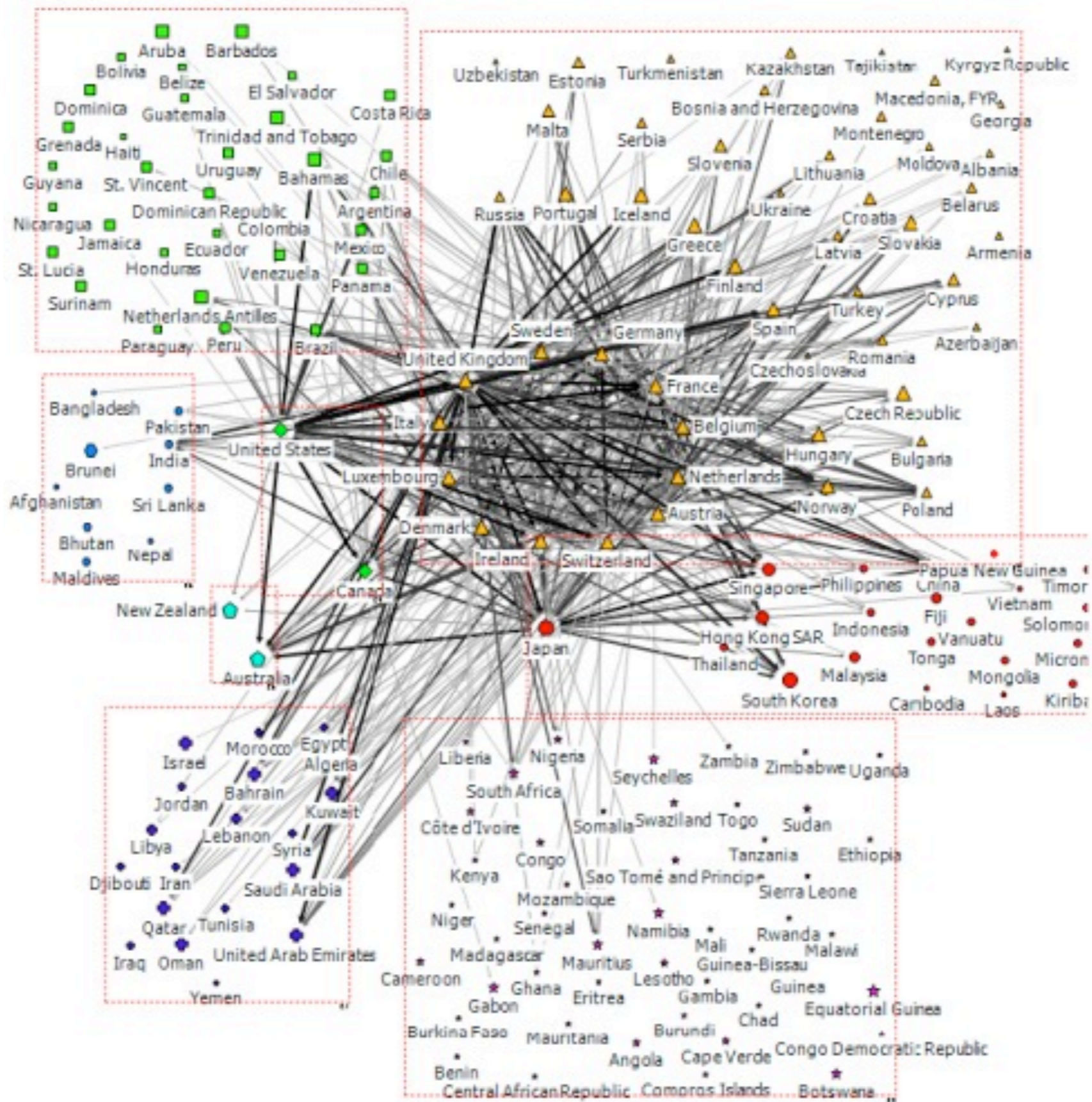


# discerning importance and structure

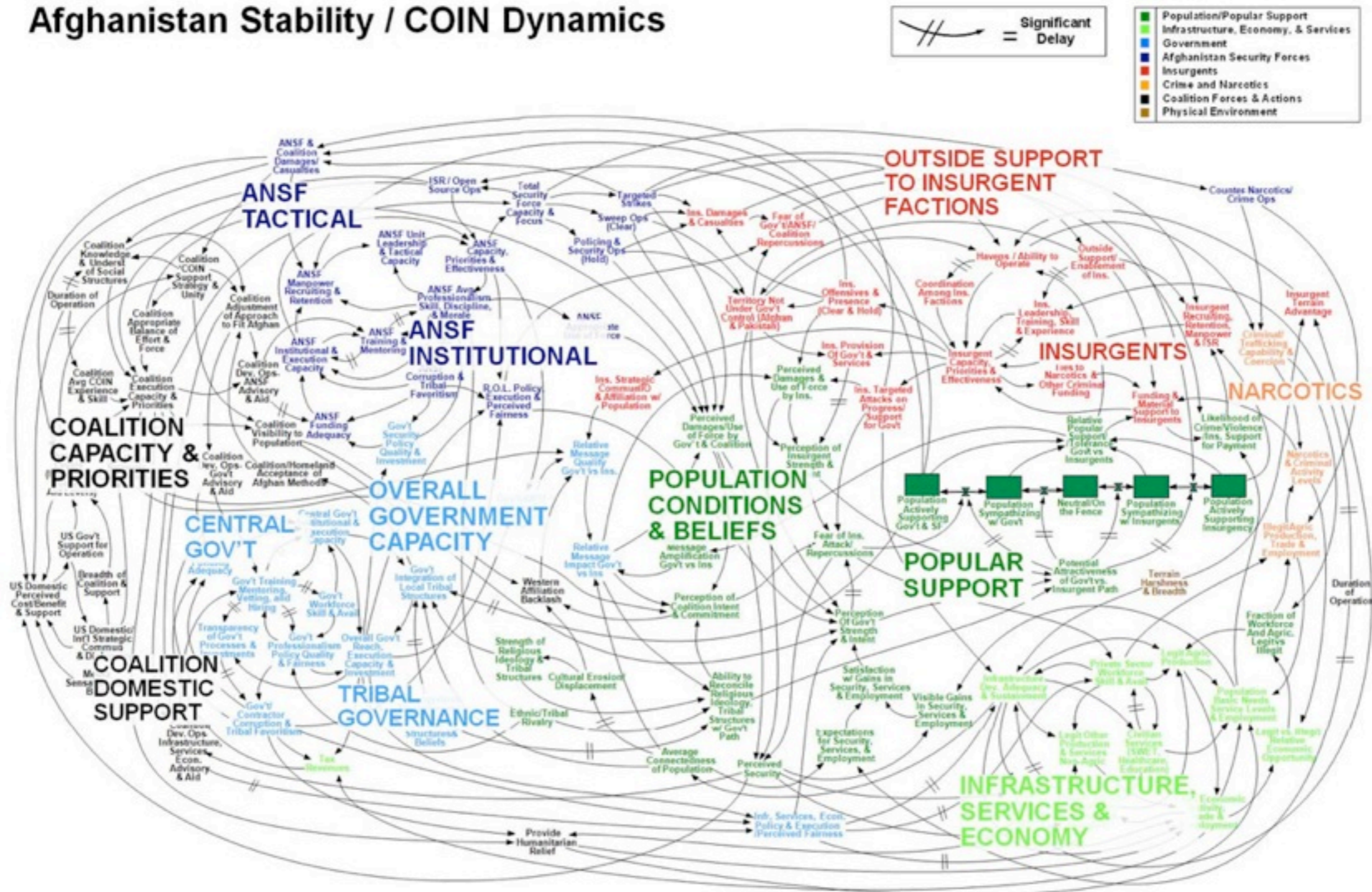
---

COMP 102  
Derek Ruths  
Oct 13, 2011



# what nodes are important?

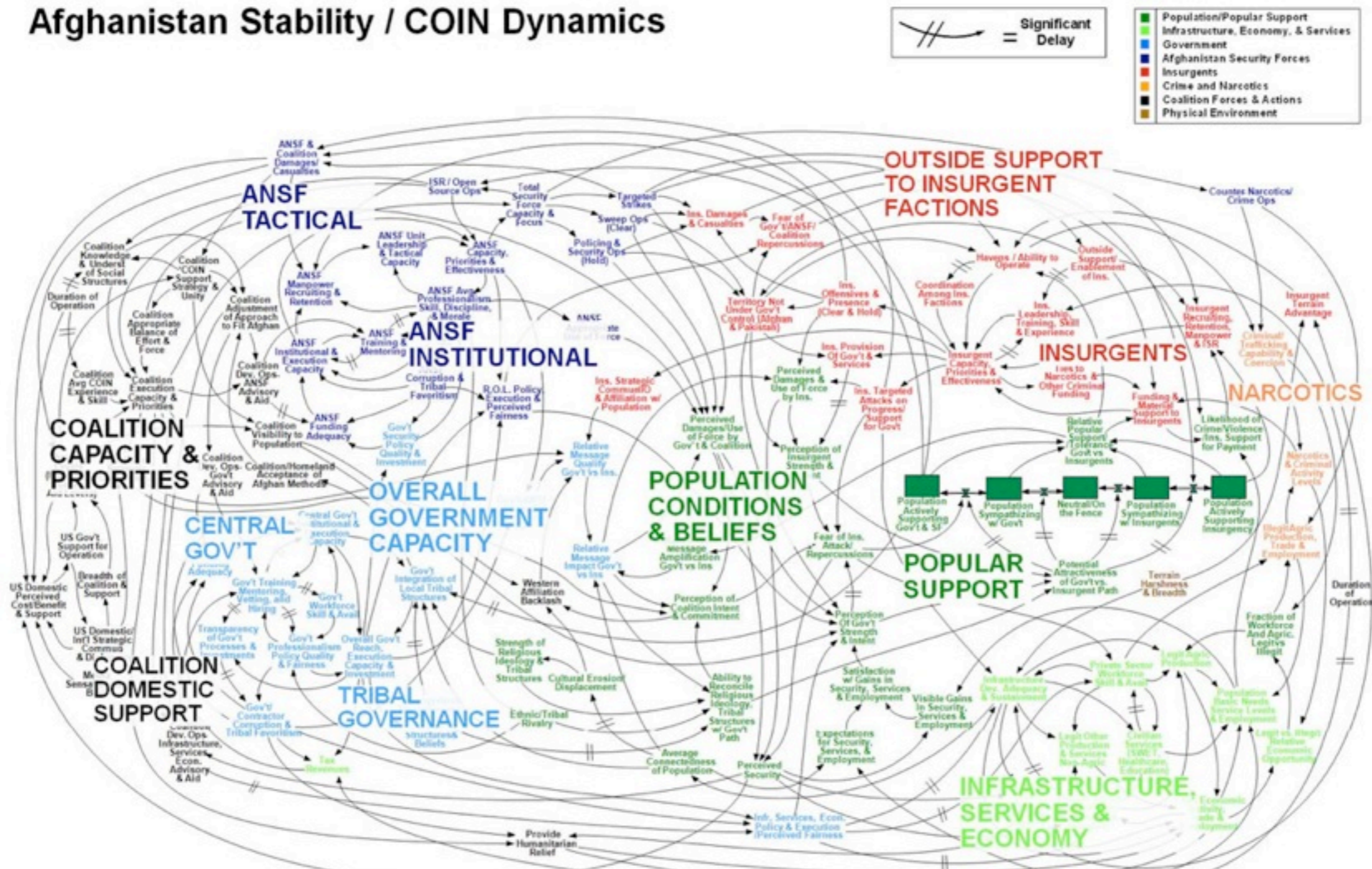
