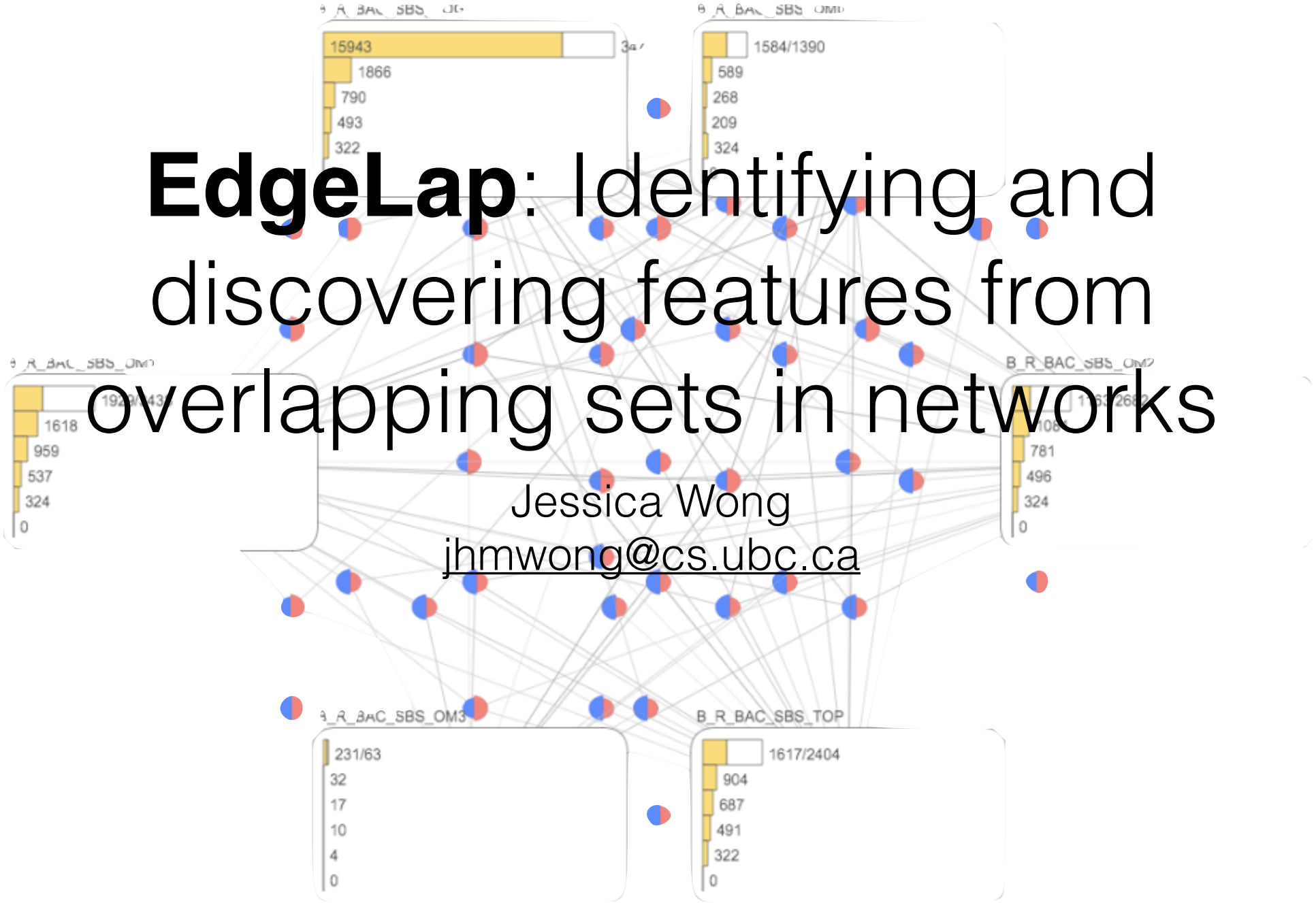
EdgeLap: Identifying and discovering features from overlapping sets in networks

