

STATISTICAL ANALYSIS ON FLOW SPLICE GRAPHS

Labcourse Graph Theory

Daniel Mayer, Benjamin Schindler

February 23, 2018

University of Leipzig

outline

- Methods
 - Software
 - own Statistics
- Results
 - task A general statistics
 - task B categorical statistics on EP True
- Discussion

methods

Software:

- scripting in python2
 - networkx as graph library
 - graph data written to CSV
- statistics in R using ggplot2

Source code available at

https://github.com/gnummig/labcourse_graphtheory

own statistics



basic statistics

Our dataset is EP True:

- no noise
- perfect alignment

We compare the **filtered** and the **splice graph** Graphs

- 5000 graphs per category
- 125.000 Problem nodes
- 500 Chimaer nodes

basic statistics

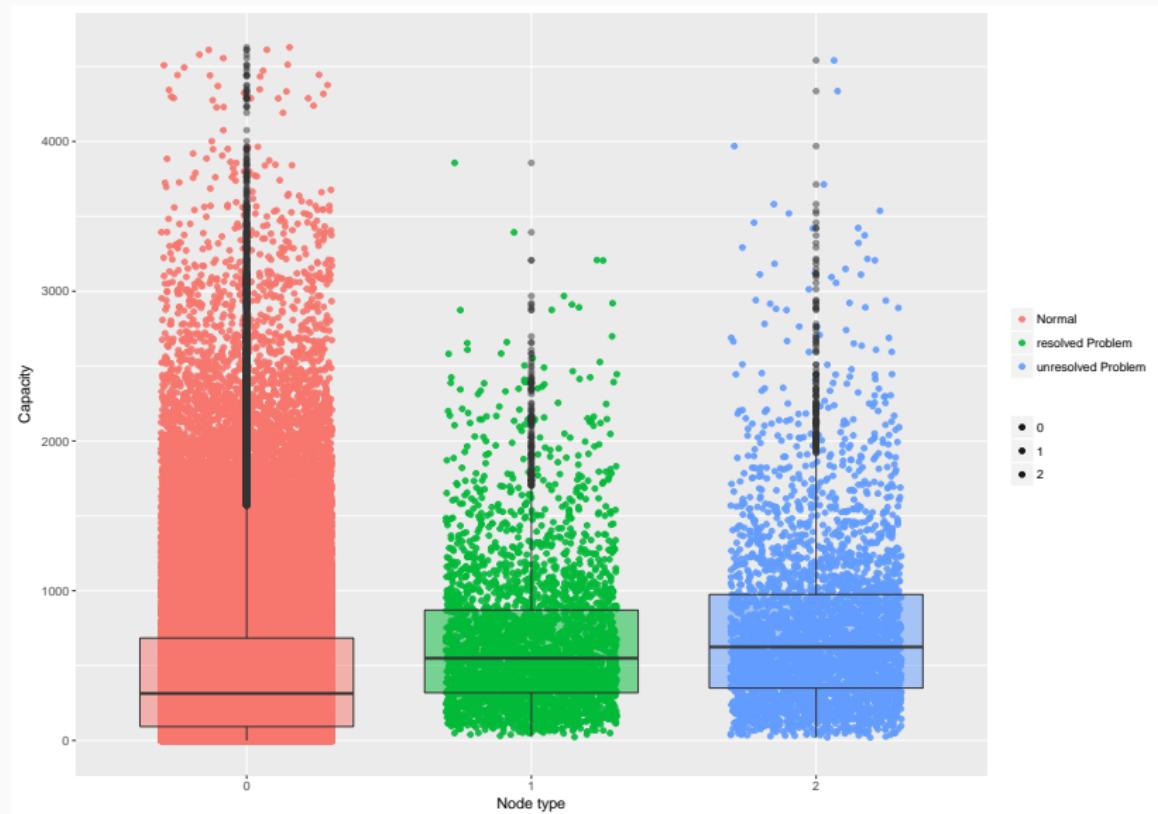


Figure: Capacity for filtered Graph

basic statistics

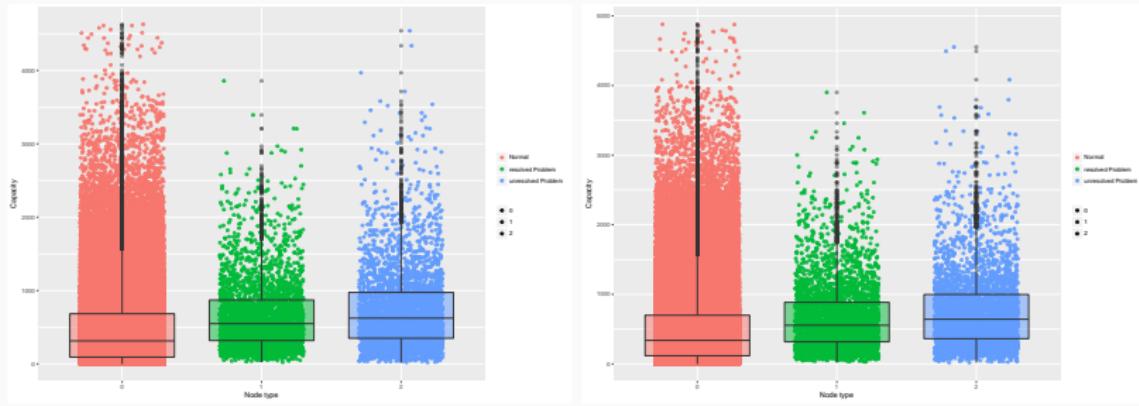


Figure: Capacity for filtered Graph and Splice Graph

basic statistics

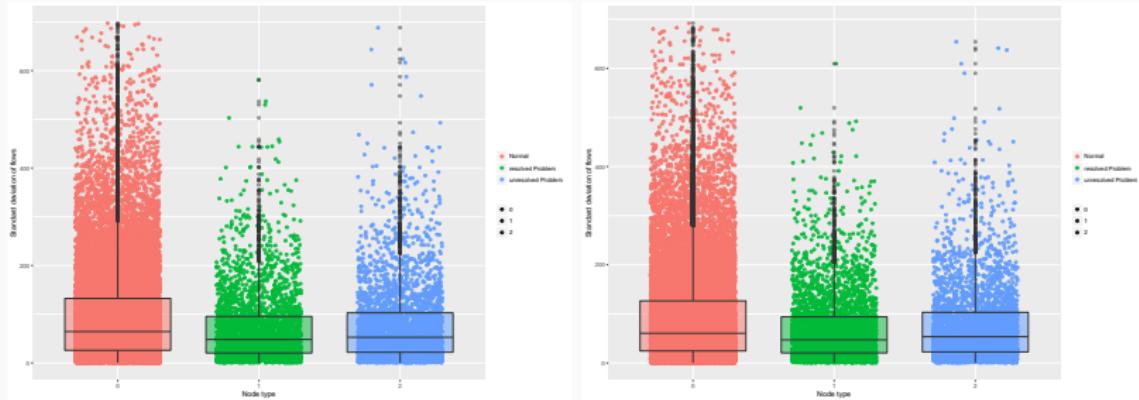


Figure: Standard deviation of in and out capacities for filtered Graph and Splice Graph

