a, b) \in Q$ if and only if $(b, a) \in Q$
  
  • Reflexivity: $(a, a) \in Q$
Taxonomy of Equivalences from Sociology

- Equivalences
  - Deterministic
    - Regular
    - Automorphic
  - Probabilistic
    - Structural
    - Stochastic
Roles find Regular Equivalences

Two nodes $u$ and $v$ are regularly equivalent if they are equally related to equivalent others.

[Everett & Borgatti, 1992]
Finding Roles in a Network

**Input**

Node × Node Matrix

$n dim space$
Finding Roles in a Network

Input
Node × Node Matrix

Recursive Feature Extraction

Node × Feature Matrix

n dim space

f dim space

[RefEx: Henderson et al., KDD 2011]
ReFeX: Recursive Feature Extraction

- [Henderson et al., KDD 2011]
- Recursively combines node-based features with egonet-based features to output regional features
ReFeX: Recursive Feature Extraction

- [Henderson et al., KDD 2011]
- Recursively combines node-based features with egonet-based features to output regional features

- Neighborhood features: What is your connectivity pattern?
- Recursive Features: To what kinds of nodes are you connected?
ReFeX: Structural Features

- **Local**
  - Essentially measures of the node degree

- **Egonet**
  - Computed based on each node’s ego network
  - Examples
    - # of within-egonet edges
    - # of edges entering & leaving the egonet

- **Recursive**
  - Some aggregate (mean, sum, max, min, …) of another feature over a node’s neighbors
  - Aggregation can be computed over any real-valued feature, including other recursive features
ReFeX: Structural Features

- **Local**
  - Essentially measures of the node degree

- **Egonet**
  - Computed based on each node’s ego network
  - Examples
    - # of within-egonet edges
    - # of edges entering & leaving the egonet

- **Recursive**
  - Some aggregate (mean, sum, max, min, …) of another feature over a node’s neighbors
  - Aggregation can be computed over any real-valued feature, including other recursive features
ReFeX: Structural Features

- **Local**
  - Essentially measures of the node degree

- **Egonet**
  - Computed based on each node’s ego network
  - Examples
    - # of within-egonet edges
    - # of edges entering & leaving the egonet

- **Recursive**
  - Some aggregate (mean, sum, max, min, …) of another feature over a node’s neighbors
  - Aggregation can be computed over any real-valued feature, including other recursive features
ReFeX (continued)

- Number of possible recursive features is infinite
ReFeX (continued)

• Number of possible recursive features is infinite
• ReFeX pruning
ReFeX (continued)

- Number of possible recursive features is infinite
- ReFeX pruning
  - Feature values are mapped to small integers via vertical logarithmic binning
    - Log binning places most of the discriminatory power among sets of nodes with large feature values
ReFeX (continued)

- Number of possible recursive features is infinite
- ReFeX pruning
  - Feature values are mapped to small integers via vertical logarithmic binning
    - Log binning places most of the discriminatory power among sets of nodes with large feature values
  - Look for pairs of features whose values never disagree by more than a threshold
    - A graph-based approach
    - Threshold automatically set
    - Details in the KDD’11 paper
Finding Roles in a Network

Input

Node × Node Matrix

Recursive Feature Extraction

Node × Feature Matrix

f dim space

n dim space

[ReFex: Henderson et al., KDD 2011]
Finding Roles in a Network

Input

Node × Node Matrix

Recursive Feature Extraction

[f dim space]

Node × Feature Matrix

Role Extraction

[Rolex: Henderson et al., KDD 2012]

Node × Role Matrix

Role × Feature Matrix

[r dim space]

Output
Finding Roles in a Network

Input

Node × Node Matrix

\( n \) dim space

Recursive Feature Extraction

\( f \) dim space


Role Extraction

[ReFex: Henderson et al., KDD 2011]

Node × Role Matrix

Node × Feature Matrix

Role × Feature Matrix

\( r \) dim space

Output

\( n \gg f \gg r \)
Role Extraction: Feature Grouping

- Soft clustering in the structural feature space
  - Each node has a mixed-membership across roles
Role Extraction: Feature Grouping

• Soft clustering in the structural feature space
  • Each node has a mixed-membership across roles
• Generate a rank $r$ approximation of $V \approx GF$
Role Extraction: Feature Grouping

- Soft clustering in the structural feature space
  - Each node has a mixed-membership across roles
- Generate a rank $r$ approximation of $V \approx GF$