## Afghanistan Stability / COIN Dynamics



WORKING DRAFT - V3

# is there a sensible way of grouping nodes?

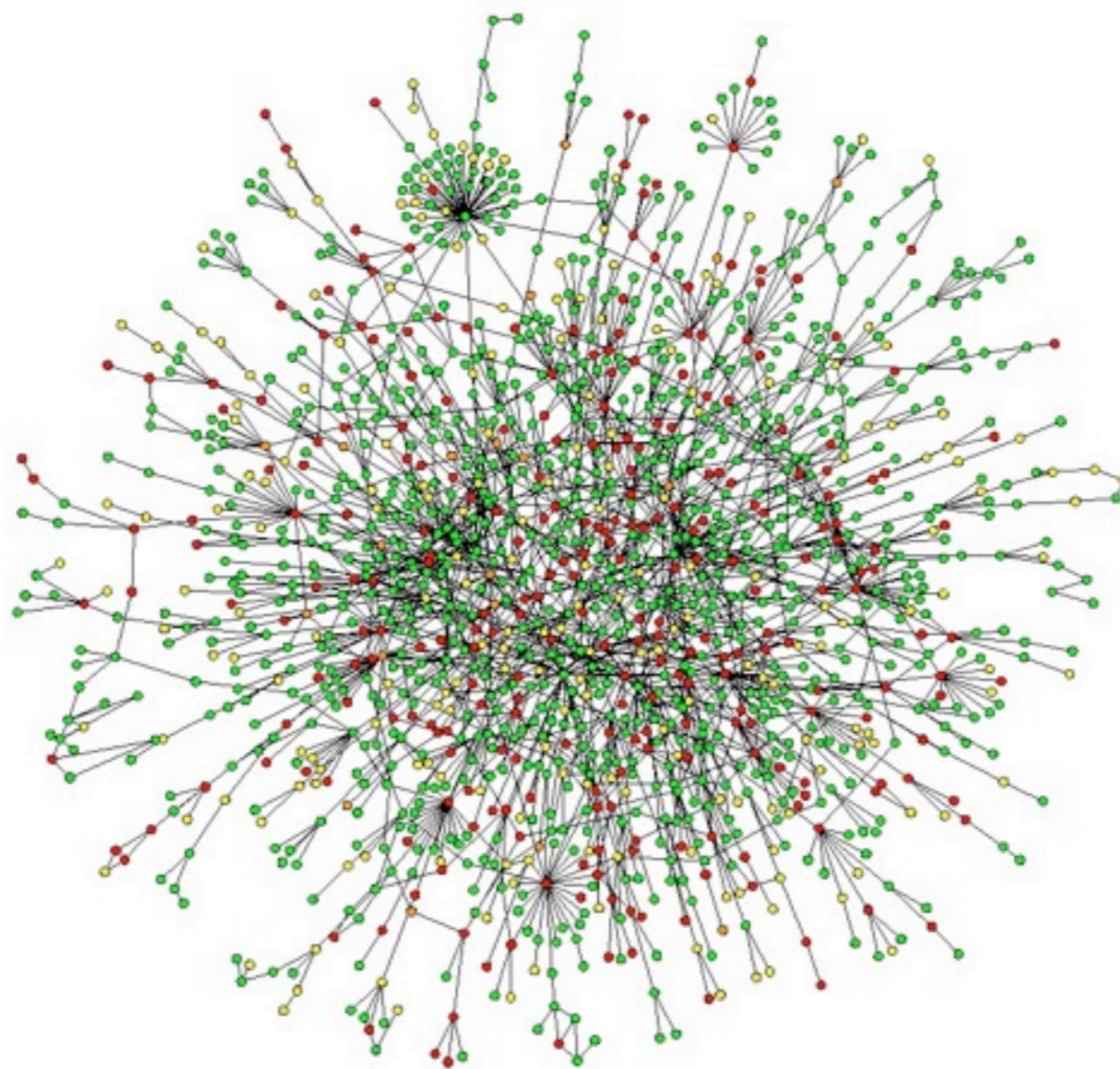
## Afghanistan Stability / COIN Dynamics