Jessica Wong
jhmwong@cs.ubc.ca

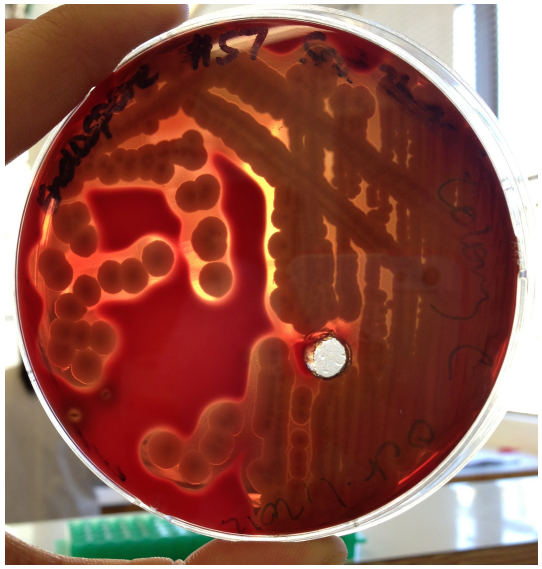


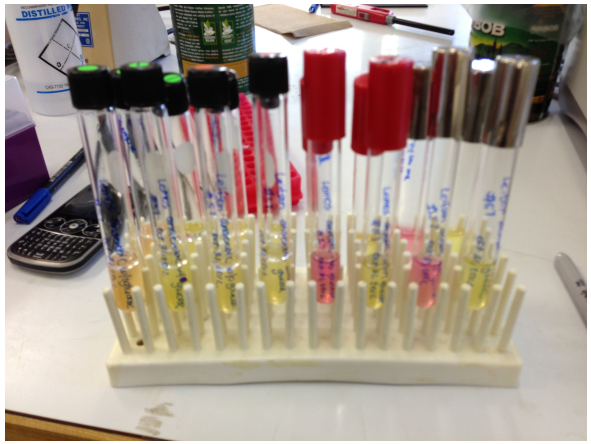
Data Domain

- Microbes are **everywhere** and commonly work together to survive
- A group of microbes working together is called a **community**
- Common task is to identify microbes present in a community and try to find patterns in their interactions

Data Domain

- You can identify microbes in many different ways





Data Domain

- Soil samples were obtained from 18 different sites
- Each site had 4 different methods of harvesting DNA
- Microbes were identified using DNA

Data Domain

- Each community has microbes and interactions between them
- Interactions can be positively correlated (copresence) or negatively correlated (mutually exclusive)
- Total of 605479 interactions

Task Abstraction

- Identify and locate common microbial interactions that happen across different networks
- Goal: to generate a hypothesis about the types, and number of correlations found between microbes in different communities

Complications

- We want to know common interactions that can occur in any combination of networks examined
  - 2^n combinations if there are n networks selected
- Also need to differentiate between positive/ negative edges