$\begin{align*}
    V &\approx GF \\
    V &\approx GF
\end{align*}$

- RolX uses NMF for feature grouping
  - Computationally efficient
  - Non-negative factors simplify interpretation of roles and memberships

$\text{argmin}_{G,F} \|V - GF\|_{fro}, \text{s.t. } G \geq 0, F \geq 0$
Role Extraction: Model Selection

• Roles summarize behavior
  • Or, they compress the feature matrix, $V$
Role Extraction: Model Selection

- Roles summarize behavior
  - Or, they compress the feature matrix, $V$
- Use MDL to select the model size $r$ that results in the best compression
  - $L$: description length
  - $M$: # of bits required to describe the model
  - $E$: cost of describing the reconstruction errors in $V - GF$
  - Minimize $L = M + E$
Role Extraction: Model Selection

• Roles summarize behavior
  • Or, they compress the feature matrix, \( V \)
• Use MDL to select the model size \( r \) that results in the best compression
  • \( L \): description length
  • \( M \): # of bits required to describe the model
  • \( E \): cost of describing the reconstruction errors in \( V - GF \)
• Minimize \( L = M + E \)
  • To compress high-precision floating point values, RolX combines Lloyd-Max quantization with Huffman codes
  • Errors in \( V - GF \) are not distributed normally, RolX uses KL divergence to compute \( E \)

\[
M = \bar{b}r(n + f)
\]

\[
E = \sum_{i,j} \left( V_{i,j} \log \frac{V_{i,j}}{(GF)_{i,j}} - V_{i,j} + (GF)_{i,j} \right)
\]
Finding Roles in a Network

Input

Node × Node Matrix

Recursive Feature Extraction

Node × Feature Matrix

Role Extraction

Node × Role Matrix

Role × Feature Matrix

Output

n ∝ f ∝ r

n dim space

f dim space

r dim space

RefEx: Henderson et al., KDD 2011

Rolx: Henderson et al., KDD 2012
Automatically Discovered Roles

Network Science Co-authorship Graph
[Newman 2006]
Automatically Discovered Roles

Network Science Co-authorship Graph
[Newman 2006]
Making Sense of Roles

Node Sense: $GE \approx M$

- Role 1
- Role 2
- Role 3
- Role 4
- Default

cliquey    bridge    periphery    isolated
Making Sense of Roles

Node Sense: \( GE \approx M \)

Neighbor Sense: \( GQ \approx N \)

- Role 1
- Role 2
- Role 3
- Role 4

cliquery, bridge, periphery, isolated
Making Sense of Roles

Topological measures & role homophily help interpret roles.

Node Sense: $GE \approx M$
Neighbor Sense: $GQ \approx N$

Role Affinities

Role 1
Role 2
Role 3
Role 4

Role 1
Role 2
Role 3
Role 4

Role-1 contribution to Node-measurement

Degree
Avg. Weight
Clustering Coeff.
Eccentricity
PageRank
GateKeeper
GateKeeper-Local
Pivot
Structural Hole
Role 1
Role 2
Role 3
Role 4

NodeSense
NeighborSense
cliquey
bridge
periphery
isolated
### Applications of role discovery

<table>
<thead>
<tr>
<th>Task</th>
<th>Use Case</th>
</tr>
</thead>
<tbody>
<tr>
<td>Role query</td>
<td>Identify individuals with similar behavior to a known target</td>
</tr>
<tr>
<td>Role outliers</td>
<td>Identify individuals with unusual behavior</td>
</tr>
<tr>
<td>Role dynamics</td>
<td>Identify unusual changes in behavior</td>
</tr>
<tr>
<td>Re-identification</td>
<td>Identify individuals in an anonymized network</td>
</tr>
<tr>
<td>Role transfer</td>
<td>Use knowledge of one network to make predictions in another</td>
</tr>
<tr>
<td>Network comparison</td>
<td>Determine network compatibility for knowledge transfer</td>
</tr>
<tr>
<td>Exploration in role space</td>
<td>Exploratory analysis of network data in the role space</td>
</tr>
</tbody>
</table>
# Applications of role discovery

