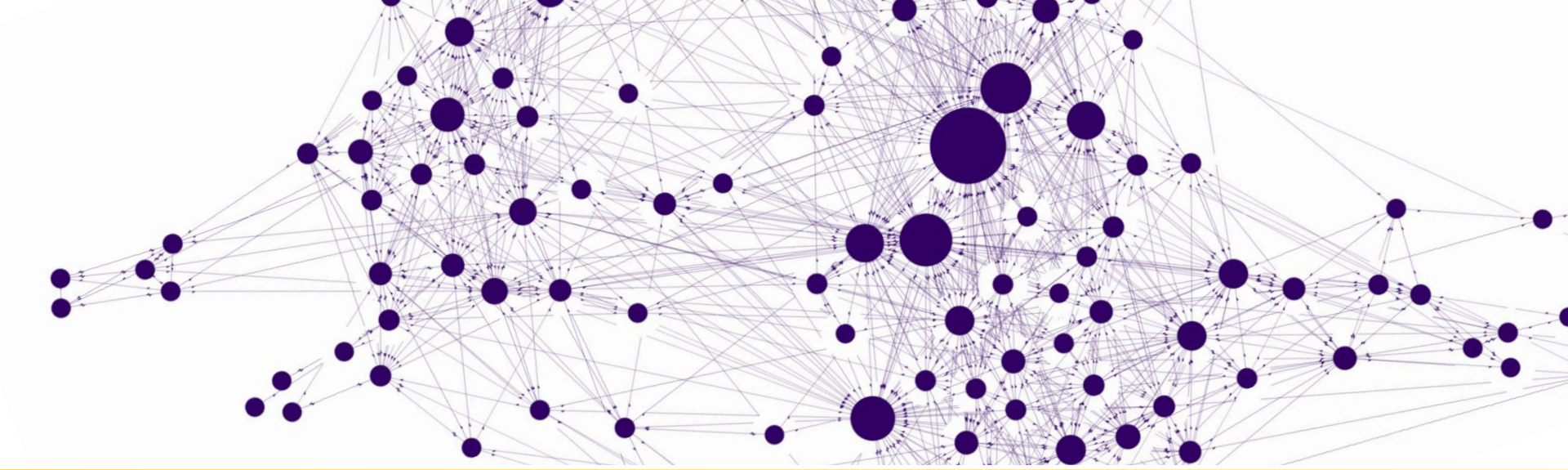


Transductive Classification on Heterogeneous Information Networks with Edge Betweenness-based Normalization

P. Bangcharoensap & T. Murata

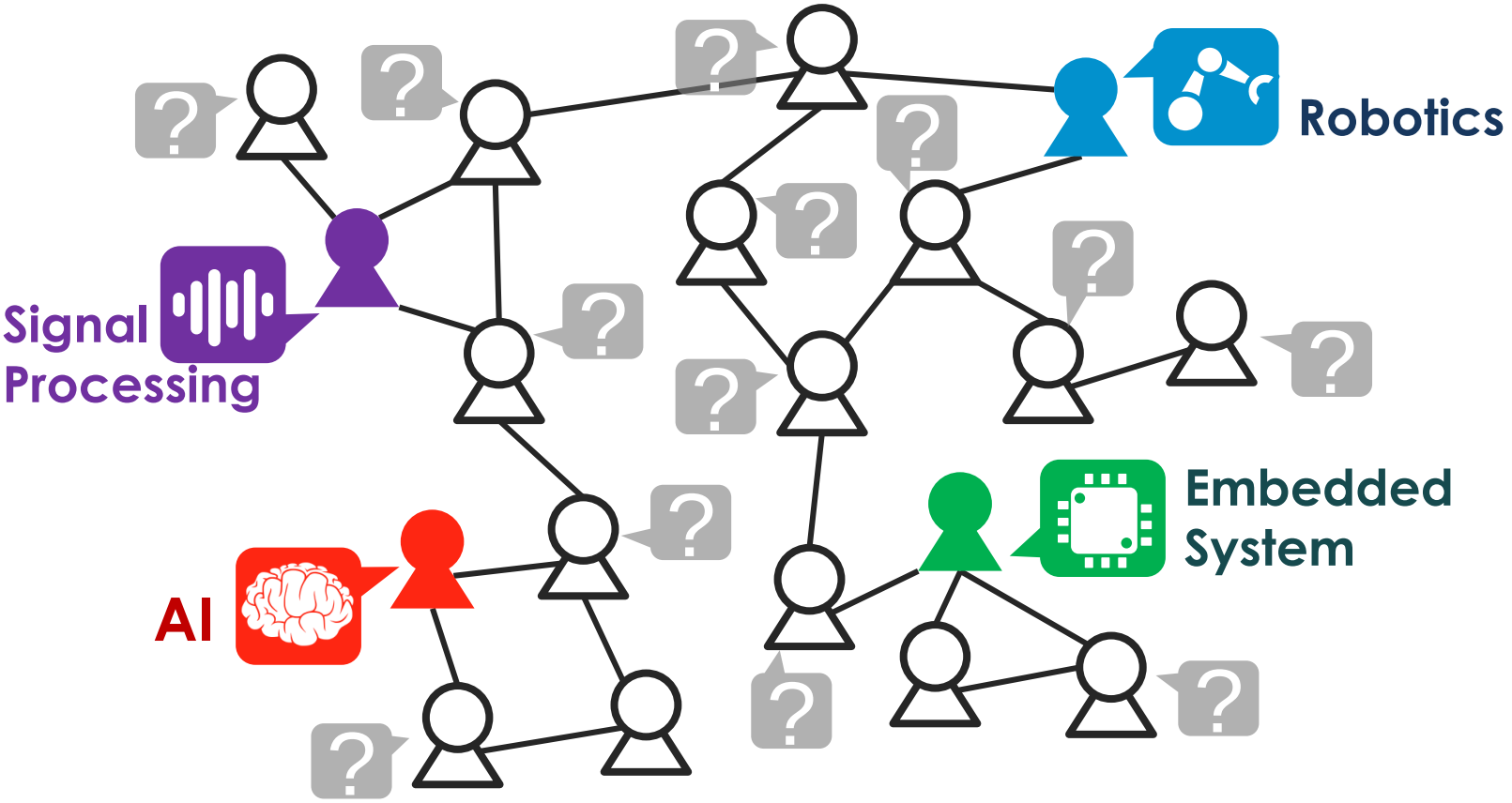
東京工業大学
Tokyo Institute of Technology

H. Kobayashi & N. Shimizu

YAHOO! JAPAN

Transductive Classification on Network

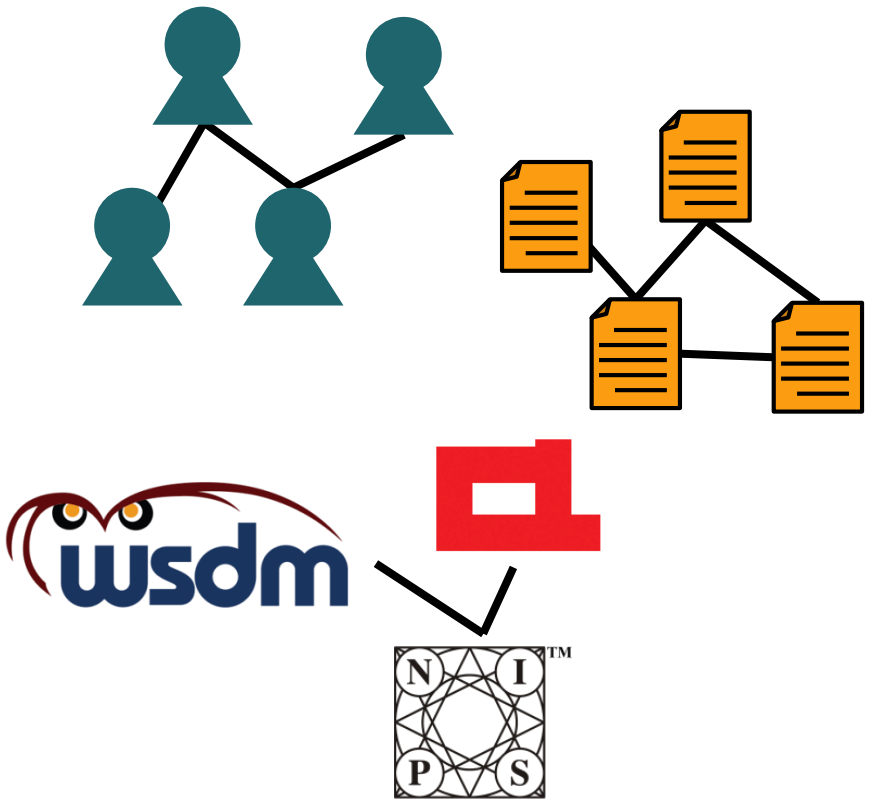
- Given a network and the labels of some vertices
- What is the labels of the remaining instances?
- Individuals tend to be linked with similar others.