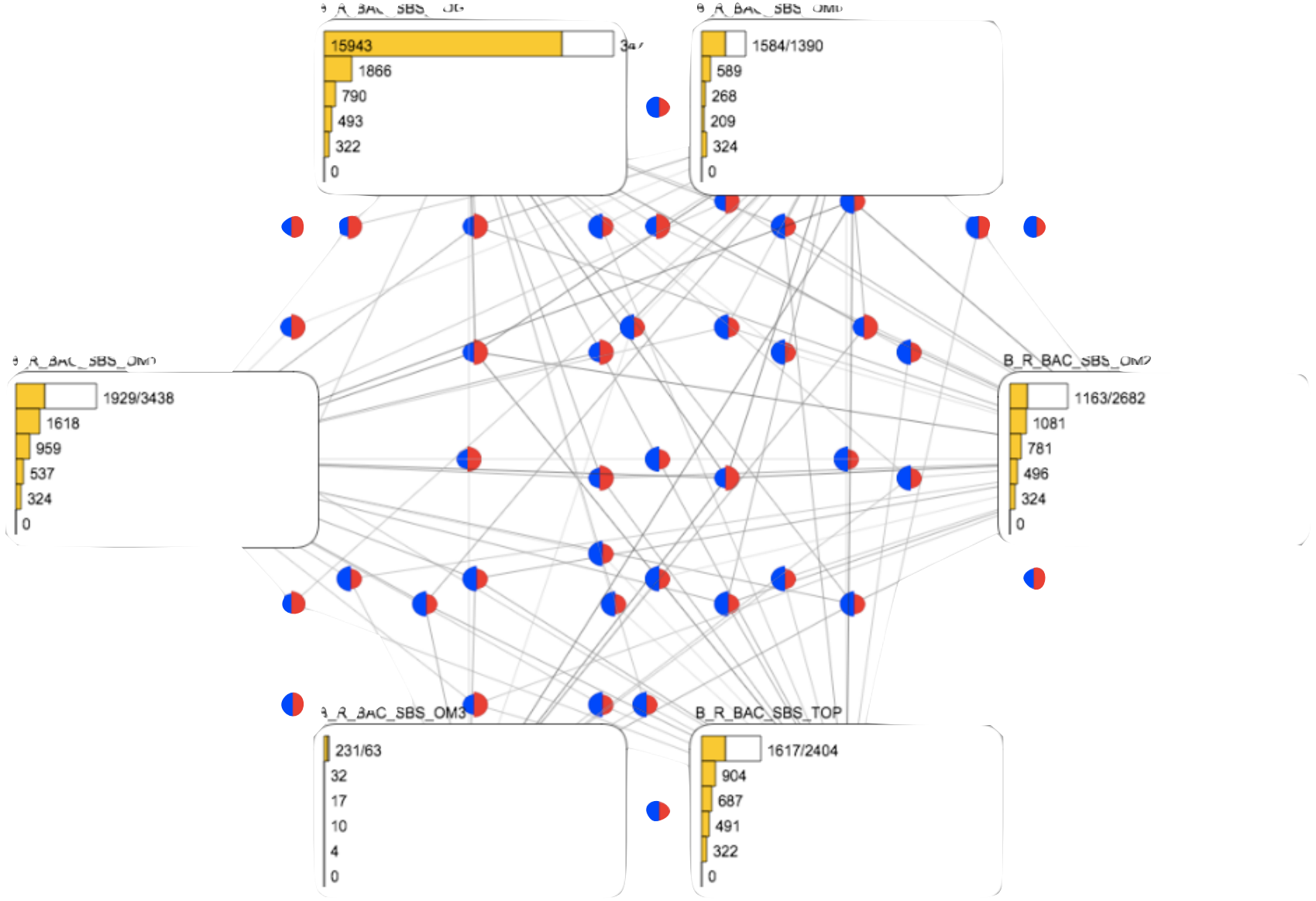
EdgeLap

- Inspired from Radial Sets
- Created from a mix of Java, HTML, PHP, JavaScript, and Processing
- Finds common interactions between 2-7 networks

Step One: Data

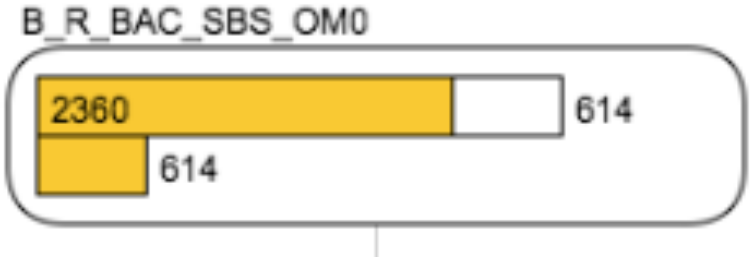
- Data needed to be processed and stored into a database
- Stored in Amazon RDS
- Sorted each interaction by name
- Created an index on the table to help speed up queries

Step Two: Display!