<table>
<thead>
<tr>
<th>Task</th>
<th>Use Case</th>
</tr>
</thead>
<tbody>
<tr>
<td>Role query</td>
<td>Identify individuals with similar behavior to a known target</td>
</tr>
<tr>
<td>Role outliers</td>
<td>Identify individuals with unusual behavior</td>
</tr>
<tr>
<td>Role dynamics</td>
<td>Identify unusual changes in behavior</td>
</tr>
<tr>
<td>Re-identification</td>
<td>Identify individuals in an anonymized network</td>
</tr>
<tr>
<td>Role transfer</td>
<td>Use knowledge of one network to make predictions in another</td>
</tr>
<tr>
<td>Network comparison</td>
<td>Determine network compatibility for knowledge transfer</td>
</tr>
<tr>
<td>Exploration in role space</td>
<td>Exploratory analysis of network data in the role space</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Mapping Big Data
A Data-Driven Market Report
By Russell Jurney
Publisher: O'Reilly
Released: September 2015

Description
To discover the shape and structure of the big data market, the San Francisco-based startup Relato took a unique approach to market research and created the first fully data-driven market report. Company CEO Russell Jurney and his team collected and analyzed raw data from a variety of sources to reveal a boatload of business insights about the big data space. This exceptional report is now available for free download.

Using data analytic techniques such as social network analysis (SNA), Relato exposed the vast and complex partnership network that exists among tens of thousands of unique big data vendors. The dataset Relato collected is centered around Cloudera, Hortonworks, and MapR, the major platform vendors of Hadoop, the primary force behind this market.

From this snowball sample, a 2-hop network, the Relato team was able to answer several questions, including:

- Who are the major players in the big data market?
- Which is the leading Hadoop vendor?
- What sectors are included in this market and how do they relate?
- Which among the thousands of partnerships are most important?
- Who's doing business with whom?

Metrics used in this report are also visible in Relato's interactive web application, via a link in the report, which walks you through the insights step-by-step.

Russell Jurney is CEO of Relato, a San Francisco area startup that maps markets to drive sales and marketing. He is the author of Agile Data Science and co-author of Big Data for Chimps (both O'Reilly). In addition, Russell is an Apache Committer on the Incubating DataFu project. Russell is a full stack engineer.
Big data business partnerships

- **Green**: equal opportunity bridges; big-data vendors
- **Red**: middle-men; general IT vendors
- **Blue**: Strong affinity for big-data vendors; small vendors
Big-data business-partnerships
Louvain Clustering After Removing Small Vendors (Blue Role)
An Interactive Market Map of the Big Data Space

Role Transfer = RolX + SL
Role Transfer = RolX + SL

External Network

(1) RolX

Role Definitions

|   | f1 | f2 | f3 | f4 |...
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>r1</td>
<td>.4</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>r2</td>
<td>1</td>
<td>.2</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>r3</td>
<td>.1</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Role Memberships

<table>
<thead>
<tr>
<th></th>
<th>r1</th>
<th>r2</th>
<th>r3</th>
</tr>
</thead>
<tbody>
<tr>
<td>n1</td>
<td>.5</td>
<td>.5</td>
<td>0</td>
</tr>
<tr>
<td>n2</td>
<td>.2</td>
<td>.2</td>
<td>.6</td>
</tr>
<tr>
<td>n3</td>
<td>.25</td>
<td>.25</td>
<td>.5</td>
</tr>
<tr>
<td>:</td>
<td>:</td>
<td>:</td>
<td>:</td>
</tr>
</tbody>
</table>
Role Transfer = RolX + SL

External Network

(1) RolX

Role Definitions

<table>
<thead>
<tr>
<th>f1</th>
<th>f2</th>
<th>f3</th>
<th>f4</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>r1</td>
<td>.4</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>r2</td>
<td>1</td>
<td>.2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>r3</td>
<td>.1</td>
<td>1</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(2) Learning

Role Memberships

<table>
<thead>
<tr>
<th></th>
<th>r1</th>
<th>r2</th>
<th>r3</th>
</tr>
</thead>
<tbody>
<tr>
<td>n1</td>
<td>.5</td>
<td>.5</td>
<td>0</td>
</tr>
<tr>
<td>n2</td>
<td>.2</td>
<td>.2</td>
<td>.6</td>
</tr>
<tr>
<td>n3</td>
<td>.25</td>
<td>.25</td>
<td>.5</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

P(class | role membership)

Classifier

<table>
<thead>
<tr>
<th>class</th>
</tr>
</thead>
<tbody>
<tr>
<td>n1</td>
</tr>
<tr>
<td>n2</td>
</tr>
<tr>
<td>n3</td>
</tr>
<tr>
<td>...</td>
</tr>
</tbody>
</table>
Role Transfer = RolX + SL

External Network

(1) RolX

Role Definitions

f1 f2 f3 f4 ... r1 r2 r3
r1 .4 1 0 3
r2 1 .2 0 0
r3 .1 1 5 0
...