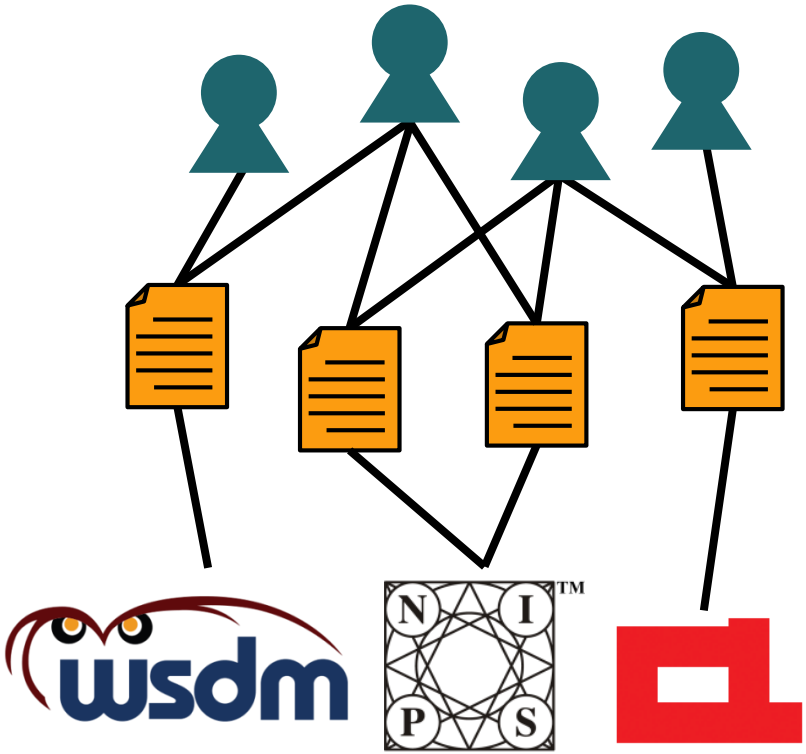
Heterogeneous Network

Networks containing multiple types of vertices

Homogeneous



Heterogeneous



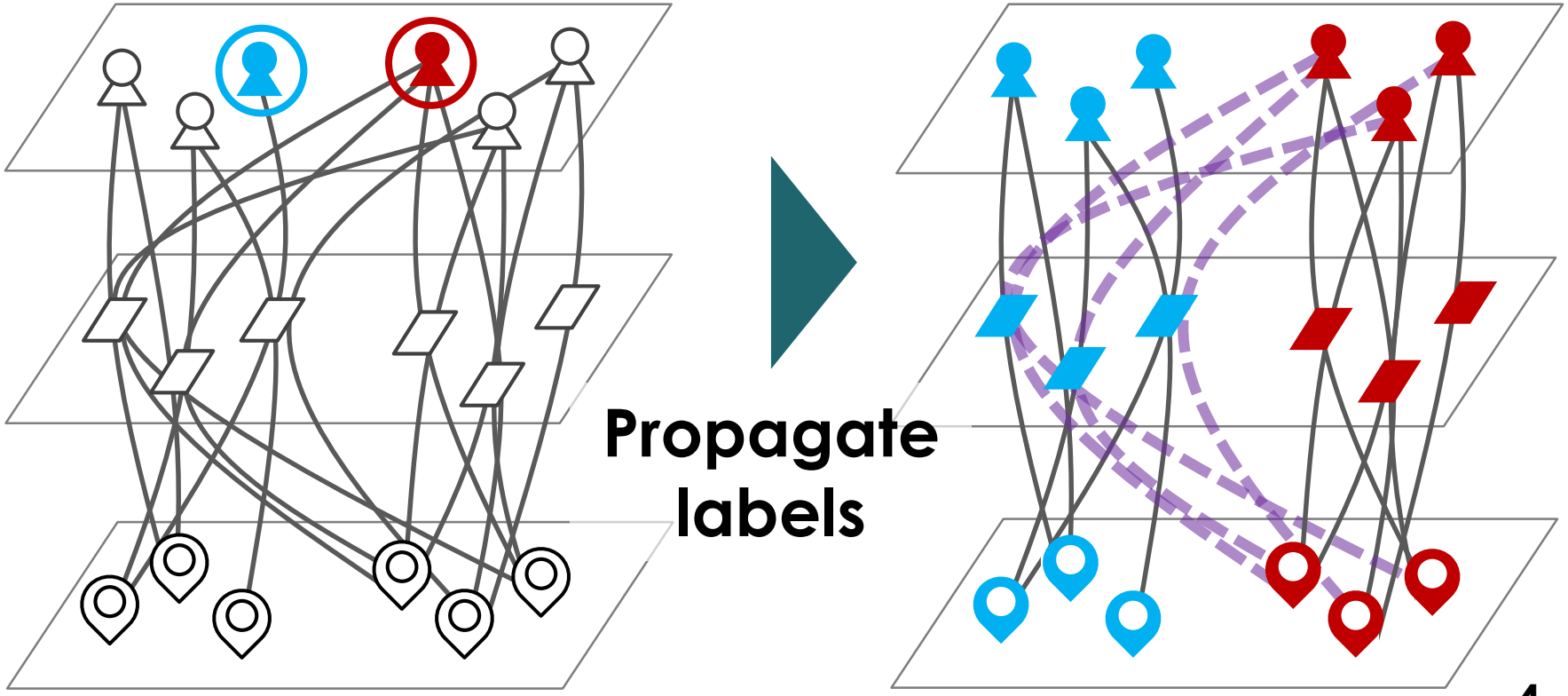
Label Propagation

Input

Heterogeneous network, G(V,E,W), and labeled vertices