basic statistics

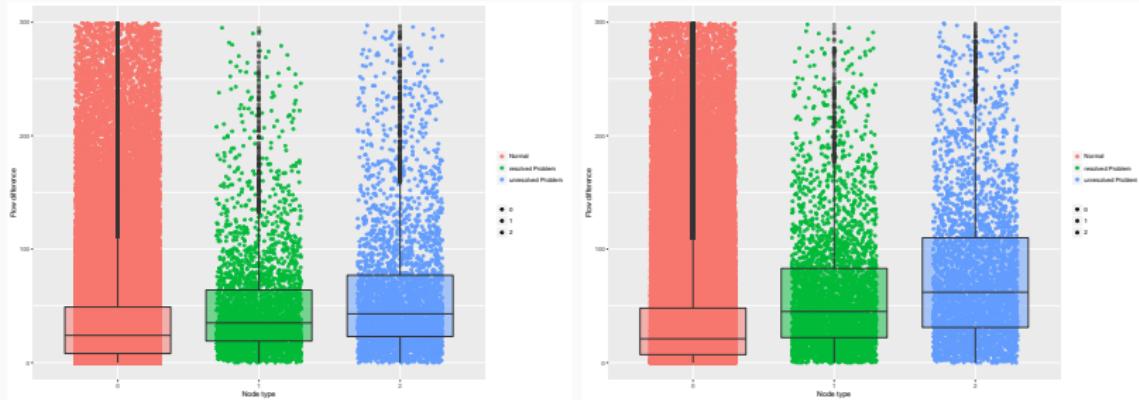
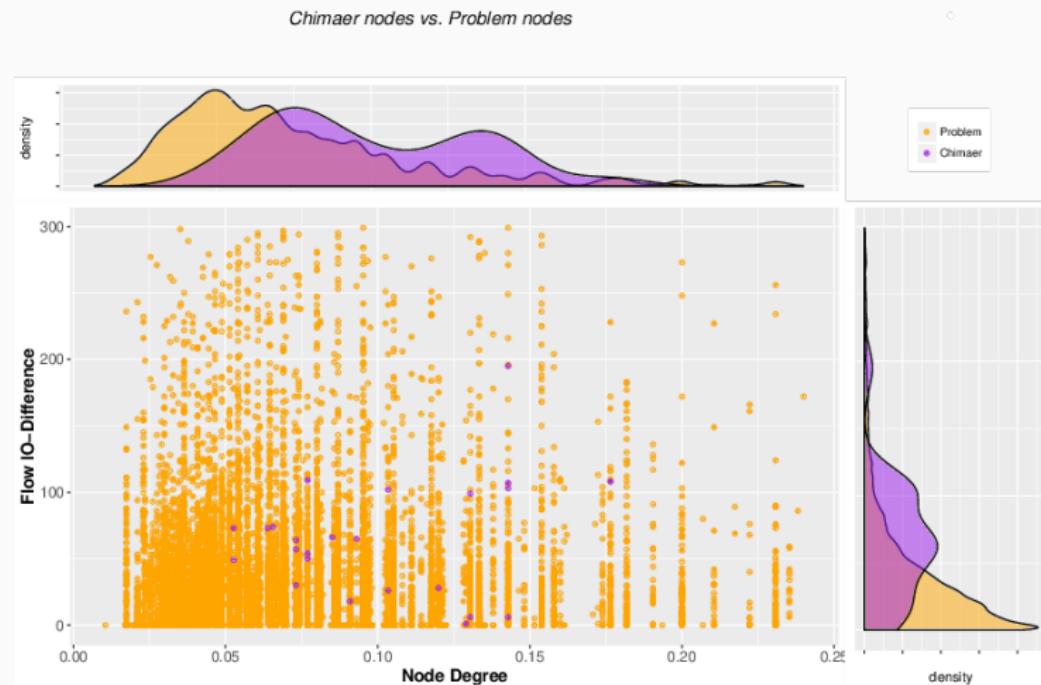


Figure: Difference in InCapacity and OutCapacity for filtered Graph and Splice Graph