WORKING DRAFT - V3

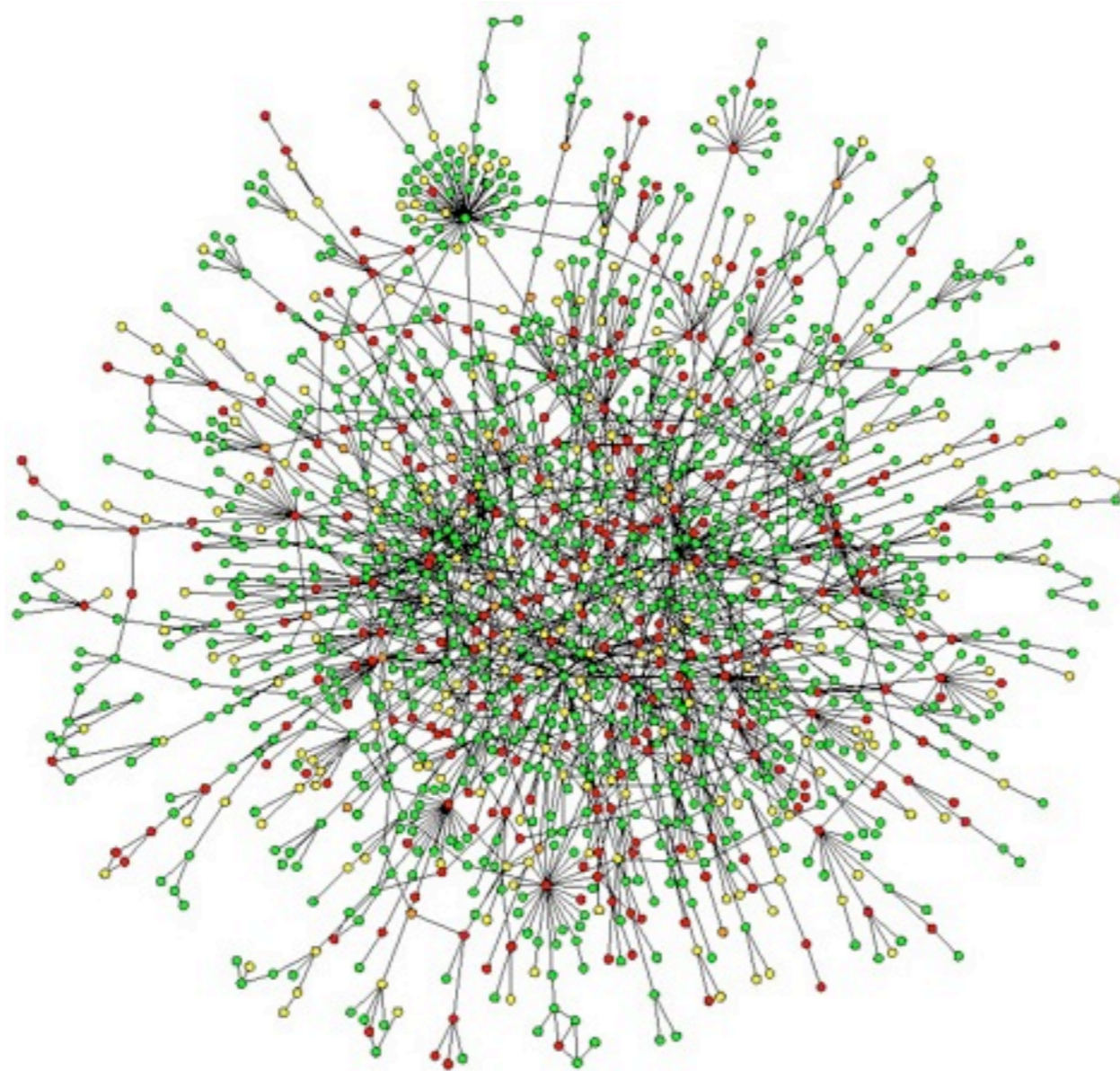
# what nodes are important?

---



is there a sensible way of grouping nodes?

---



discerning important nodes = *centrality*

discerning node groups = *community structure*

# the undirected network

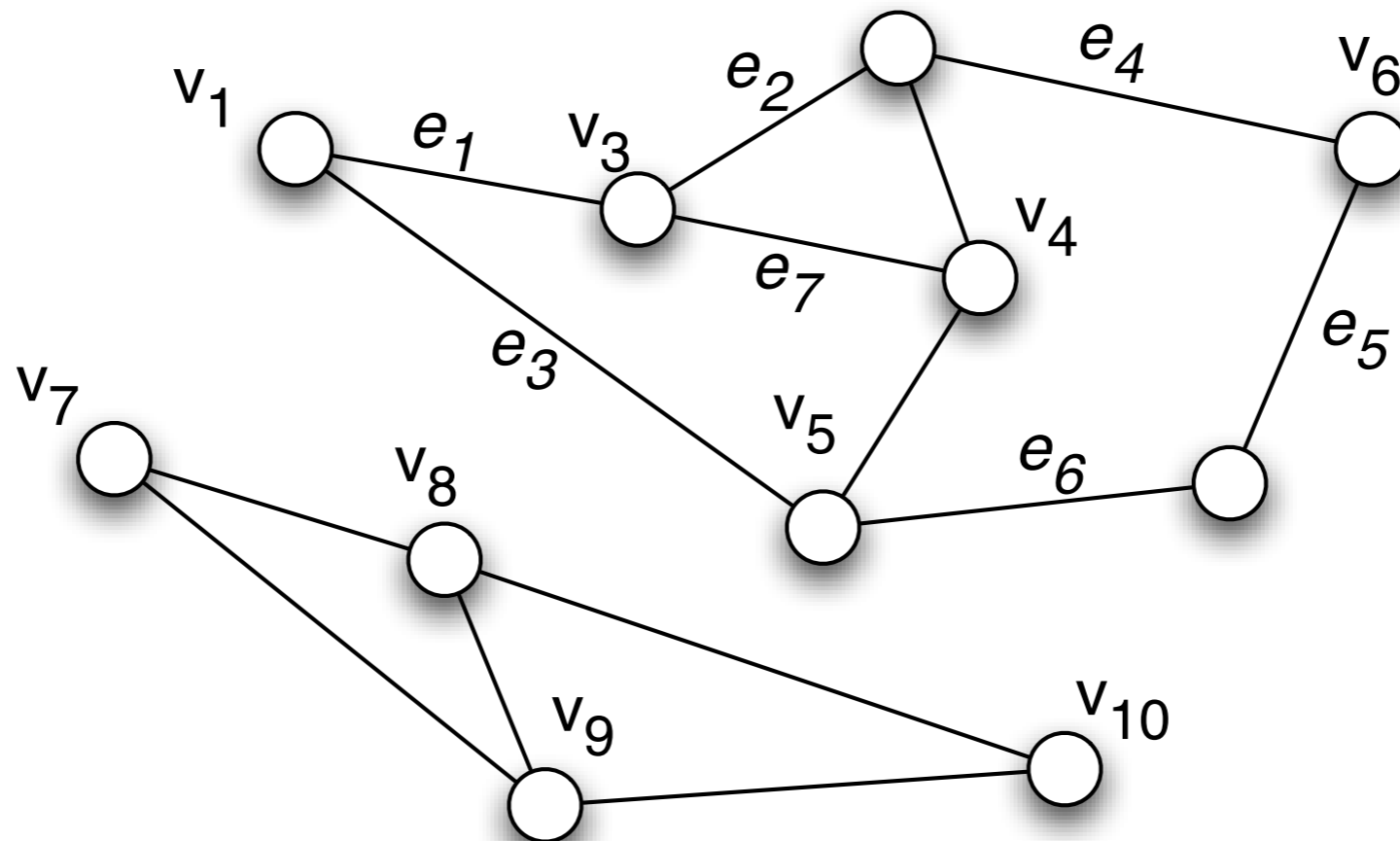
---

$$G = (V, E)$$

“nodes”