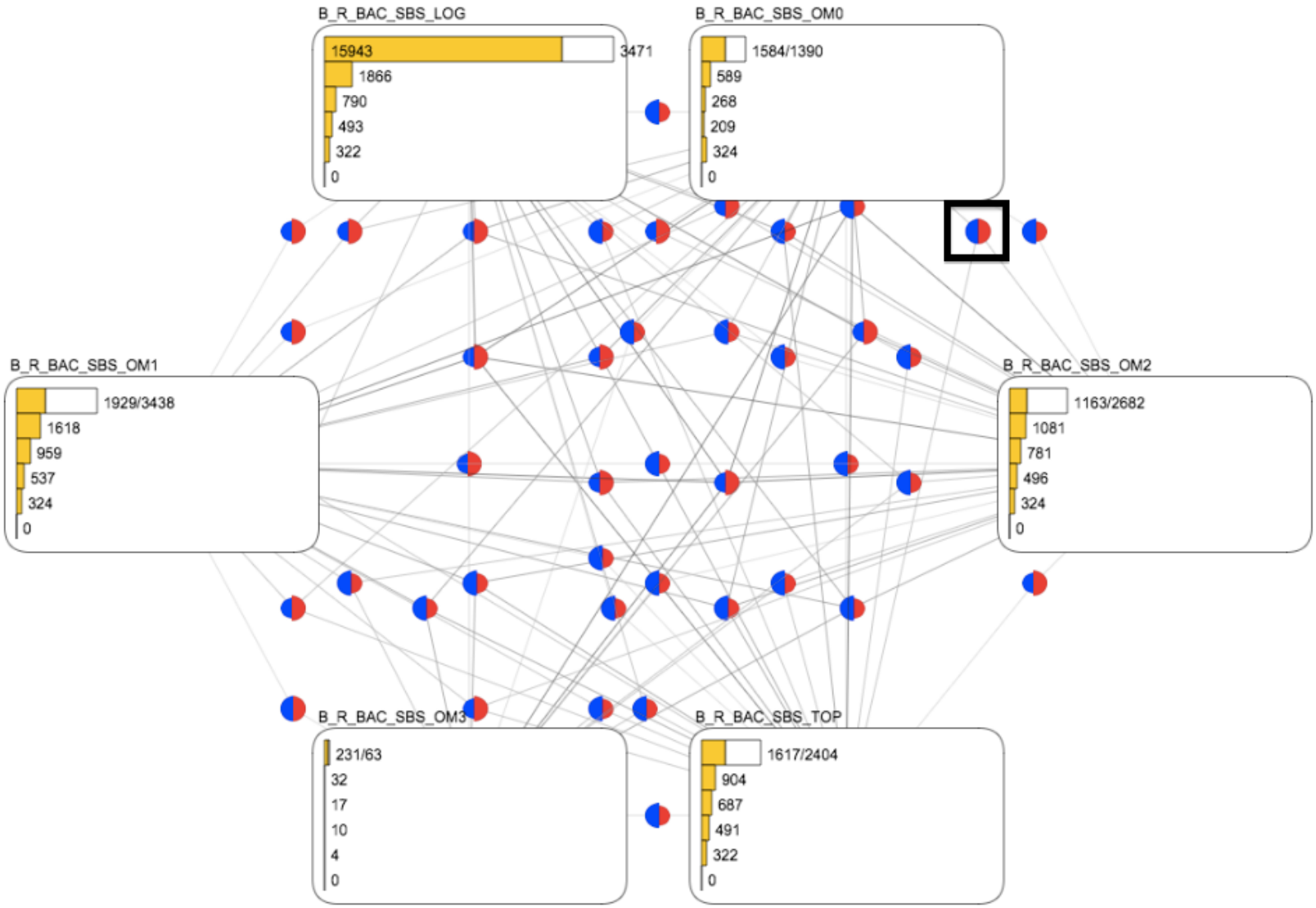
Network Glyph

- Network glyphs are meant to show information about how many edges are shared between 0 to n networks

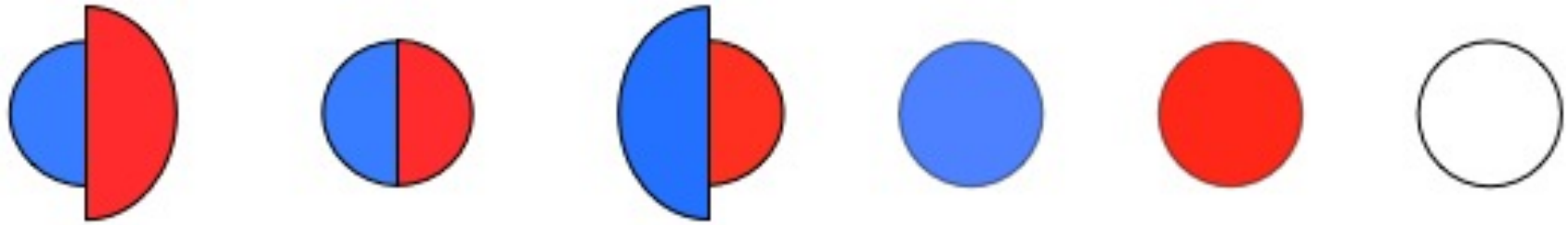


- Meant as secondary information about the network

Step Two: Display!



