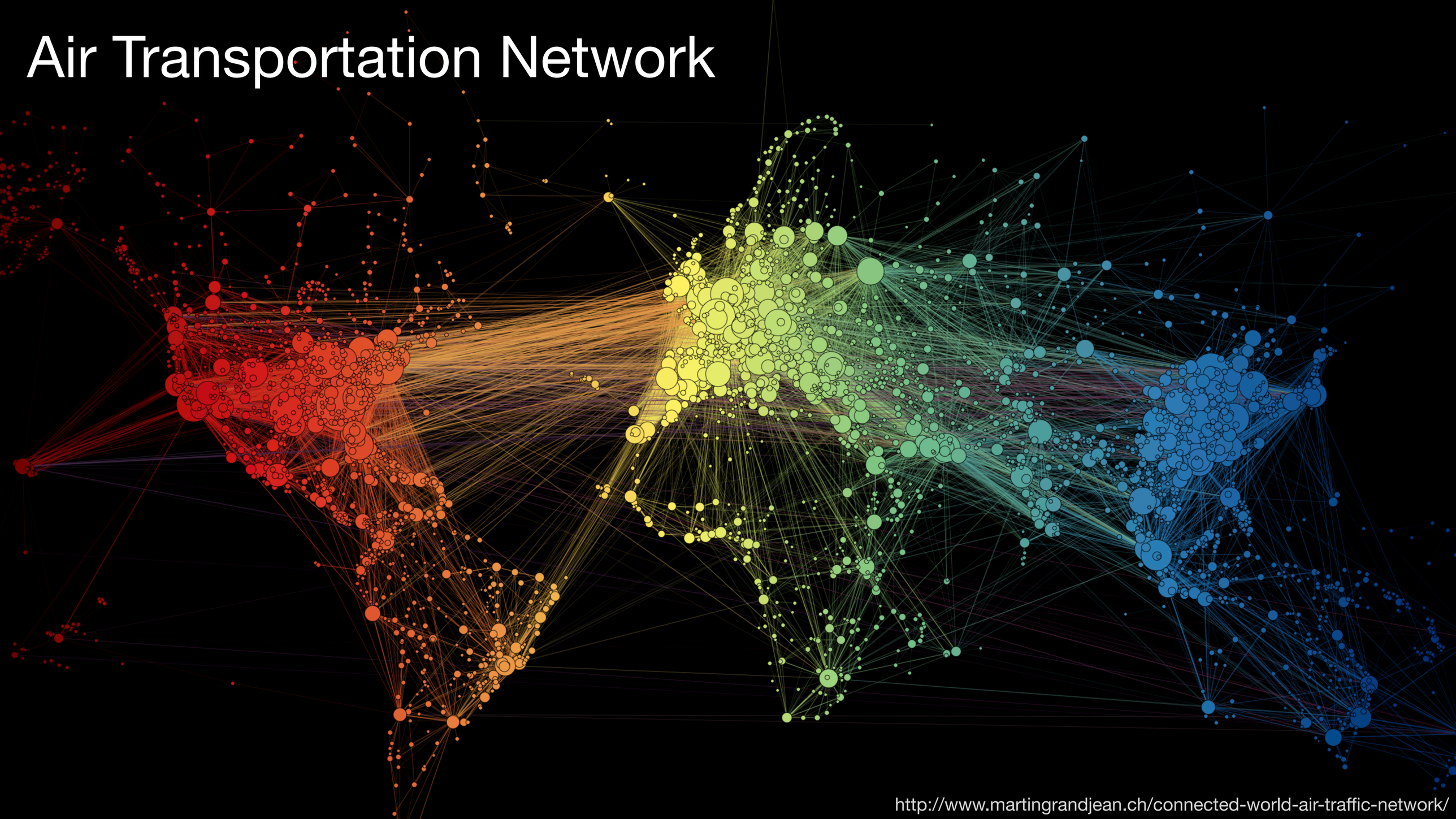
An Introduction to Graph Mining

Procheta Sen

UNIVERSITY OF LIVERPOOL

Networks are everywhere

Air Transportation Network

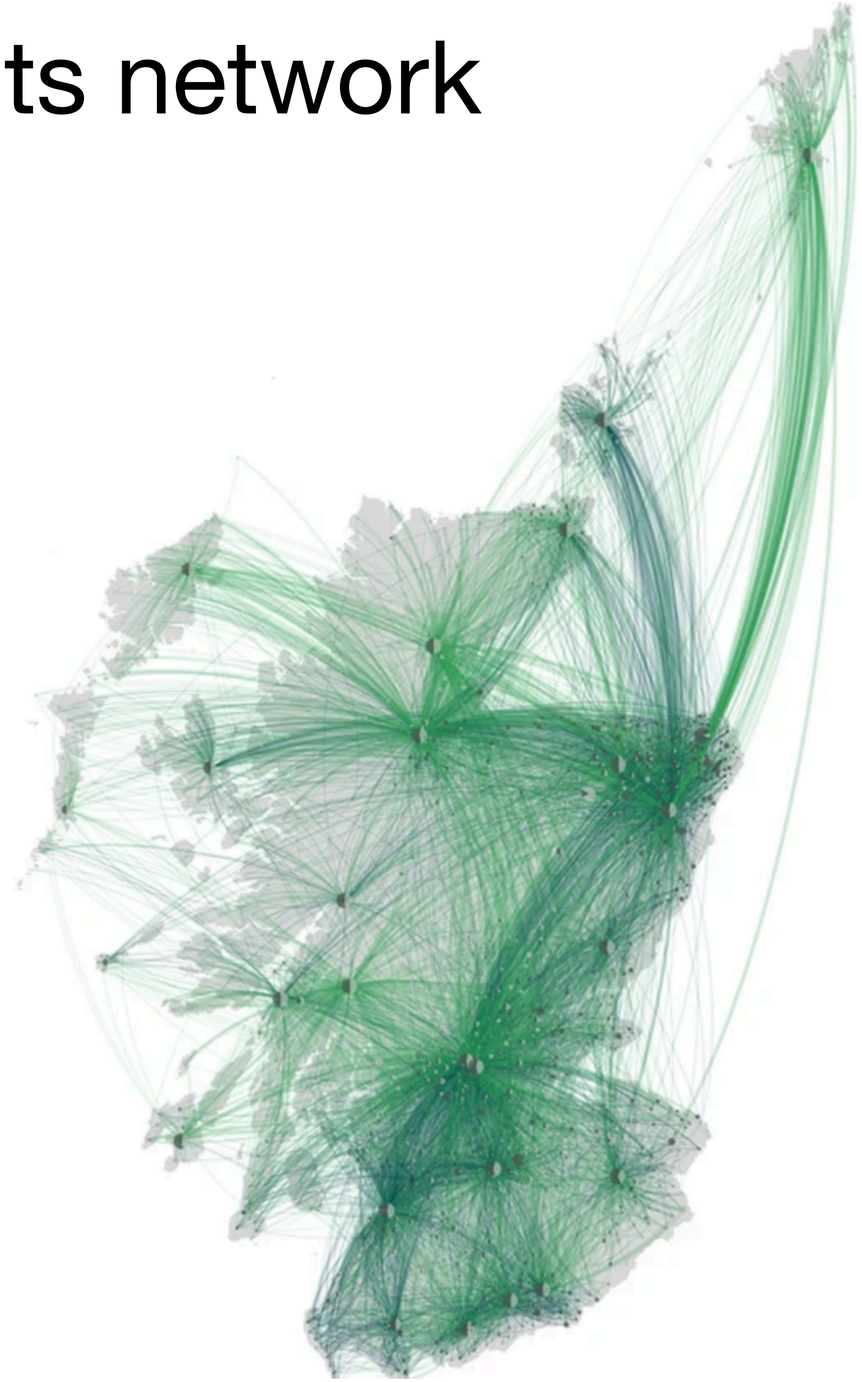


