Interactive Data Visualization for Rapid Understanding of Scientific Literature

Cody Dunne
Dept. of Computer Science and Human-Computer Interaction Lab, University of Maryland
cdunne@cs.umd.edu

Links from this talk:

STM Annual Spring Conference 2011
April 26-28, 2010  Washington, DC
Roadmap

• Network Analysis 101
• Action Science Explorer (ASE)
• Getting Started with Visualization
Network Theory

• Central tenet
  – Social structure emerges from the aggregate of relationships (ties)

• Phenomena of interest
  – Emergence of cliques and clusters
  – Centrality (core), periphery (isolates)

Terminology

- **Node**
  - “actor” on which relationships act; 1-mode versus 2-mode networks
- **Edge**
  - Relationship connecting nodes; can be directional
- **Cohesive Sub-Group**
  - Well-connected group; clique; cluster; community
- **Key Metrics**
  - **Centrality** (group or individual measure)
    - Number of direct connections that individuals have with others in the group (usually look at incoming connections only)
    - Measure at the individual node or group level
  - **Cohesion** (group measure)
    - Ease with which a network can connect
    - Aggregate measure of shortest path between each node pair at network level reflects average distance
  - **Density** (group measure)
    - Robustness of the network
    - Number of connections that exist in the group out of 100% possible
  - **Betweenness** (individual measure)
    - # shortest paths between each node pair that a node is on
    - Measure at the individual node level
- **Node roles**
  - Peripheral – below average centrality
  - Central connector – above average centrality
  - Broker – above average betweenness
**Action Science Explorer**

The figure shows a software interface, likely related to data analysis or data exploration. The interface includes a citation network, a table with data entries, and various panes for selecting and manipulating data. The table seems to be related to dependency parsing, with columns for references, authors, title, and other metadata. The citation network highlights connections between different data points, possibly indicating relationships or dependencies in the data set being analyzed.
NodeXL
FOSS Social Network Analysis add-in for Excel 2007/2010

<table>
<thead>
<tr>
<th>Vertex</th>
<th>In Degree</th>
<th>Out Degree</th>
<th>Betweenness Centrality</th>
<th>Closeness Centrality</th>
<th>Eigenvector Centrality</th>
<th>PageRank</th>
</tr>
</thead>
<tbody>
<tr>
<td>politics</td>
<td>112</td>
<td>2</td>
<td>6345.3406</td>
<td>0.001</td>
<td>0.002</td>
<td>6.29</td>
</tr>
<tr>
<td>scottwgraves</td>
<td>153</td>
<td>10</td>
<td>45379.123</td>
<td>0.001</td>
<td>0.007</td>
<td>6.95</td>
</tr>
<tr>
<td>conservative1</td>
<td>146</td>
<td>93</td>
<td>24054.904</td>
<td>0.001</td>
<td>0.009</td>
<td>6.06</td>
</tr>
<tr>
<td>arizona_freedom</td>
<td>150</td>
<td>102</td>
<td>28549.038</td>
<td>0.001</td>
<td>0.009</td>
<td>5.79</td>
</tr>
<tr>
<td>conservative432</td>
<td>135</td>
<td>74</td>
<td>20294.618</td>
<td>0.001</td>
<td>0.009</td>
<td>5.36</td>
</tr>
<tr>
<td>kosmosnet</td>
<td>111</td>
<td>5</td>
<td>15621.759</td>
<td>0.001</td>
<td>0.005</td>
<td>4.16</td>
</tr>
<tr>
<td>karylinidy</td>
<td>138</td>
<td>24</td>
<td>12879.665</td>
<td>0.001</td>
<td>0.007</td>
<td>4.76</td>
</tr>
<tr>
<td>goodperbadpork</td>
<td>5</td>
<td>46</td>
<td>12209.131</td>
<td>0.001</td>
<td>0.001</td>
<td>2.33</td>
</tr>
<tr>
<td>larrycsbato</td>
<td>49</td>
<td>3</td>
<td>5405.256</td>
<td>0.001</td>
<td>0.007</td>
<td>2.38</td>
</tr>
<tr>
<td>tennishampa</td>
<td>33</td>
<td>58</td>
<td>5920.756</td>
<td>0.001</td>
<td>0.001</td>
<td>2.23</td>
</tr>
<tr>
<td>newsbusters</td>
<td>104</td>
<td>14</td>
<td>8663.937</td>
<td>0.001</td>
<td>0.004</td>
<td>3.75</td>
</tr>
<tr>
<td>whisper1111</td>
<td>34</td>
<td>37</td>
<td>8514.875</td>
<td>0.001</td>
<td>0.000</td>
<td>3.67</td>
</tr>
<tr>
<td>stale</td>
<td>92</td>
<td>4</td>
<td>8418.530</td>
<td>0.001</td>
<td>0.005</td>
<td>3.17</td>
</tr>
</tbody>
</table>
Import from Twitter User’s Network

Optionally clear the NodeXL workbook, then get the network of people followed by a Twitter user or the network of people following the user.
Hello!

We are a group of researchers who want to create open tools, generate and host open data, and support open scholarship related to social media.

Social media is the term for all the ways people connect to people through computation. Mobile devices, social networks, micro-blogging and location sharing are just a few of the ways people engage in computer-mediated collective action.

Mapping, measuring and understanding the landscape of social media is our mission. We support tool projects that enable the collection, analysis and visualization of social media data. We host data sets that are relevant to social media research. And we will support graduate students studying and building research related to social media.
MAY 25TH - 26TH, 2011

HCIL’s 28th Annual Symposium will highlight the cutting-edge research being conducted in the Human-Computer Interaction Laboratory at the University of Maryland. The Symposium will take place **Wednesday, May 25th and Thursday, May 26th.**

This year for the first time, the Symposium, Tutorials and Workshops merge into an extended program with parallel sessions. Attendees can focus on suggested tracks or sample talks from all topics.

**Social Media | Consumer Health Informatics | Electronic Health Record Informatics | Information Visualization | Interaction Design & Children | All Topics**

[REGISTER]

Special thanks to our Sponsors

[SAP] [Palantir] [Deloitte] [Palantir] [The Washington Post Company] [Spotfire] [TELUS] [interactions] [virtusa]
Treemaps can find papers & patents missed by searches
Overview

• Network Analysis 101
• Action Science Explorer (ASE)
• Getting Started with Visualization
Interactive Data Visualization for Rapid Understanding of Scientific Literature

Cody Dunne
Dept. of Computer Science and Human-Computer Interaction Lab, University of Maryland
cdunne@cs.umd.edu

This work has been partially supported by NSF grant "iOPENER: A Flexible Framework to Support Rapid Learning in Unfamiliar Research Domains", jointly awarded to UMD and UMich as IIS 0705832.

Links from this talk: