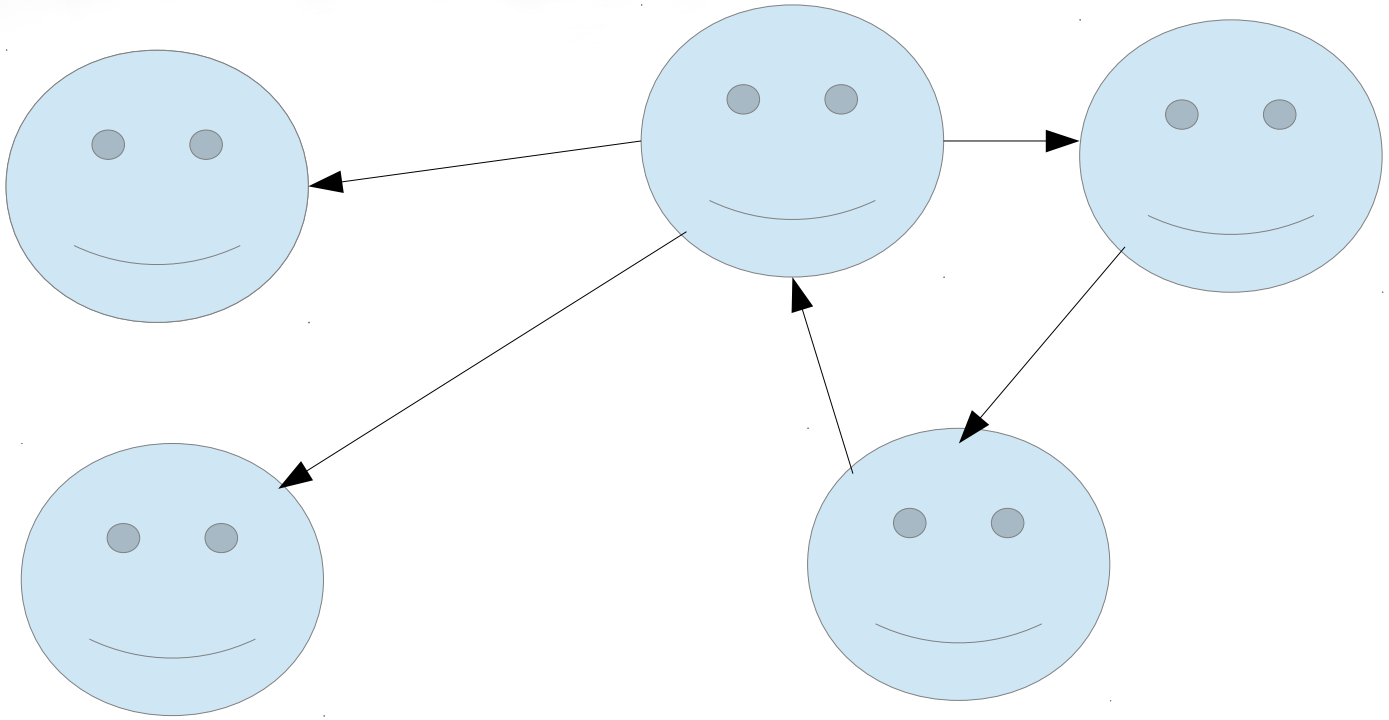
Social Network Analysis

Abram Hindle
abram.hindle@ualberta.ca
Department of Computing Science
University of Alberta
http://softwareprocess.es/

Social Network Analysis

- Not Facebook!
- Social Networks
  - Social relationships represented by networks
  - Networks can be considered to be graphs
  - Focus on SOCIAL

Social Network Analysis



Social Network Analysis

- Focus on modelling and measuring social networks
- Measure the entire graph
- Measure a single node and its relationships

Social Network Analysis

- **Nodes**
  - ***Things***
  - ***People***
  - ***Artifacts***
  - ***Aspects of people***
  - ***Documents***

Software Engineering

- **Programmers**
- **Projects**
- **Stakeholders**
- **Artifacts: bugs, files, methods, functions, classes, lines of code, revisions, commits, issues, tickets, requirements, documentation, etc.**

Social Network Analysis

- Arcs/Edges
  - Relationships
  - Relations
  - Associations
  - Communication
- Software Engineering
  - Communication
  - Discussion
  - Relationship between authors and files
  - Method Calls
  - Dependencies
  - Coupling
  - Committed Together

Social Network Measures

- **Entire Graph**
  - **Radius**
  - **Diameter**
  - **Size**
- **Nodes/Vertices**
  - **Degree Centrality**
  - **Between Centrality**

Entire Graph Metrics

- |V| : number of nodes/vertices
- |E| : number of edges/arcs
- eccentricity(v): given a node, what's the maximum shortest path to other nodes.
- radius(g): min eccentricity(v) for all v in V
- diameter(g): max eccentricity(v) for all v in V

Node Measures

- Centrality is often Importance
- In SE centrality often indicates expertise [Meneely et al.]
- Degree Centrality: deg(v) , how many edges v has.
- Betweeness Centrality:
  - o_st(v) number of paths
    between s & t that include v
  - o_st – number of paths
    between s and t

$$C_B(v) = \sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

  - Wikipedia Closeness Centrality
  - http://en.wikipedia.org/wiki/Closeness_centrality

Results from Meneely et al.

- In contribution networks
  - Edges are often collaboration
  - Higher network centrality indicate project experts
  - Distance between developers is perceived distance
    - Beyond degree 2 there is limited relation

- Socio-Technical Developer Networks: Should We Trust Our Measurements? Andrew Meneely and Laurie Williams, ICSE 2011