chimaer nodes in the context of problem nodes



chimaer nodes in the context of problem nodes

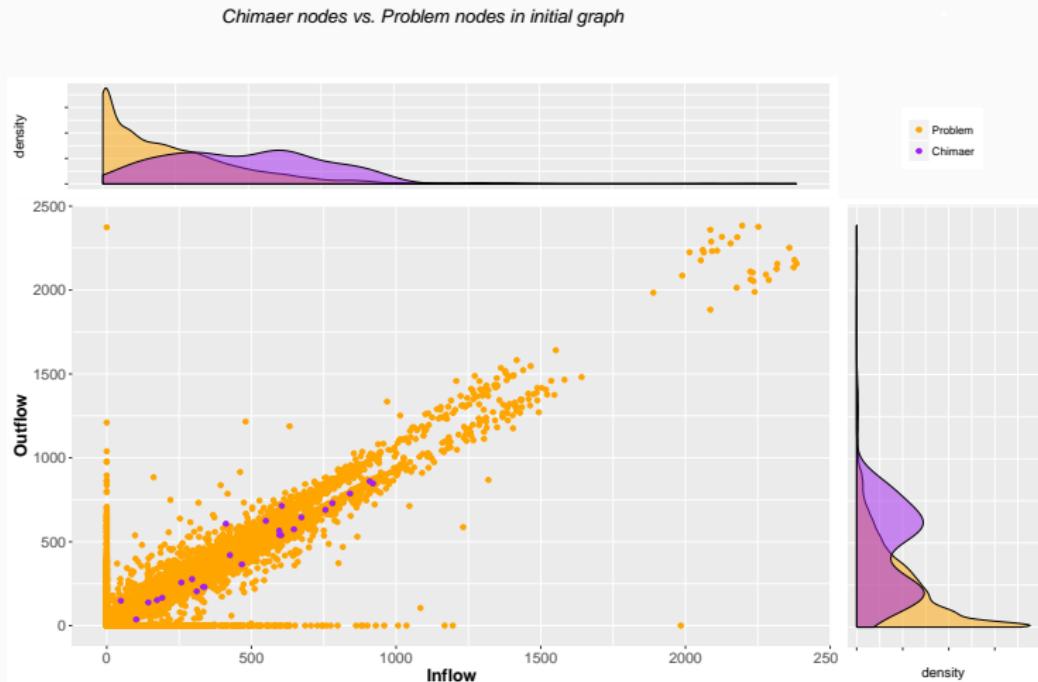


Figure: Input flow and output flow for Chimaer and other Problem Nodes

chimaer nodes in the context of problem nodes

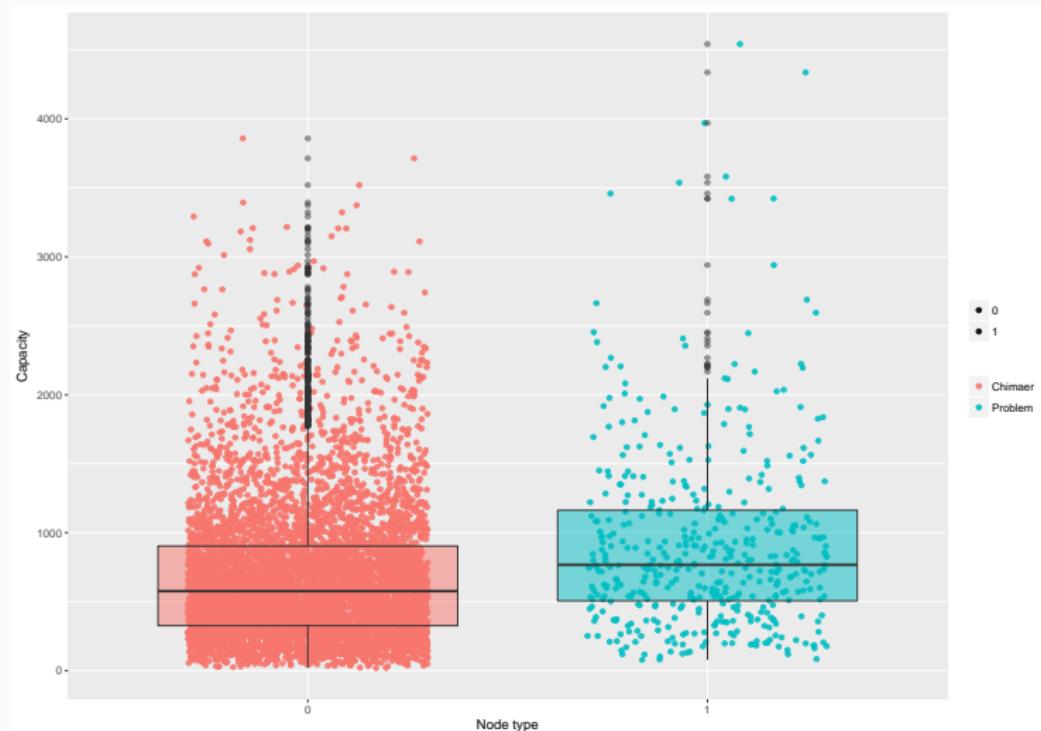


Figure: Capacities for Chimaer and other Problem Nodes

QUESTIONS?