http://www.martingrandjean.ch/connected-world-air-traffic-network/

Social Networks

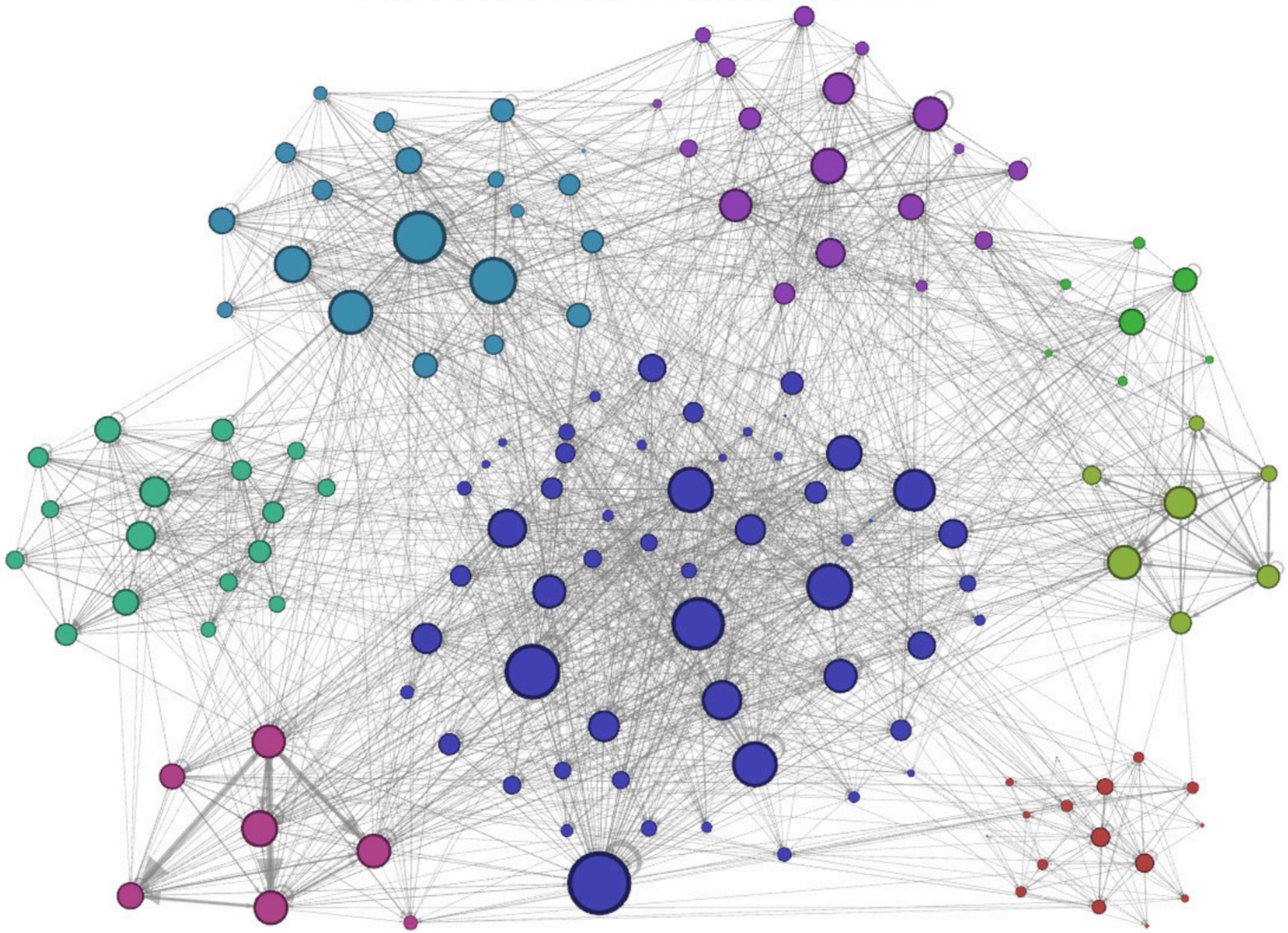
https://www.euroscientist.com/imagine-a-social-network-like-facebook-with-no-facebook/

Cattle movements network



Ruget, Anne-Sophie, et al. "Multi-species temporal network of livestock movements for disease spread." *Applied Network Science* 6.1 (2021): 1-20.