Output

Labels of all vertices in the given network



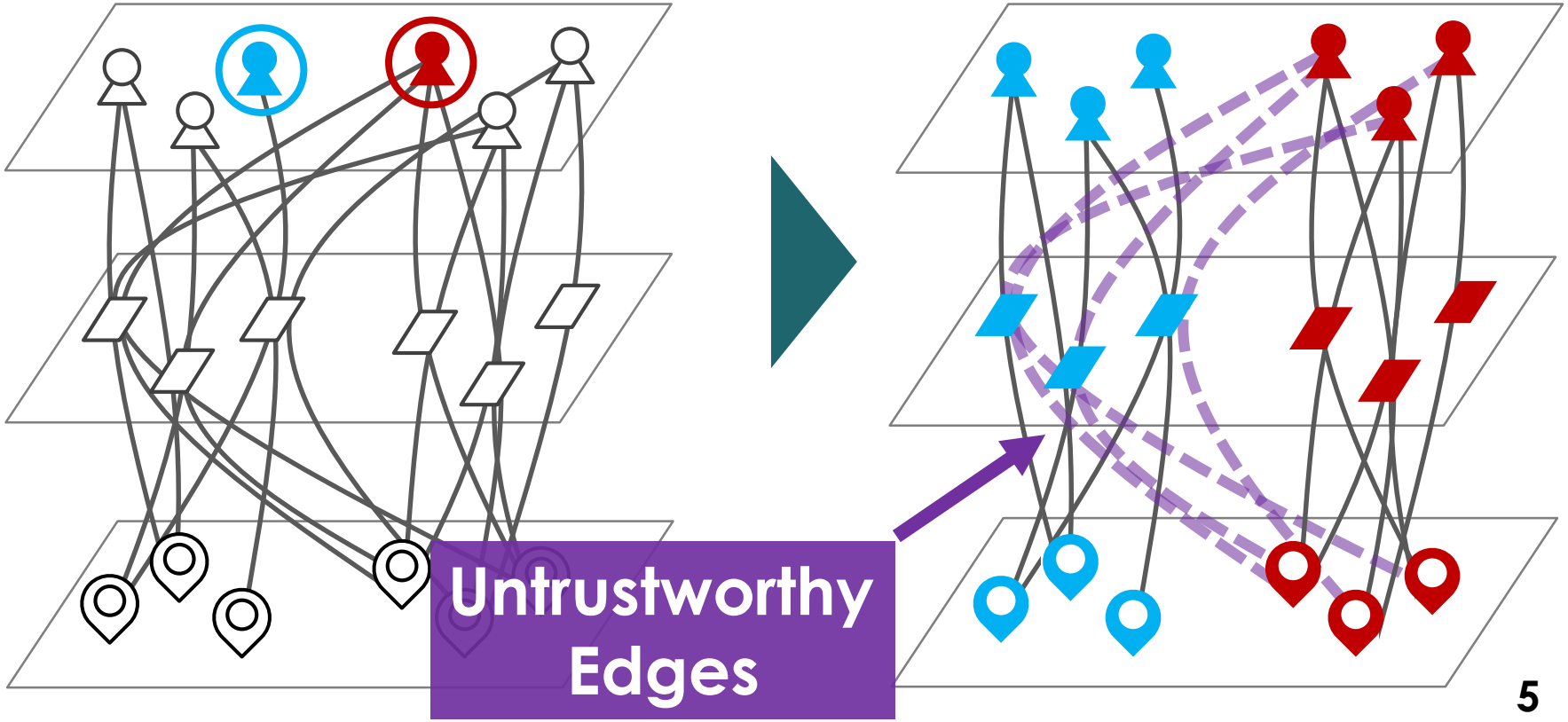
Challenge

Input

Heterogeneous network, G(V,E,W), and labeled vertices

Output

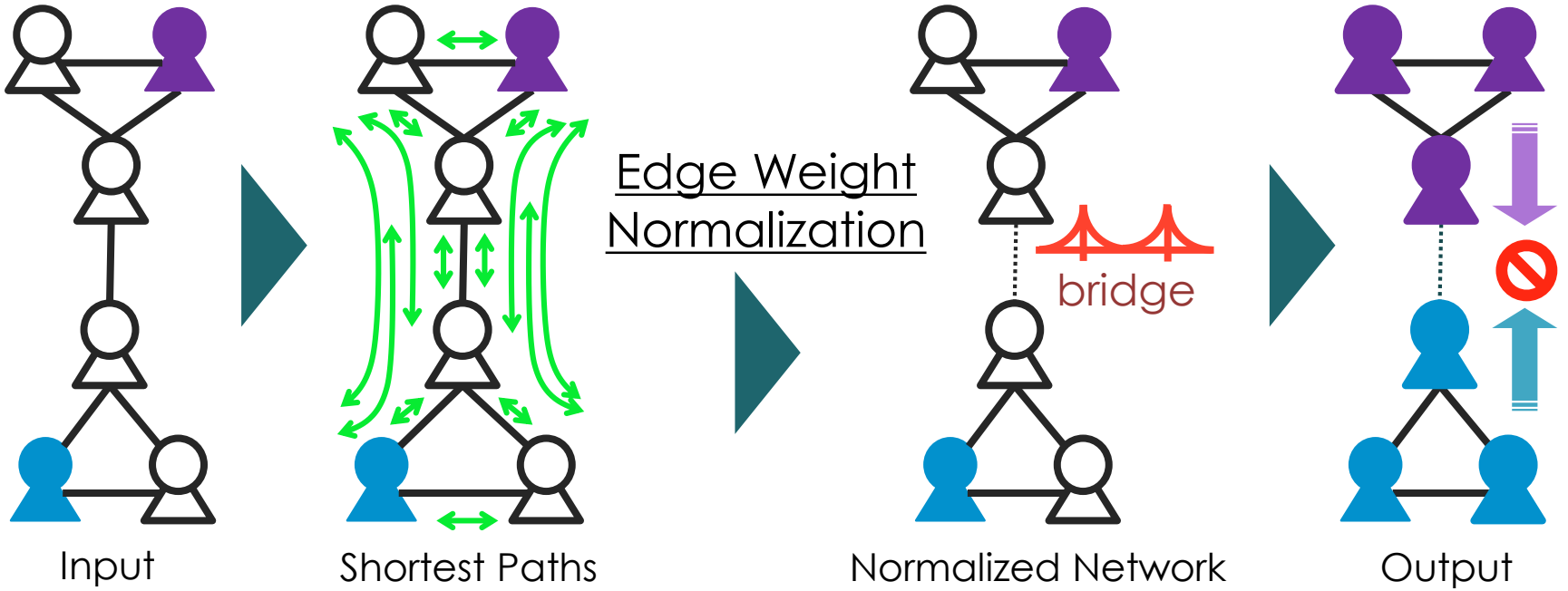
Labels of all vertices in the given network



Contributions

1. Edge Betweenness Centrality = Untrustworthiness
   - # of shortest paths between all vertex pairs passing through it (Gervan and Newman, 2002)
2. A novel definition of edge betweenness for heterogeneous networks

5 pp increase in accuracy



Thank you