Social Network Analysis for Computer Scientists

F.W. Takes

Leiden University, The Netherlands

Code review session
In the next weeks . . .

- Finish writing code
- Run experiments
- Evaluate results
- Write remaining sections of the paper
- Dec 12: project deadline
- Report any questions, issues, difficulties or problems
Code review

- **Peer review**: evaluation of work by one or more individuals with similar competence
- “Pair programming”
- Four eyes see more than two
- Go beyond your current knowledge and skills
- Output: list of “best practices”
Today

- Explain your work to the other team
- Mention what you have done and not yet done
- Introduce the other team to your code
- Ask questions about the other team’s code
- Explain to the other team positive and less positive constructive points about their work
- Together, derive useful “best practices” and add them to the Google Doc (see website)
Evaluation criteria

- Correctness guarantees
- Time and memory constraints
- Are input and output data validated for consistency?
- Is the output easily reusable for result tables or diagrams?
- Is there a pipeline of experiments to run different algorithms on different datasets?
About the Data and Experiments

- Is the data relevant and sufficient?
- Is the data “diverse” in relevant dimensions?
- What do you measure in each experiment? Quality, running time, error?
- Why is this data good for these experiments?
- Is the data possibly biased and how may this affect the experiments?
Remember . . .

Please be constructive!
### Team pairs

<table>
<thead>
<tr>
<th>Betweenness centrality 1</th>
<th>Betweenness centrality 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Influence spread 1</td>
<td>Influence spread 2</td>
</tr>
<tr>
<td>Visualization 1</td>
<td>Visualization 2</td>
</tr>
<tr>
<td>Counting triangles</td>
<td>Diameter computation</td>
</tr>
<tr>
<td>Network dissolution</td>
<td>Personalized PageRank</td>
</tr>
<tr>
<td>Network de-anonymization</td>
<td>Sampling methods 1</td>
</tr>
<tr>
<td>Link prediction</td>
<td>Signed link prediction</td>
</tr>
<tr>
<td>Disease spread</td>
<td>Minimum weight cycles</td>
</tr>
<tr>
<td>Neighborhoods</td>
<td>Graph compression</td>
</tr>
<tr>
<td>Network motifs</td>
<td>Community detection</td>
</tr>
<tr>
<td>Shortest paths 1</td>
<td>Shortest paths 2</td>
</tr>
<tr>
<td>Closeness centrality 2</td>
<td>Sampling methods 2</td>
</tr>
</tbody>
</table>