Email Exchange Networks



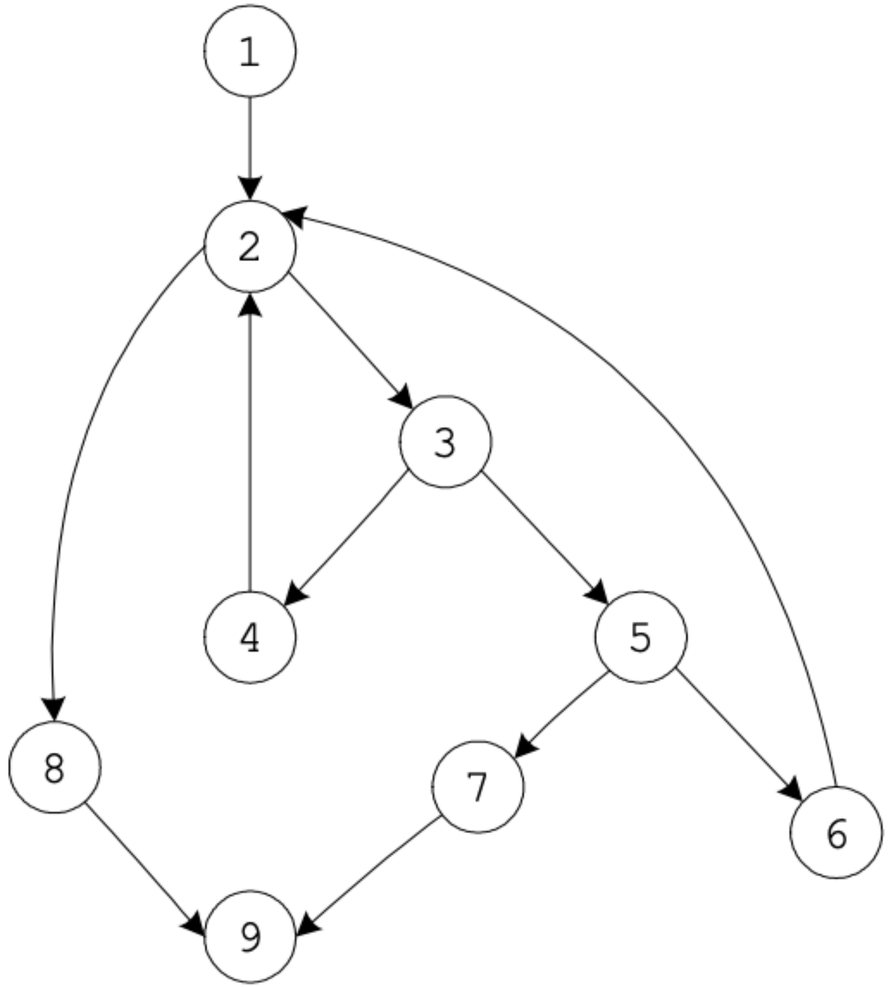
https://www.linkedin.com/pulse/visualizing-email-networks-damon-wade/

Program flow graph

Source Program:

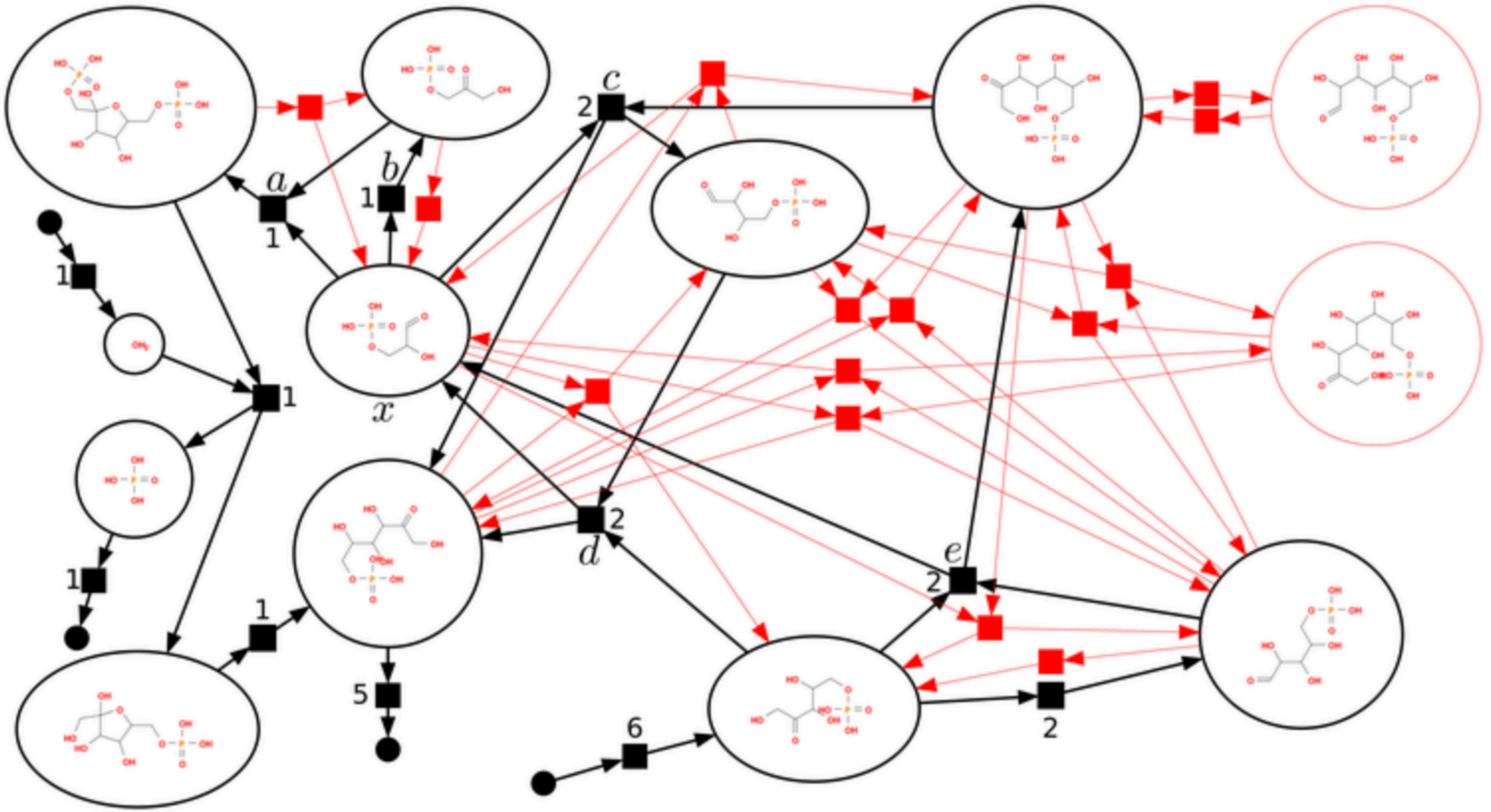
```
int binsearch(int x, int v[], int n)
{
     | int low, high, mid;
   1 | low = 0;
     | high = n - 1;
       while (low <= high) | 2
       {
           3 | mid = (low + high)/2;
             | if (x < v[mid])
                       high = mid - 1; | 4
           5 | else if (x > v[mid])
                       low = mid + 1;  | 6
           7 | else return mid;
       }
       return -1; | 8
} | 9
```

CFG:



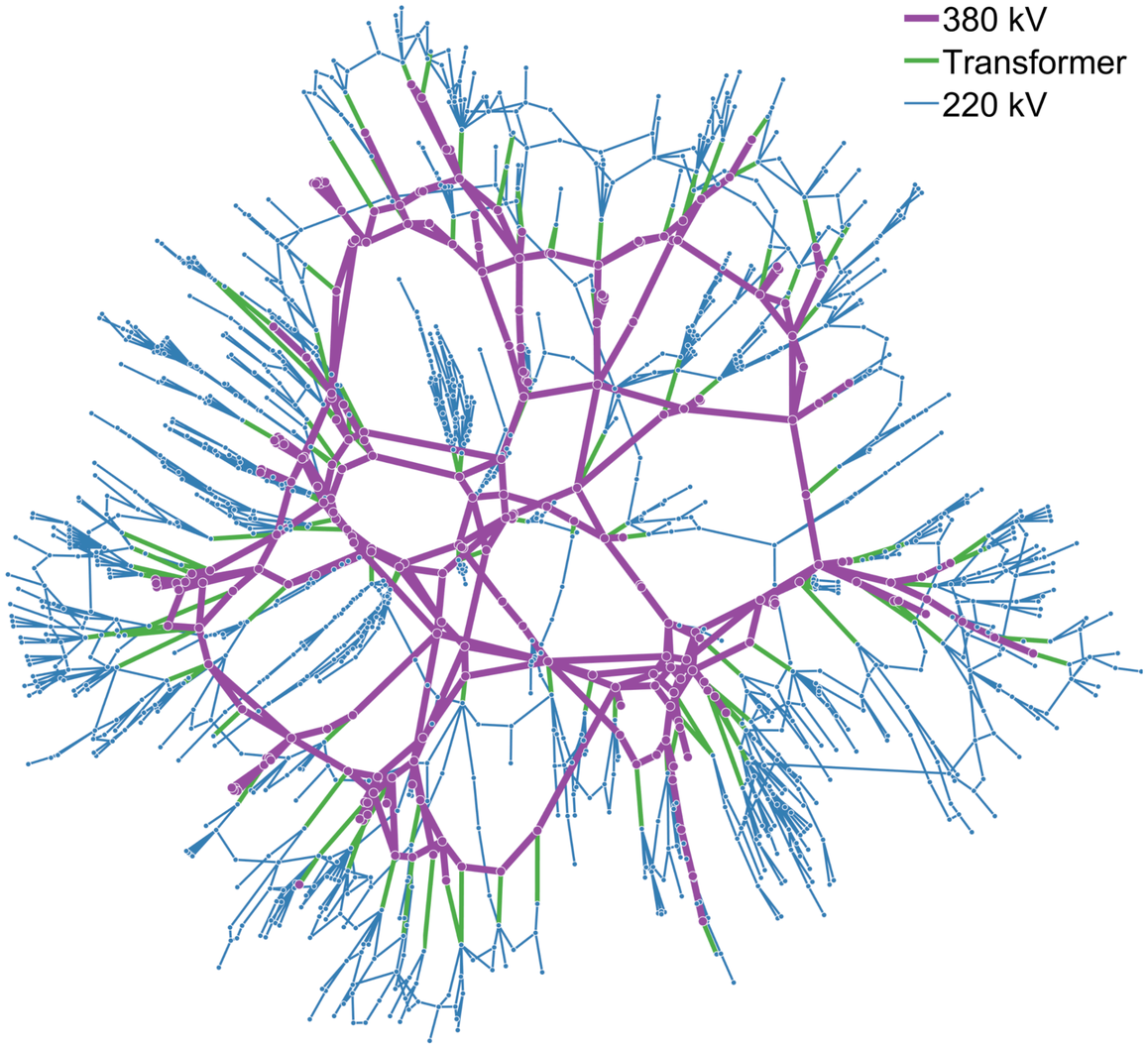
Al-Ekram, Raihan, and Kostas Kontogiannis. "Source code modularization using lattice of concept slices." *Eighth European Conference on Software Maintenance and Reengineering, 2004. CSMR 2004. Proceedings.*. IEEE, 2004.

Chemical Reaction Networks



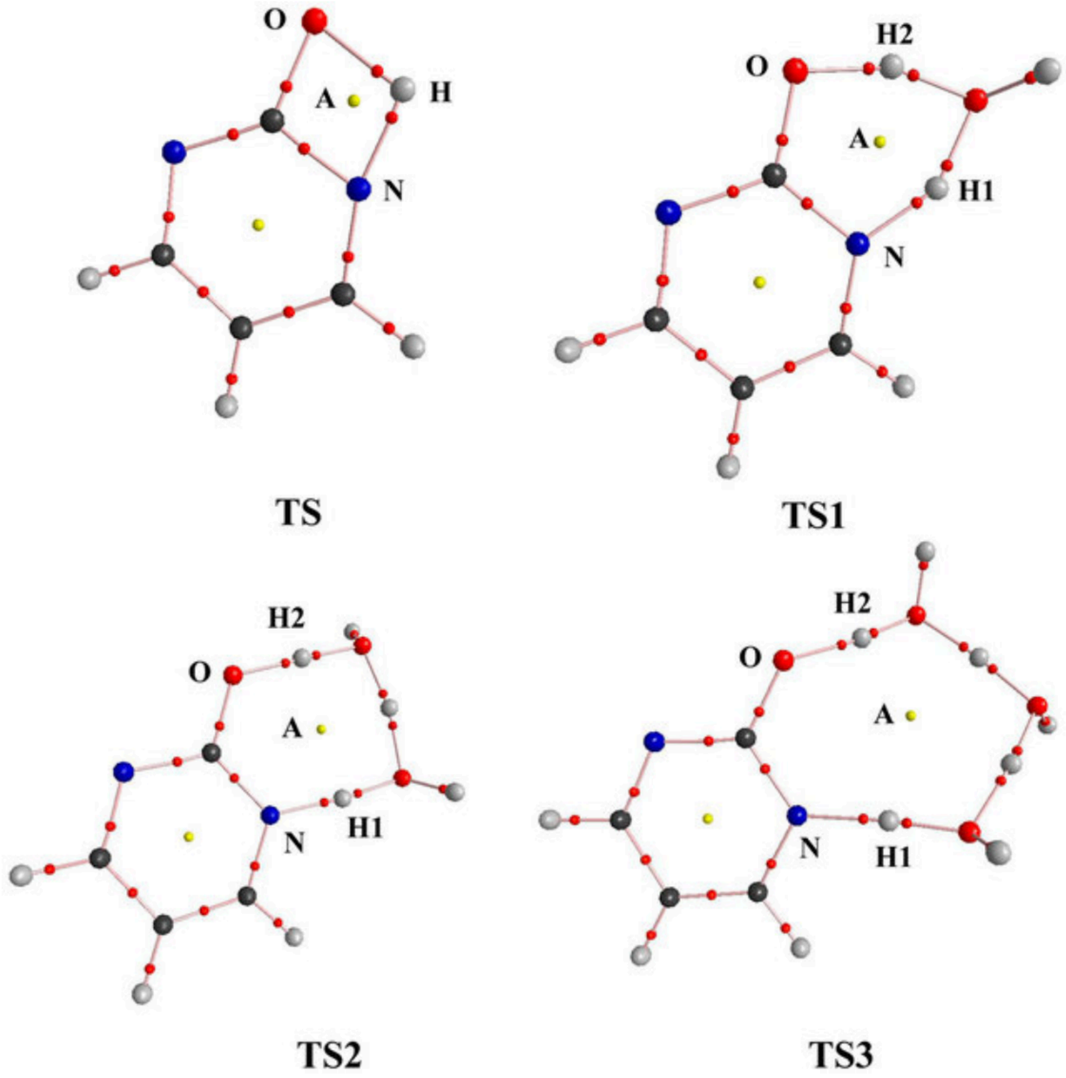
https://www.mis.mpg.de/stadler/research/chemical-reaction-networks.html

Power network (electrical grid)



https://en.wikipedia.org/wiki/Electrical_grid

Molecular graph or chemical graph



Heidarpoor Saremi, L., Ebrahimi, A., & Lagzian, M. (2019). Substituent effects on direct and indirect tautomerism of pyrimidin-2 (1H)-one/pyrimidin-2-ol. *Molecular Simulation*, *45*(1), 58-67.

Graph mining topics

- Graph Classification
- Graph Clustering
- Graph Pattern Mining
- Graph Compression
- Graph Dynamics
- **Social Network Analysis**
- Graph Visualisation
- Link Analysis
- ...

Some of possible settings

Database of many small graphs

- **Examples of applications**: chemical and biological data, program flow analysis
- **Examples of tasks**: graph pattern mining, graph classification, graph clustering

A single large graph

- **Examples of applications**: web graph analysis, social network analysis, organisational network analysis, transportation network analysis, etc.
- **Examples of tasks**: community detection, influential nodes identification, node ranking, link prediction