edges

Text<sub>2</sub>





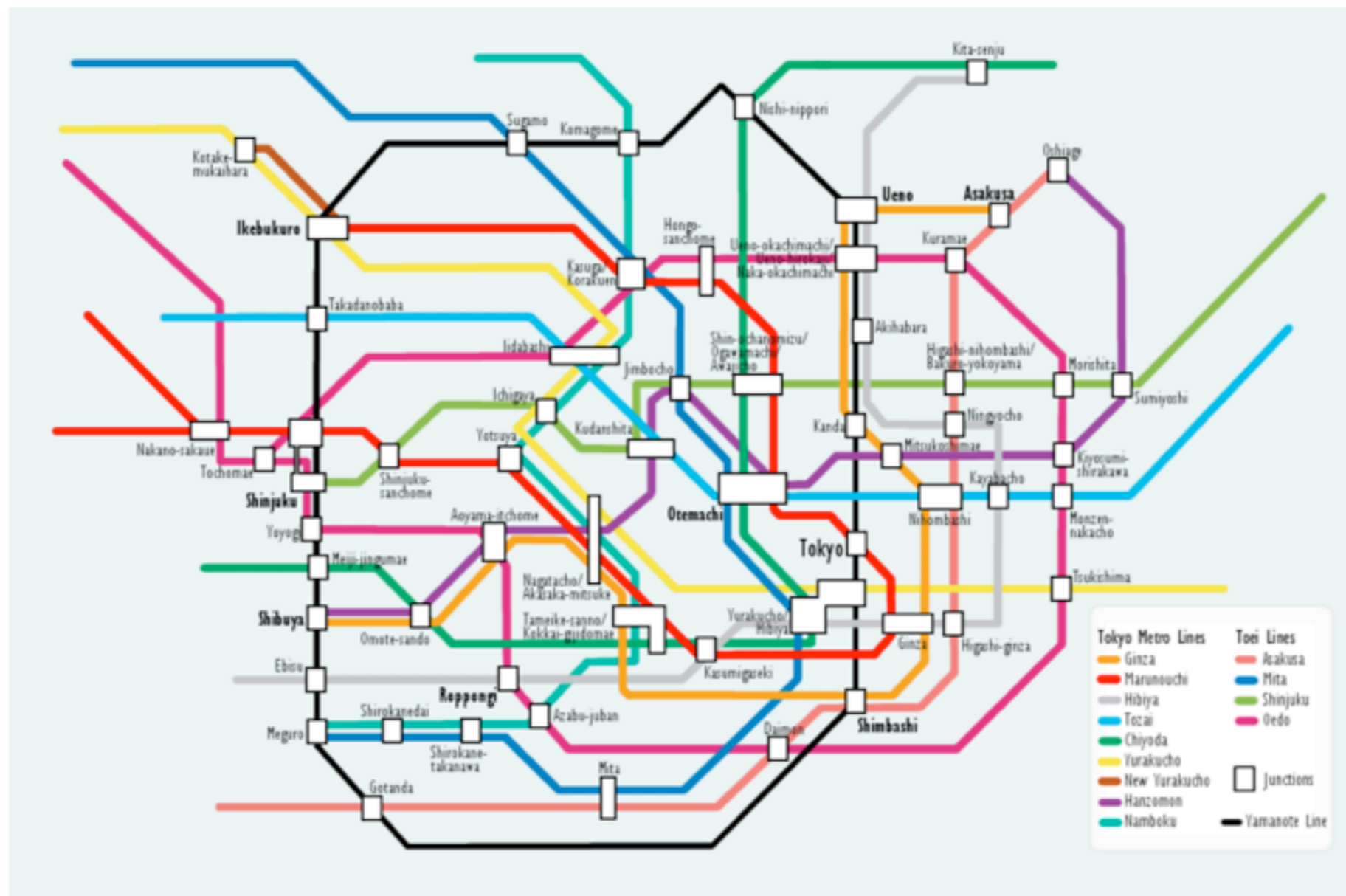
# degree-based centrality

---

- degree centrality
- Katz prestige

# degree centrality

*the number of neighbors a node has often implies its importance*

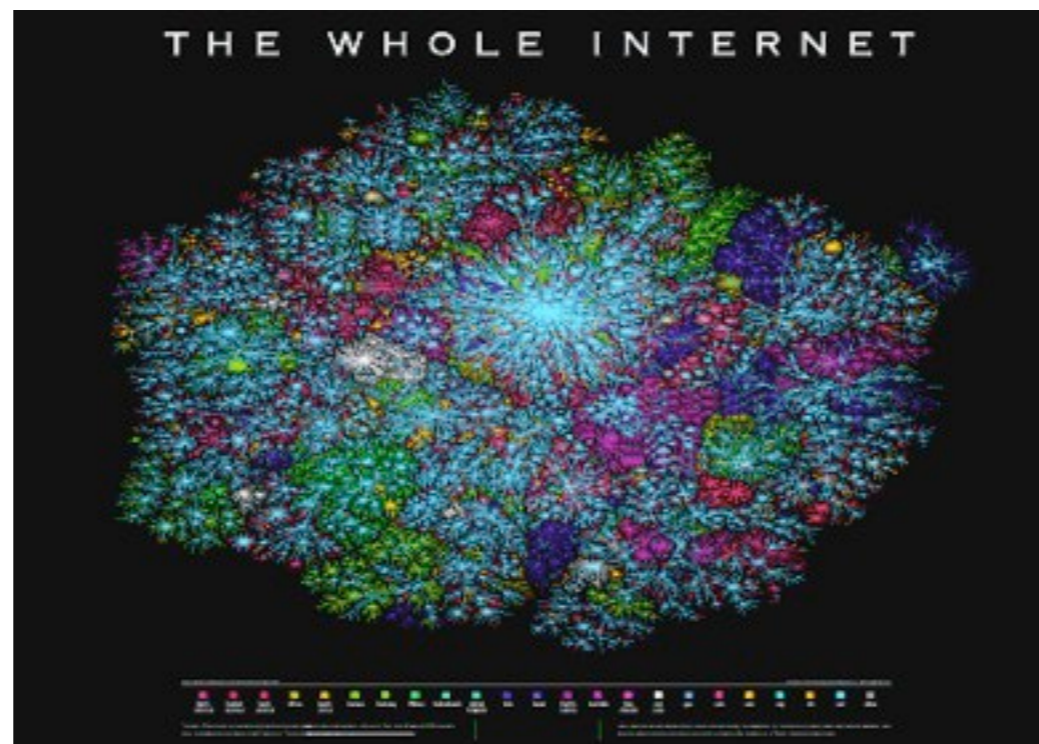


# Katz prestige (Page Rank)

---

*you may know people, but who says you get all their attention?*

$$P_i^K(g) = \sum_{j \neq i} g_{ij} \frac{P_j^K(g)}{d_j(g)}$$



