Introduction to ORA Reports

Jeff Reminga

The CASOS Center
COS Program, School of Computer Science, Carnegie Mellon
Summer Institute 2018

Agenda

• What is a report?
• How do you run one?
• How to understand report tables
• Some important things to keep in mind
What is a report?

- Analysis of a meta-network using:
  - Network visualizations
  - Charts
  - Measure tables
  - E.g. similar to reports in other software packages

- Each report answers a specific question by running measures and providing graphics
- Quickly generated via “wizard” style inputs

A Common Report Workflow

- Given a new meta-network, this sequence of reports is often run to answer the following questions:

  1. What are the key people, knowledge, tasks?
     → Key Entity or Standard Network Analysis

  2. Show me all there is to know about an entity of interest.
     → Sphere of Influence

  3. What if I removed a key knowledge? Or the manager?
     → Immediate Impact
General Report Properties

- Reports take as input one or more meta-networks
  - Tailor output based on the number of inputs
  - “comparison” mode and “multi-mode”

- Reports run only the measures listed in the Measures Manager (main menu Analysis \ Measures Manager)
  - Turn off slow measures via the menu: Preferences\Measures

Key Entity Report

Load the Software Company dataset that we used for data importing.

With the Software Company selected in the meta-network manager, click the Generate Reports button (circled in red).
Key Entity Report...

Select a report: start typing a name, and they are filtered, or select by category.

Here, the Key Entity report is selected.

Select the meta-network(s) on which to run the report.

- **Output formats** shows what the report does when more than one meta-network is given to be analyzed.
  - Single meta-network exists for all reports.
  - Side-by-side comparison means that for two meta-networks the report creates side-by-side analysis output.
  - Multiple meta-network comparison there is special formatting for 3 or more meta-networks.

- If there is no comparison or multi modes, then each meta-network is run individually, and the output “stacked”.
Next we select the meta-networks for input.

We have only one, but if there were more, we could choose as many as we would want.

Use the Filter Data page to select which nodesets and networks to analyze.

This is optional, as by default data are selected.

This is useful when you have large data.
Select which analyses to run for the report.

Using K-Betweenness Centrality is useful for large data.

Clicking Next we arrive at the last panel where we give the file formats to create and the filename.

Use CSV to generate useful values of measure tables to input into other statistics packages.
The bar chart at the top lists the Agents that repeatedly rank highest in the measures that follow.

What are those measures...

The Measure List directly below the bar chart lists the measures.

Click on one to go to a measure table.

We will go to: In-the-Know
Key Entity Report: Table...

Input network: describes the network upon the measure was run – it had 16 nodes, and a density of 0.30833

Values in red are for nodes that have a score >= 1 std.dev above the mean score (reported below the table).

Values in green are for nodes that score within a standard deviation of the mean.

Values in blue are for nodes that score one std.dev below the mean.

The value column lists the scaled node value, almost always scaled into [0,1].

Unscaled means the measure score directly computed.

The nodes are ranked highest to lowest by measure value.

The Context column gives the expected score for the node in a random network of the same size and density.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Agent</th>
<th>Value</th>
<th>Unscaled</th>
<th>Context*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Software Engineer</td>
<td>0.533</td>
<td>8</td>
<td>1.949</td>
</tr>
<tr>
<td>2</td>
<td>Technical Lead</td>
<td>0.533</td>
<td>8</td>
<td>1.949</td>
</tr>
<tr>
<td>3</td>
<td>Application Architect</td>
<td>0.467</td>
<td>7</td>
<td>1.371</td>
</tr>
<tr>
<td>4</td>
<td>Design Lead</td>
<td>0.409</td>
<td>6</td>
<td>0.794</td>
</tr>
<tr>
<td>5</td>
<td>Project Manager</td>
<td>0.409</td>
<td>6</td>
<td>0.794</td>
</tr>
<tr>
<td>6</td>
<td>Web Developer</td>
<td>0.409</td>
<td>6</td>
<td>0.794</td>
</tr>
<tr>
<td>7</td>
<td>Art Director</td>
<td>0.333</td>
<td>5</td>
<td>0.217</td>
</tr>
<tr>
<td>8</td>
<td>Business Analyst 1</td>
<td>0.267</td>
<td>4</td>
<td>-0.361</td>
</tr>
<tr>
<td>9</td>
<td>Business Analyst 2</td>
<td>0.267</td>
<td>4</td>
<td>-0.361</td>
</tr>
<tr>
<td>10</td>
<td>Data Architect</td>
<td>0.267</td>
<td>4</td>
<td>-0.361</td>
</tr>
</tbody>
</table>

* Number of standard deviations from the mean of a random network of the same size and density

Min: 0.067  Max: 0.533
Mean: 0.308  Mean in random network: 0.308
Std.dev: 0.137  Std.dev in random network: 0.115
When two meta-networks are selected as input to the report, then a **Comparison** format is used in the tables.

The input meta-networks should have nodesets and networks with the same name. Only networks with the same name are compared against each other.

The %Diff columns compare the same node with its score the other period. **N/A** is reported if the node does not exist in both meta-networks.

---

**Things to Keep In Mind**

- You need to know what the measure is calculating, that is, what is being measured.

- You need to know whether the networks are:
  - Symmetric (links go in both directions)
    - Most measures work directly on the given data and do not first make the network symmetric before computing
  - Weighted
    - Most measures use link weights
    - Affects measure interpretations (e.g. total degree centrality)

- See the Measure Manager for a list of such measures