Role Memberships

(2) Learning

Classifier

P(class | role membership)

(3) Role Assignment

Role Memberships

<table>
<thead>
<tr>
<th>class</th>
<th>n1</th>
<th>n2</th>
<th>n3</th>
</tr>
</thead>
<tbody>
<tr>
<td>m1</td>
<td>.4</td>
<td>.4</td>
<td>.2</td>
</tr>
<tr>
<td>m2</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>m3</td>
<td>.2</td>
<td>.2</td>
<td>.6</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Role Transfer = RolX + SL
Roles Generalize Across Disjoint Networks

![Classification Accuracy Chart]

**Table:**

<table>
<thead>
<tr>
<th>TrainSet-TestSet</th>
<th>IP-A1</th>
<th>IP-A2</th>
<th>IP-A3</th>
<th>IP-A4</th>
<th>IP-B</th>
</tr>
</thead>
<tbody>
<tr>
<td># Nodes</td>
<td>81,450</td>
<td>57,415</td>
<td>154,103</td>
<td>206,704</td>
<td>181,267</td>
</tr>
<tr>
<td>% labeled</td>
<td>36.7%</td>
<td>28.1%</td>
<td>20.1%</td>
<td>32.9%</td>
<td>15.3%</td>
</tr>
<tr>
<td># Links</td>
<td>968,138</td>
<td>432,797</td>
<td>1,266,341</td>
<td>1,756,082</td>
<td>1,945,215</td>
</tr>
<tr>
<td>(# unique)</td>
<td>206,112</td>
<td>137,822</td>
<td>358,851</td>
<td>465,869</td>
<td>397,925</td>
</tr>
</tbody>
</table>

**Legend:**
- Red: RolX
- Light Blue: Feat
- Black: Default

**Graph Description:**
- Y-axis: Classification Accuracy

**Note:** The graph illustrates the classification accuracy for different train-test sets across various disjoint networks, highlighting how roles generalize across these networks.
RolX Model Selection

RolX selects high accuracy model sizes

9 roles automatically discovered by RolX
Classification accuracy is highest when RolX selection criterion is minimized.
2nd Generation Algorithms for Role Discovery

- **GLRD**: guided learning for role discovery
  - [Gilpin et al., KDD 2013]
- **DBMM**: dynamic behavioral mixed-membership model
  - [Rossi et al., WSDM 2013]
- **RC-Joint**: simultaneous detection of communities and roles
  - [Ruan & Parthasarathy, COSN 2014]
- **Motif-Role-Fingerprints**
  - [McDonnell et al., PLoS ONE 9(12), 2014]
- Dynamic inference of social roles in information cascades
  - [Choobdar et al., DMKD 29(5), 2015]
- **MRD**: multi-relational role discovery
  - [Gilpin et al., ArXiv 2016]
- **DERM**: dynamic edge role mixed-membership model
  - [Ahmed et al., ArXiv 2016]
- A combinatorial approach to role discovery
  - [Arockiasamy et al., ICDM 2016]
2nd Generation Algorithms for Role Discovery

- **GLRD**: guided learning for role discovery
  - [Gilpin et al., KDD 2013]
- **DBMM**: dynamic behavioral mixed-membership model
  - [Rossi et al., WSDM 2013]
- **RC-Joint**: simultaneous detection of communities and roles
  - [Ruan & Parthasarathy, COSN 2014]
- Motif-Role-Fingerprints
  - [McDonnell et al., PLoS ONE 9(12), 2014]
- Dynamic inference of social roles in information cascades
  - [Choobdar et al., DMKD 29(5), 2015]
- **MRD**: multi-relational role discovery
  - [Gilpin et al., ArXiv 2016]
- **DERM**: dynamic edge role mixed-membership model
  - [Ahmed et al., ArXiv 2016]
- A combinatorial approach to role discovery
  - [Arockiasamy et al., ICDM 2016]
Multi-relational Role Discovery (MRD)

- Moving beyond simple networks
- Suppose you have a multi-relational networks
- Example: Congressional co-sponsorship data

[Gilpin et al., ArXiv 2016]
No longer have an adjacency matrix

- We have a person \(\times\) person \(\times\) committee tensor

- Entry at \((i, j, k)\) indicates how often congress-person \(i\) and \(j\) co-sponsored a bill that was sent to committee \(k\) for a particular congressional committee

\[A_{i,j,k}\] indicates how often representative \(i\) and \(j\) co-sponsored a bill that was sent to committee \(k\) for a particular congressional committee.

\[\text{Figure 5: How multi-relational graphs are created}\]

\[\text{Table 2: The re-identification accuracy of identifying a congress-person after embedding in a common role space}\]

\[\begin{array}{c|c|c}
\text{Number} & \text{Accuracy (Var)} & \text{Re-ident.} \\
\hline
5 & 67.3\% (2.5) & 65.9\% (3.1) \\
10 & 65.9\% (3.1) & 65.4\% (3.1) \\
20 & 67.3\% (2.5) & 65.9\% (3.1) \\
\end{array}\]
Multi-relational Role Discovery (MRD)

- Multi-relational Role Discovery (MRD)
  - No orthogonality constraint on factors
  - Nonnegative Tucker decomposition
  - Alternating least squares

[Gilpin et al., ArXiv 2016]
Multi-relational Role Discovery (MRD)

- **Multi-relational Role Discovery** (MRD)
  - No orthogonality constraint on factors
  - Nonnegative Tucker decomposition
  - Alternating least squares

\[
\begin{align*}
\text{argmin}_{G,F,R,H} & \quad \|V - \sum_i \sum_j \sum_k h_{ijk} \ast g_k \circ f_k \circ r_k\|_{Fro} \\
\text{subject to:} & \quad G \geq 0, F \geq 0, R \geq 0, H \geq 0 \\
& \quad g_i(H) \leq d_{H_i}, i = 1 \ldots t_H \\
& \quad \text{where } g_i \text{ is a convex function}
\end{align*}
\]

Figure 2: The Tucker decomposition for role discovery. The diagrammatic explanation of Equation 3 is shown in the figure.

[Girolami et al., ArXiv 2016]
Multi-relational Role Discovery (MRD)

- Multi-relational Role Discovery (MRD)
  - No orthogonality constraint on factors
  - Nonnegative Tucker decomposition
  - Alternating least squares

The factor matrices are:
- groups of features (role definitions)
- groups of entities (groups)
- groups of relations (topics)
- Tucker core

[Gilpin et al., ArXiv 2016]
Experiments

- Data from U.S. House of Representatives
- Bill co-sponsorship data from 1979 (the start of the 96th Congress) to 2009 (the end of the 110th Congress)
- 15 committees, for which there were legislation in each congress from 96th to 110th
- 110th Congress (from 2007-09)
  - 453 representatives & 10,613 bills
  - Average degree in aggregated graph = 8.37
  - Median value of average degree across committee co-sponsorship graphs = 0.48
Multi-relational Role Discovery (MRD)

Multi-relational Role Discovery (MRD) [Gilpin et al., ArXiv 2016]
Groups of representatives

Group Members 1

- Millender–McDonald Juanita
- Obey David R.
- Tsongas Niki
- Speier Jackie
- Faleomavaega Eni F.H.
- Meehan Martin T.
- Edwards Donna F.
- Visclosky Peter J.
- Hoyer Steny H.
- Foster Bill
- Clyburn James E.
- Richardson Laura
- Becerra Xavier
- Pelosi Nancy
- Waters Maxine
- Velazquez Nydia M.
- Lantos Tom
- Childers Travis
- Cazayoux Donald J. Jr.
- Dicks Norman D.
Group 1 of representatives

<table>
<thead>
<tr>
<th>Name</th>
<th>Party</th>
<th>Exp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Millender-McDonald</td>
<td>D</td>
<td>11</td>
</tr>
<tr>
<td>Obey, David</td>
<td>D</td>
<td>38</td>
</tr>
<tr>
<td>Tsongas, Niki</td>
<td>D</td>
<td>0</td>
</tr>
<tr>
<td>Speier, Jackie</td>
<td>D</td>
<td>0</td>
</tr>
<tr>
<td>Faleomavaega, Eni</td>
<td>D</td>
<td>18</td>
</tr>
<tr>
<td>Meehan, Martin</td>
<td>D</td>
<td>14</td>
</tr>
<tr>
<td>Edwards, Donna</td>
<td>D</td>
<td>0</td>
</tr>
<tr>
<td>Visclosky, Peter</td>
<td>D</td>
<td>22</td>
</tr>
<tr>
<td>Hoyer, Steny</td>
<td>D</td>
<td>26</td>
</tr>
<tr>
<td>Foster, Bill</td>
<td>D</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 7: Samples of congressional representatives from each E-group (found in the 110th Congress Cosponsorship Graph) along with their party affiliation and years of service in U.S. House of Representatives at beginning of congress (2007).