# distance-based centrality

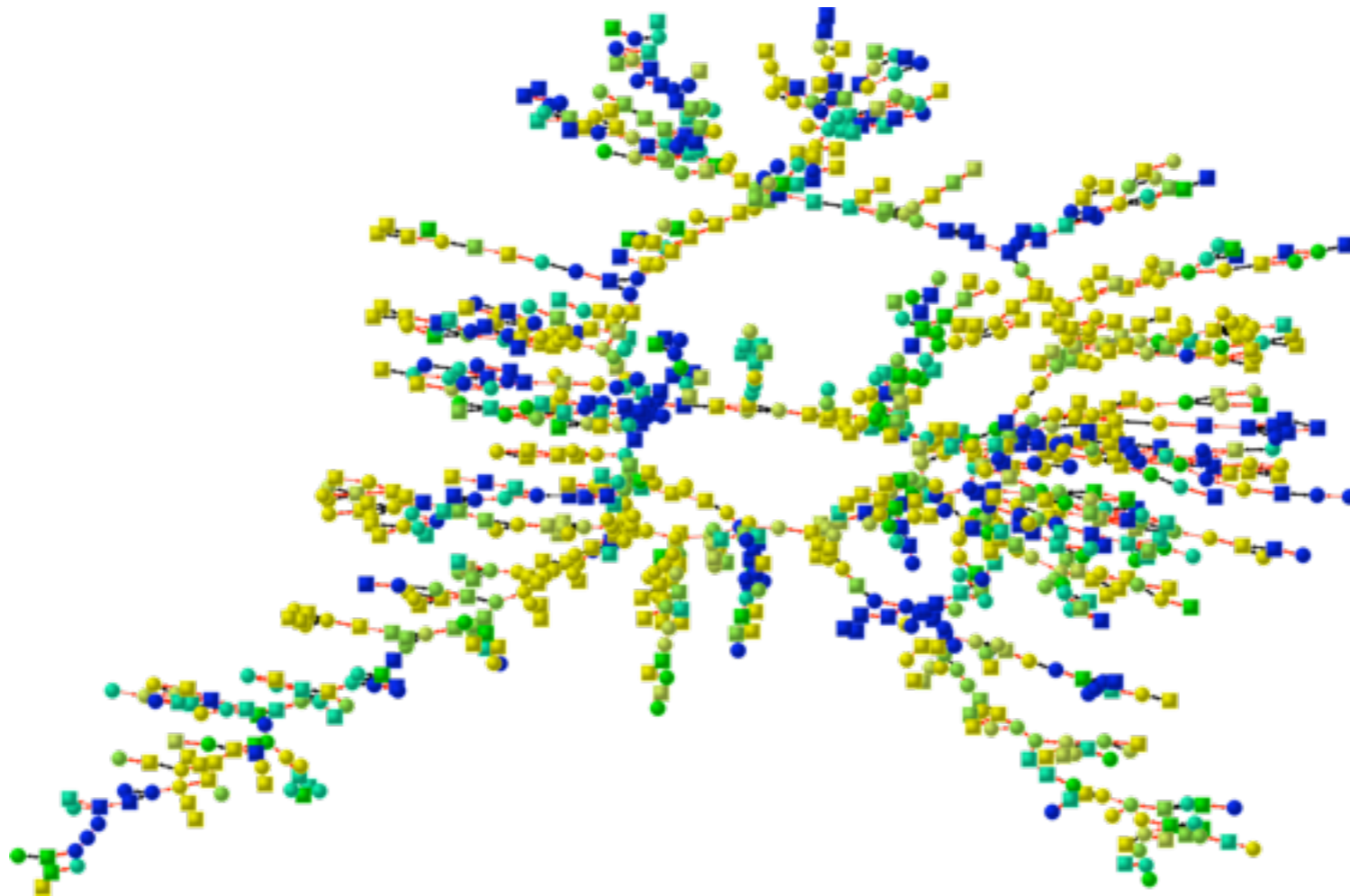
---

- closeness centrality
- betweenness

# closeness centrality

---

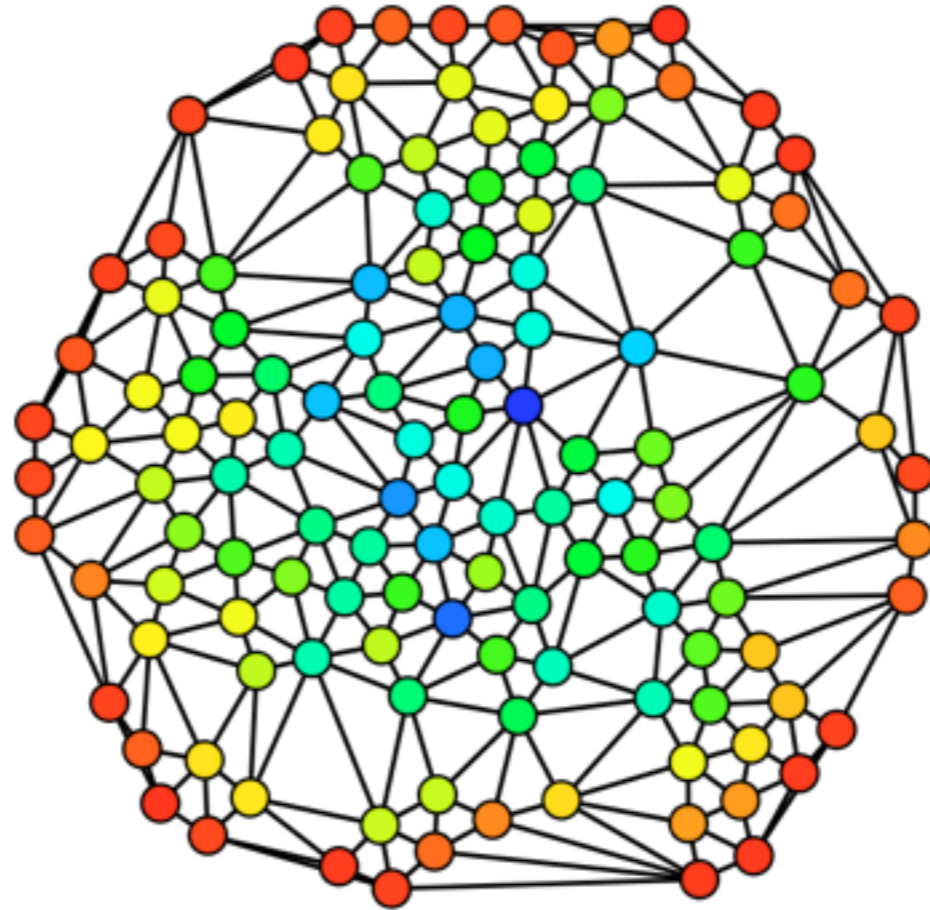
$$C_C(u) = \frac{1}{\sum_{v \neq u} \delta(u, v)}$$



betweenness

---

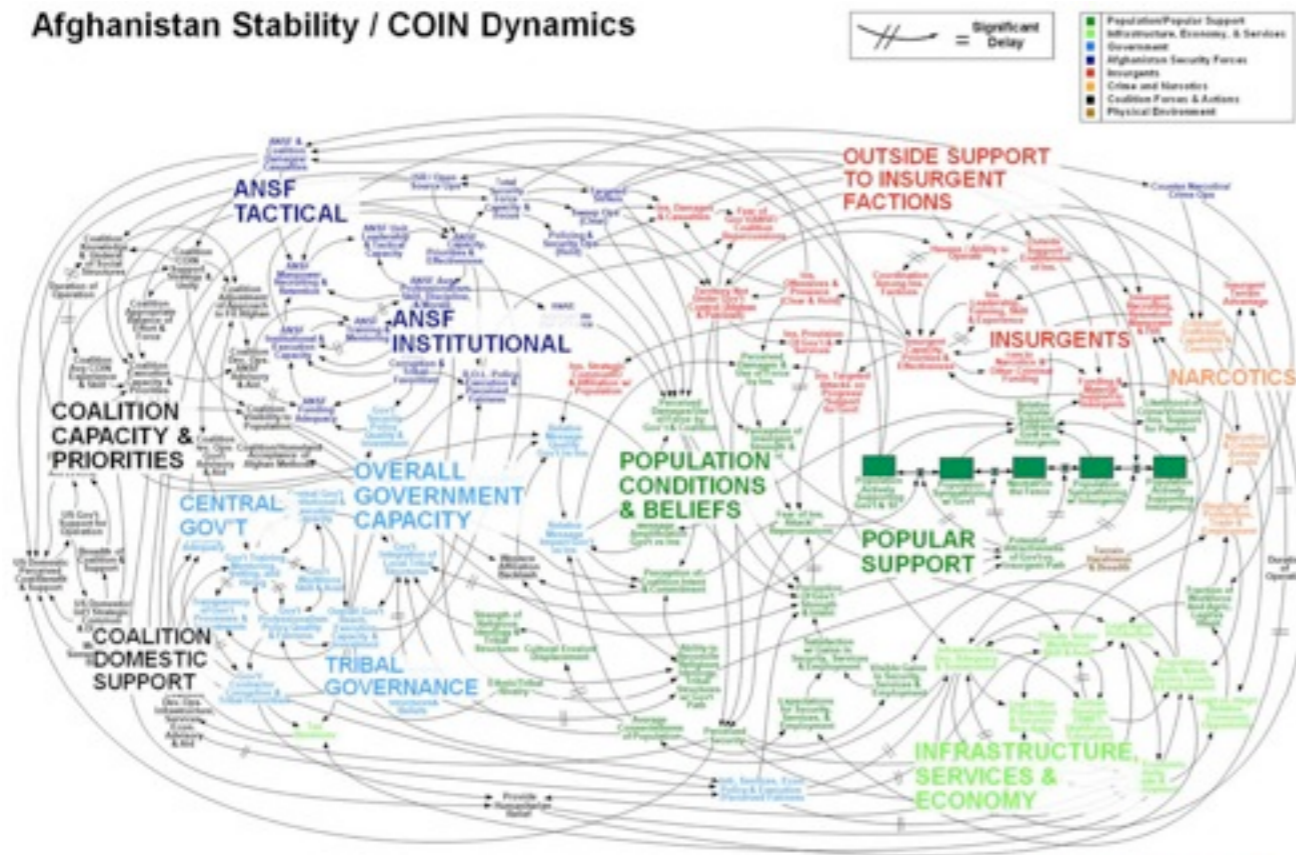
$$C_B(v) = \sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}}$$