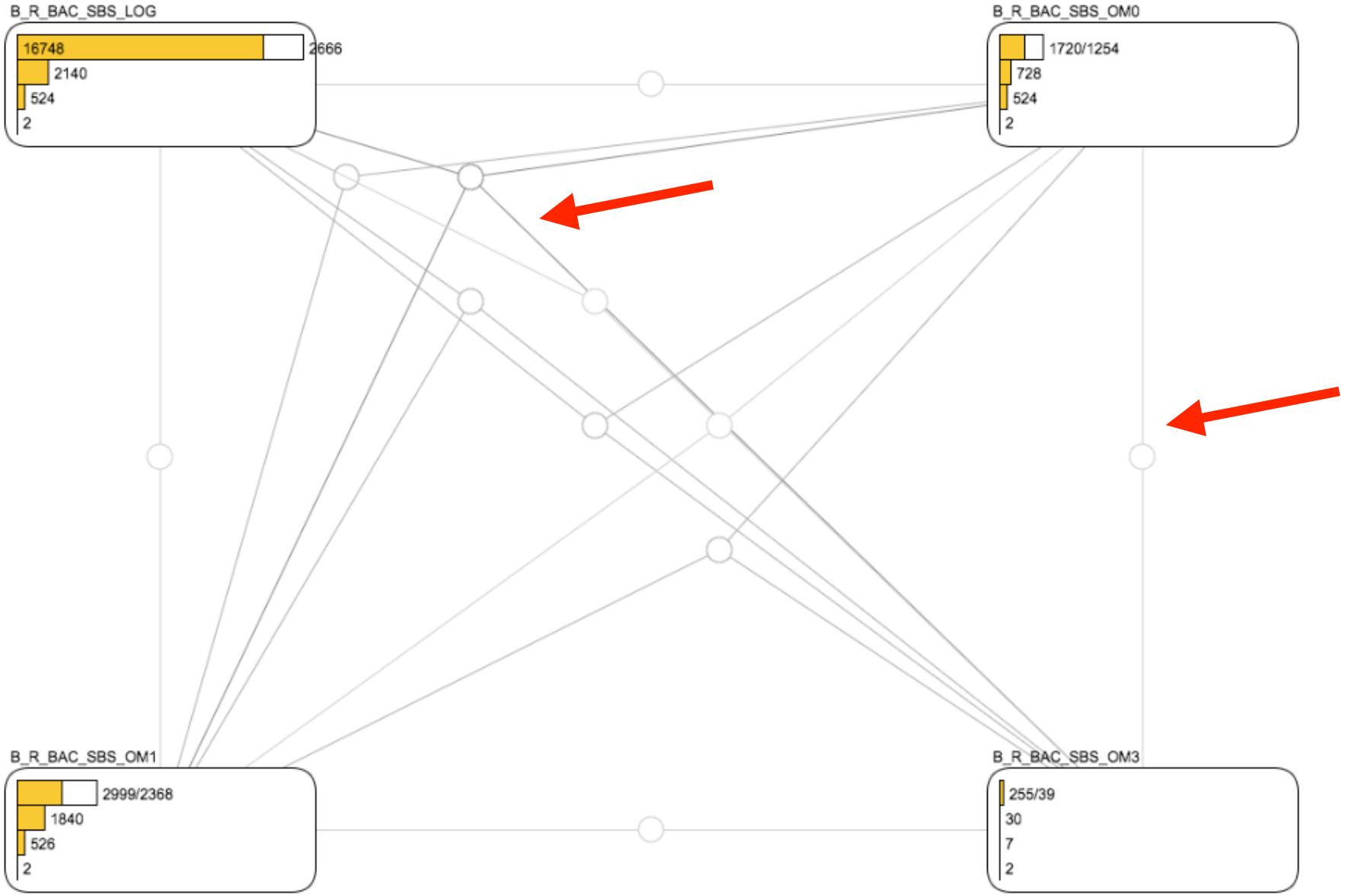
Step Two: Display!