Figure 8: R-groups for 100th congress. Each bar plot corresponds to a single R-group and the bars show how much each relation contributes to the respective relation R-group.
More insights into Group 1

**Group 1**

- **Democrats; mostly not mid-career**
- **Active in oversight & gov’t reform**
- **On the periphery, but lots of triangles**
Multi-relational Role Discovery (MRD)

[Gilpin et al., ArXiv 2016]
Interaction Graph of a Tucker Core

From the 110th Congress Co-sponsorships
## Measure Properties on the Interaction Graph

<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
<th>Computation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simplicity</td>
<td>To what extent are nodes connected to (role) similar types of nodes?</td>
<td>Average Node Degree</td>
</tr>
<tr>
<td>Sharing</td>
<td>How much can a group be separated into independent parts?</td>
<td>Mincut cost</td>
</tr>
<tr>
<td>Variability</td>
<td>How does the simplicity of nodes vary across the interaction graph?</td>
<td>Variance of node degree</td>
</tr>
<tr>
<td>Stability</td>
<td>How stable are the interactions between roles, groups, and topics?</td>
<td>Spectral gap</td>
</tr>
</tbody>
</table>
Macroscopic Properties of the Interaction Graphs from Tucker Cores

More bipartisan bills passed

Congress

Contribution

0 0.1 0.2 0.3 0.4 0.5 0.6

96 98 100 102 104 106 108 110

Cut Cost
Macroscopic Properties of the Interaction Graphs from Tucker Cores

More bipartisan bills passed

Bush Sr.

Congress

Contribution

Cut Cost
Macroscopic Properties of the Interaction Graphs from Tucker Cores

More bipartisan bills passed

Reagan’s 2nd term

Bush Sr.

Cut Cost

Congress

Contribution

0

0.1

0.2

0.3

0.4

0.5

0.6

96

98

100

102

104

106

108

110
Macroscopic Properties of the Interaction Graphs from Tucker Cores

More bipartisan bills passed

Congress

Reagan’s 1st term

Bush Sr.

Clinton

Bush Jr.’s 2nd term

Reagan’s 2nd term

Contribution

Cut Cost
Role Transfer in MRD

• Extract roles on one multi-relational network

• How well do the extracted roles transfer to another multi-relational network?

\[ \mathcal{V}_1 \xrightarrow{\text{MRD}} \mathcal{G} \xrightarrow{\text{role definitions}} \mathcal{R} \]

\[ \mathcal{V}_2 \xrightarrow{\text{?}} \mathcal{G} \xrightarrow{\text{role definitions}} \mathcal{R} \]
Role Transfer on Multi-relational Networks

Heatmap of fit quality = 1 – normalized reconstruction error
Role Transfer on Multi-relational Networks

Heatmap of fit quality = 1 – normalized reconstruction error
Role Transfer

Hastert Rule: the Speaker will not allow a floor vote on a bill unless a majority of the majority party supports the bill.

Heatmap of fit quality = 1 – normalized reconstruction error
Why are Roles Effective in Many Applications?

• Encode complex behavior
• Map nodes into a useful lower dimensional space
• Generalize across networks
Lots more to do …

• An in-depth study on properties of these latent role spaces

• Information spread through roles
  • How roles affect influence & susceptibility?
Lots more to do …

• An in-depth study on properties of these latent role spaces

• Information spread through roles
  • How roles affect influence & susceptibility?

• Combining physics of networks (PoN) with the mining of graphs (MoG)
  • What are the functional roles in an ensemble of networks?
  • How do we incorporate functional roles from instances of networks into PoN models?
Thank You!

- Papers at http://eliassi.org/pubs.html
- Tutorials at http://eliassi.org
- Open-source code at https://snap.stanford.edu/snap-2.3/
- Joint work with
  - LLNL (Keith Henderson & Brian Gallagher)
  - CMU (Christos Faloutsos, Leman Akoglu et al.)
  - Google (Sugato Basu)
  - UC Davis (Ian Davidson et al.)