# betweenness

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

## Afghanistan Stability / COIN Dynamics

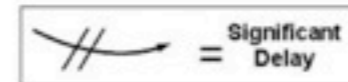


WORKING DRAFT - V3

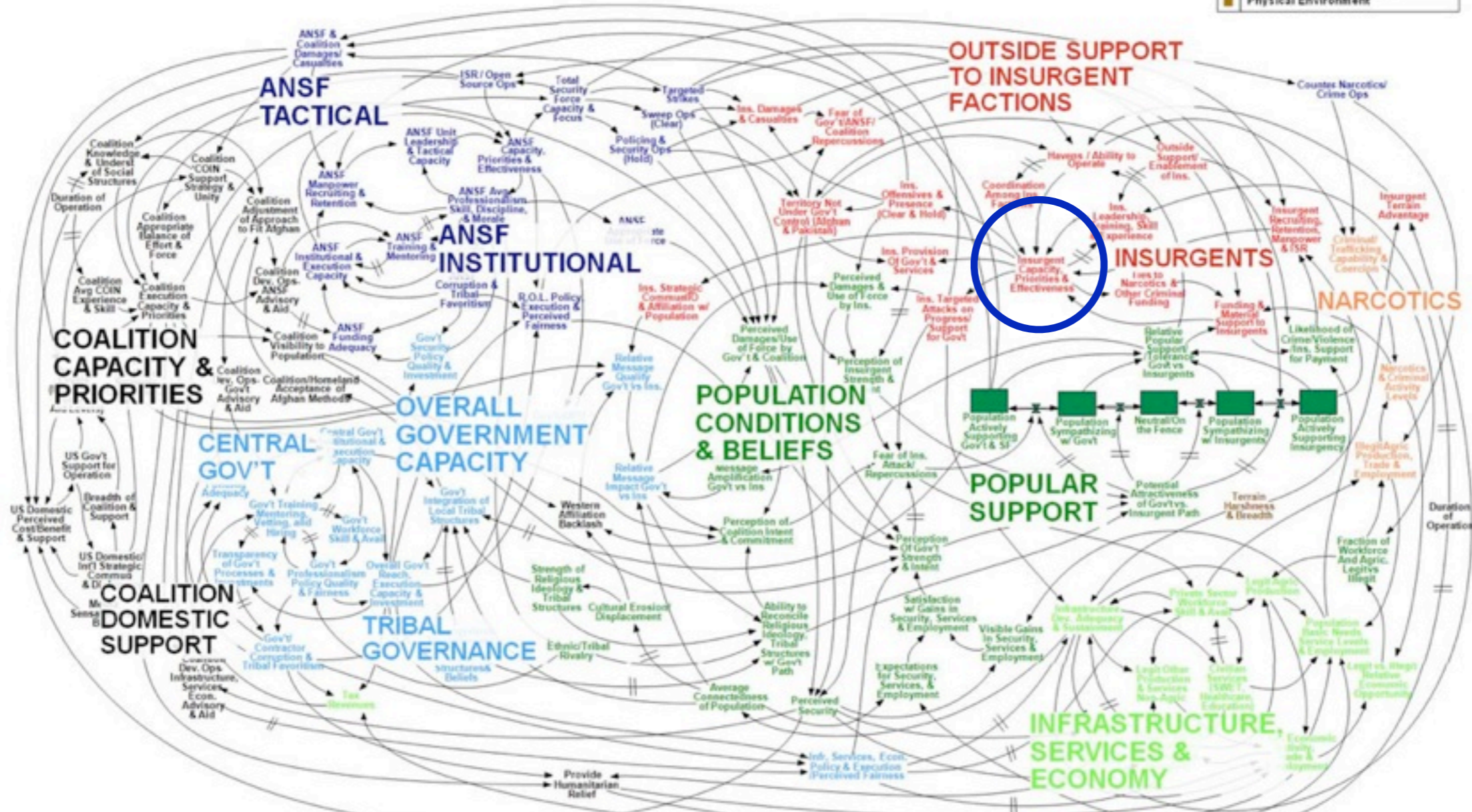
*when doesn't betweenness make sense?*

# most important node?

## Afghanistan Stability / COIN Dynamics



- Population/Popular Support
- Infrastructure, Economy, & Services
- Government
- Afghanistan Security Forces
- Insurgents
- Crime and Narcotics
- Coalition Forces & Actions
- Physical Environment

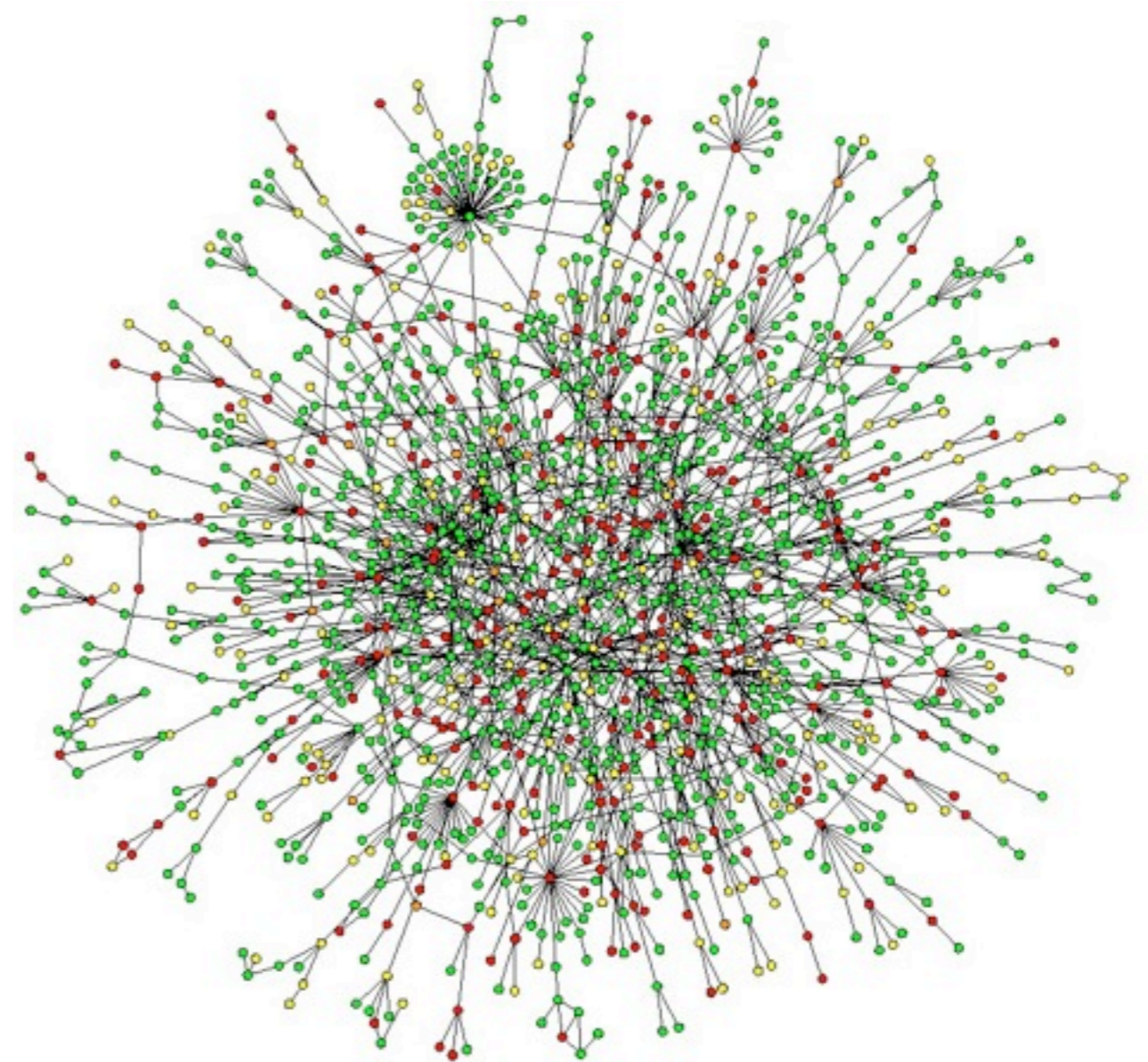
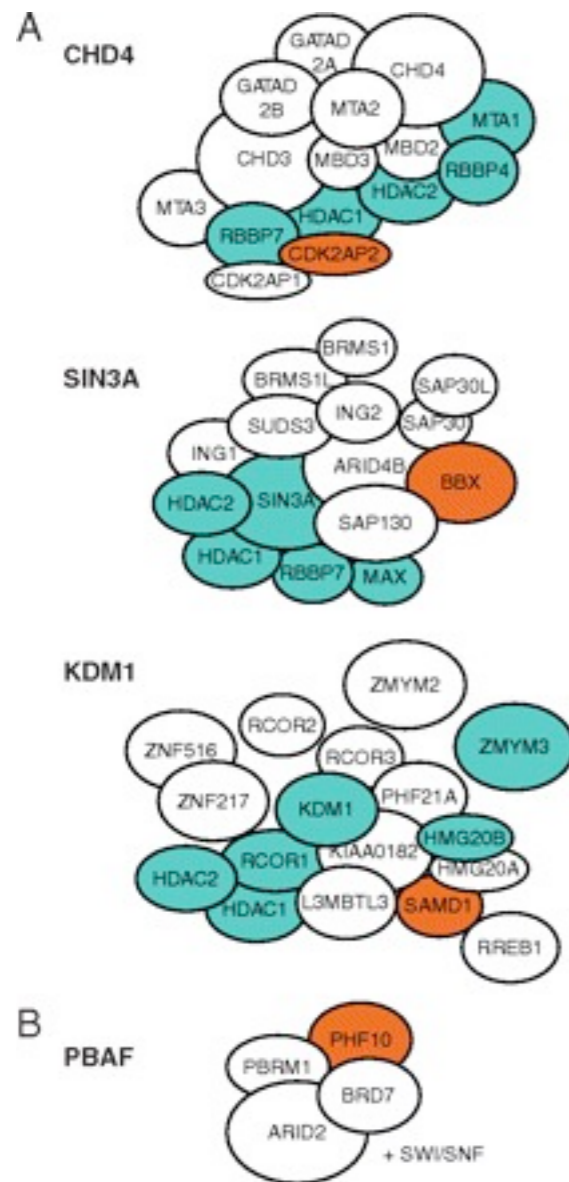


WORKING DRAFT - V3



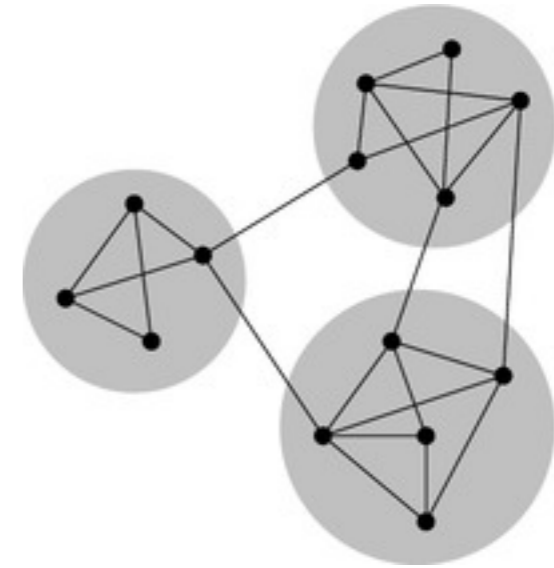
# what proteins form complexes?

---



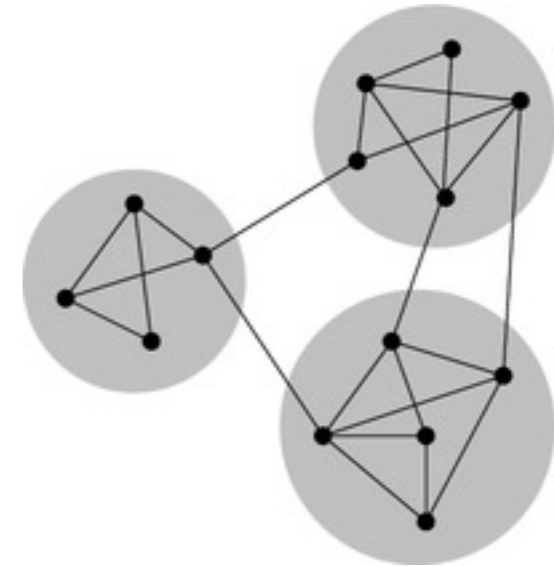
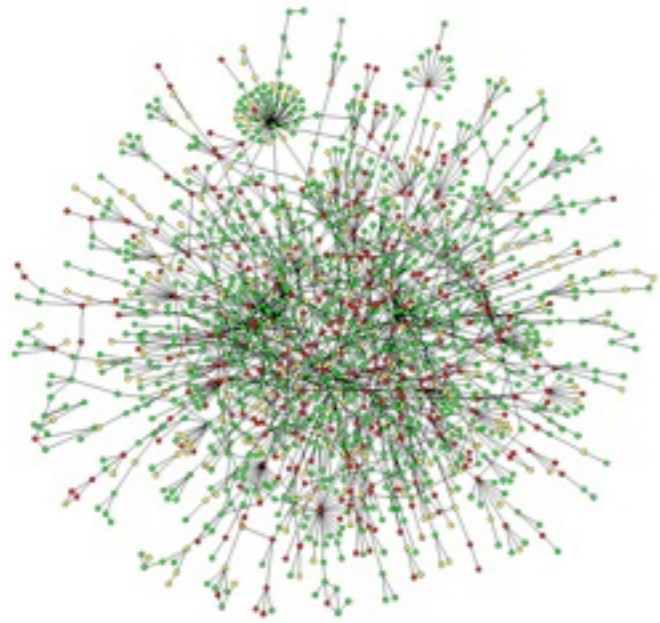
# community (module) structure

---



# community (module) structure

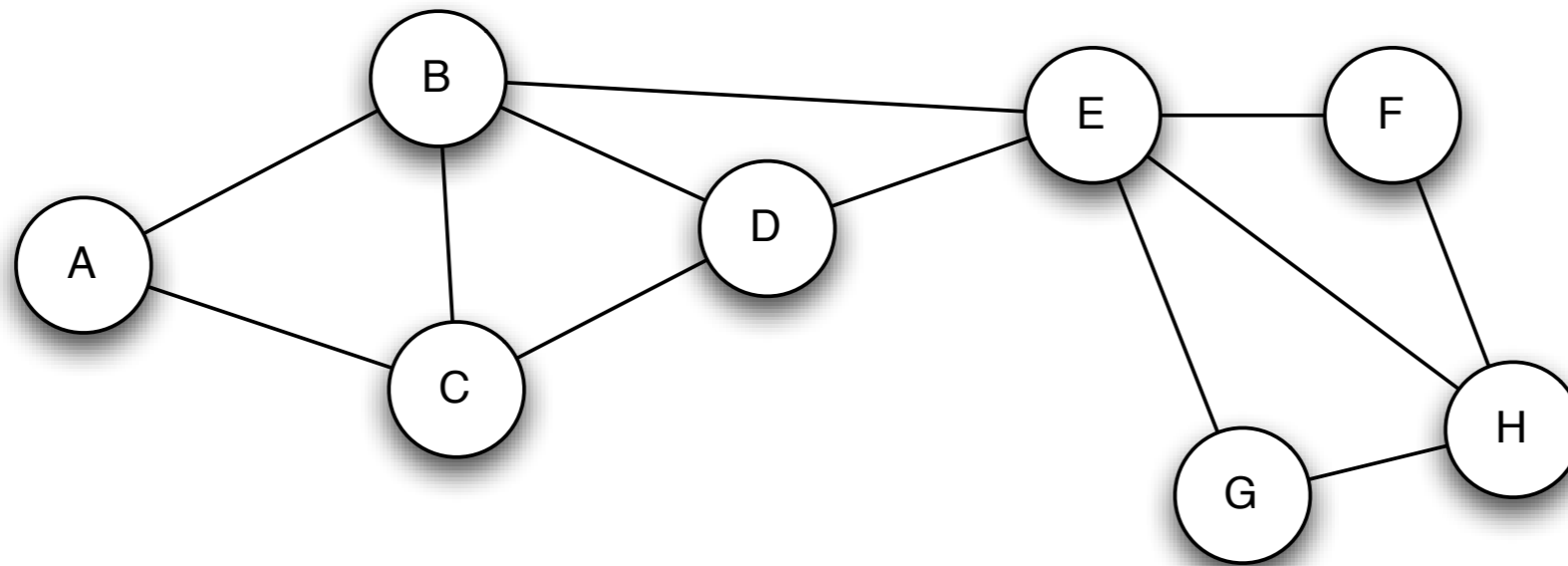
---



*Intuition: modules have more edges within them than between them*

# minimum-cut network partitioning

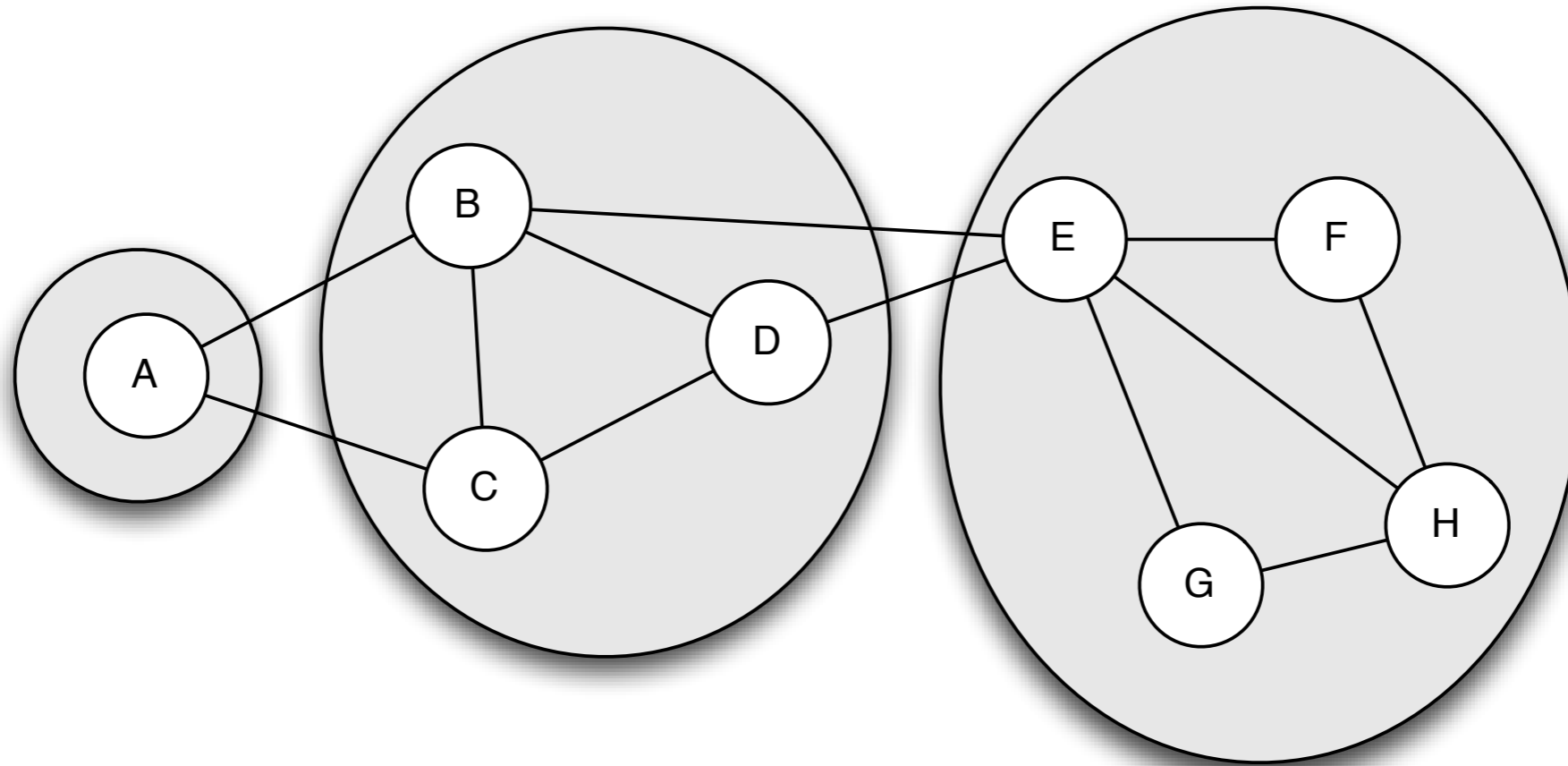
---



Group nodes to minimize the number of edges  
between groups

# minimum-cut network partitioning

---

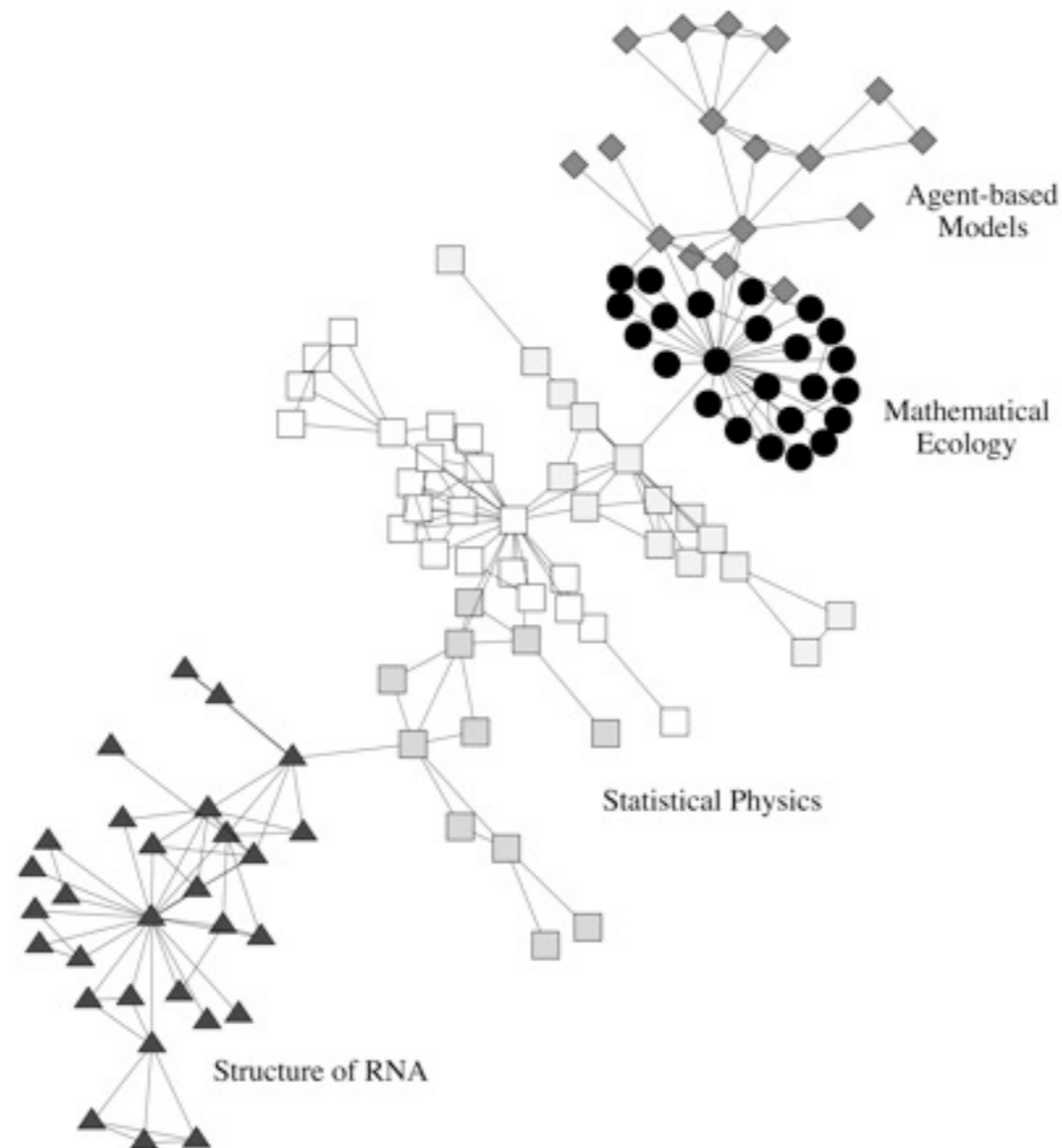


Group nodes to minimize the number of edges  
between groups

# Girvan-Newman algorithm

---

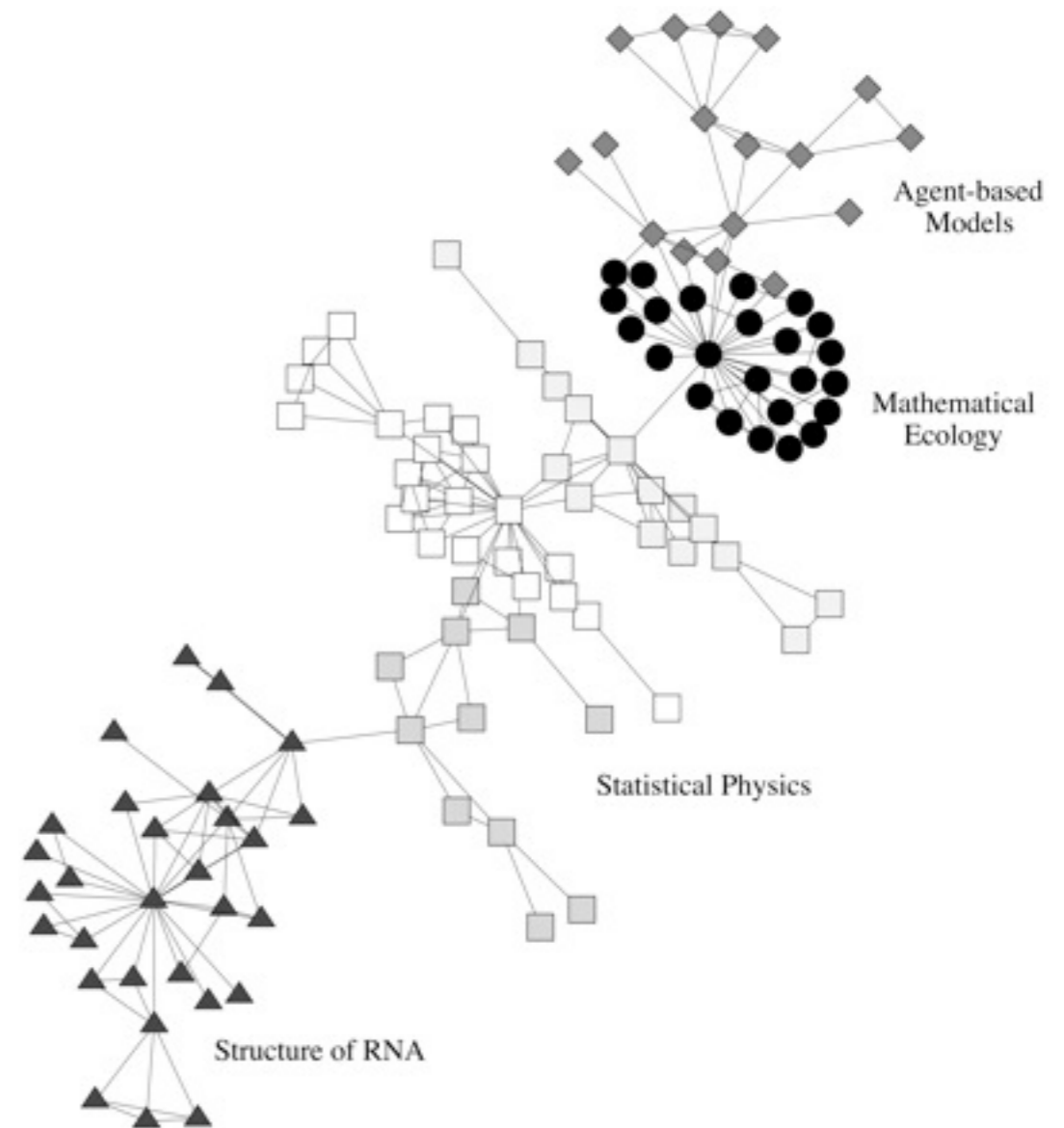
idea: edges between communities lie along many shortest paths



# Girvan-Newman algorithm

---

1. Calculate the betweenness for all edges in the network
2. Remove the edge with the highest betweenness
3. Recalculate betweenness for all edges
4. Repeat from step 2 until no edges remain



# we can quantify a variety of complex features

---

- *centrality: relative important of parts of a system*
- *community structure: communities and modules present within the system*
- formation processes: how edges and node join/leave the network
- motifs: how often very small, meaningful structures occur in the network
- dynamics: how content/information moves through a network