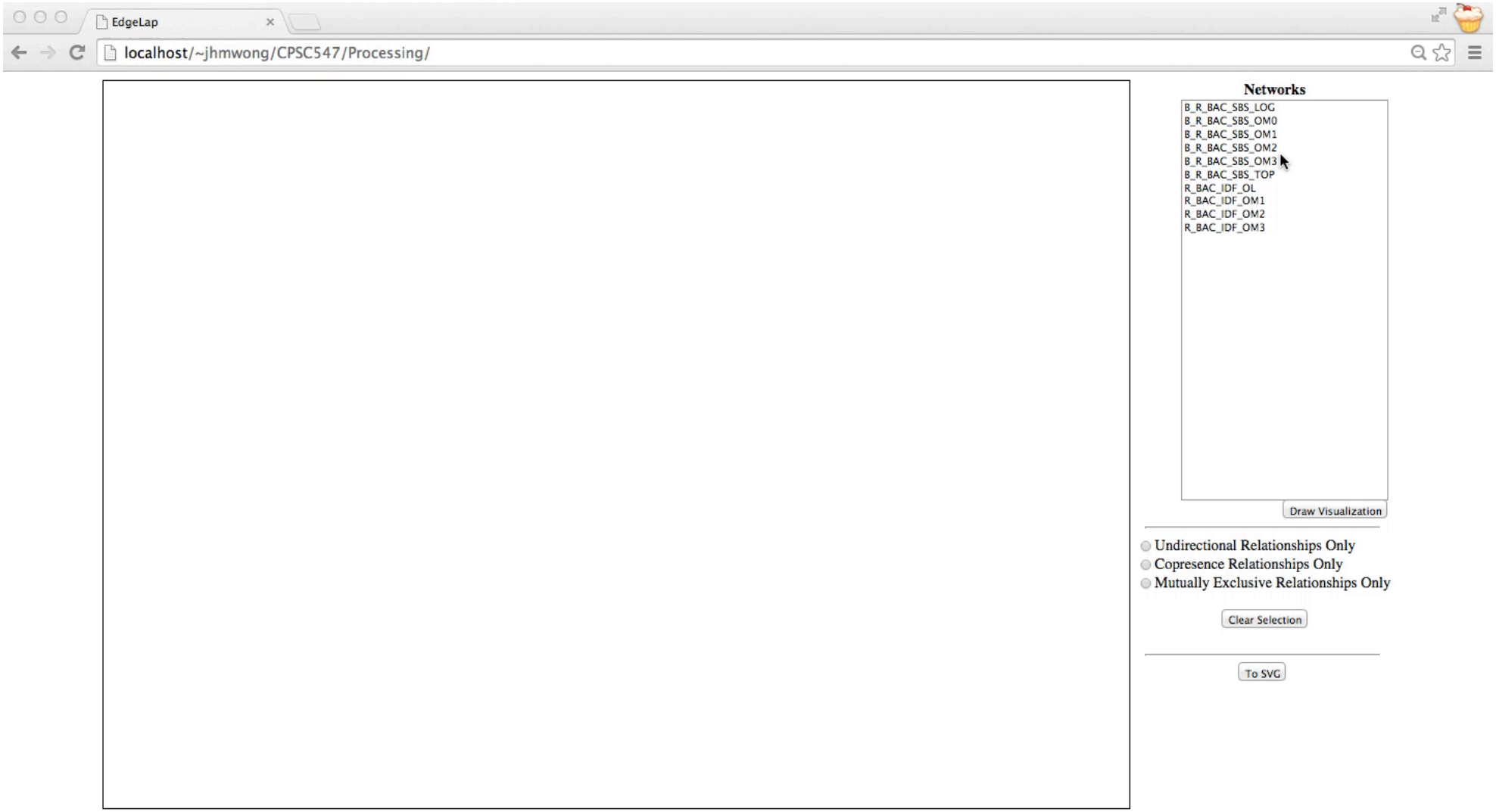
- Used to show properties of set we are looking at

- Blue = copresence (positive correlation)
- Red = mutual exclusion (negative correlation)
- White doesn’t take the type of interaction into account

Step Two: Display!



Step Two: Display!



Future Work

- Filtering option
- Hovering over the circle glyphs will highlight associated lines, display summary information
- Better method to draw lines to avoid too much overlap
- Improve performance

Q&A

- Questions?