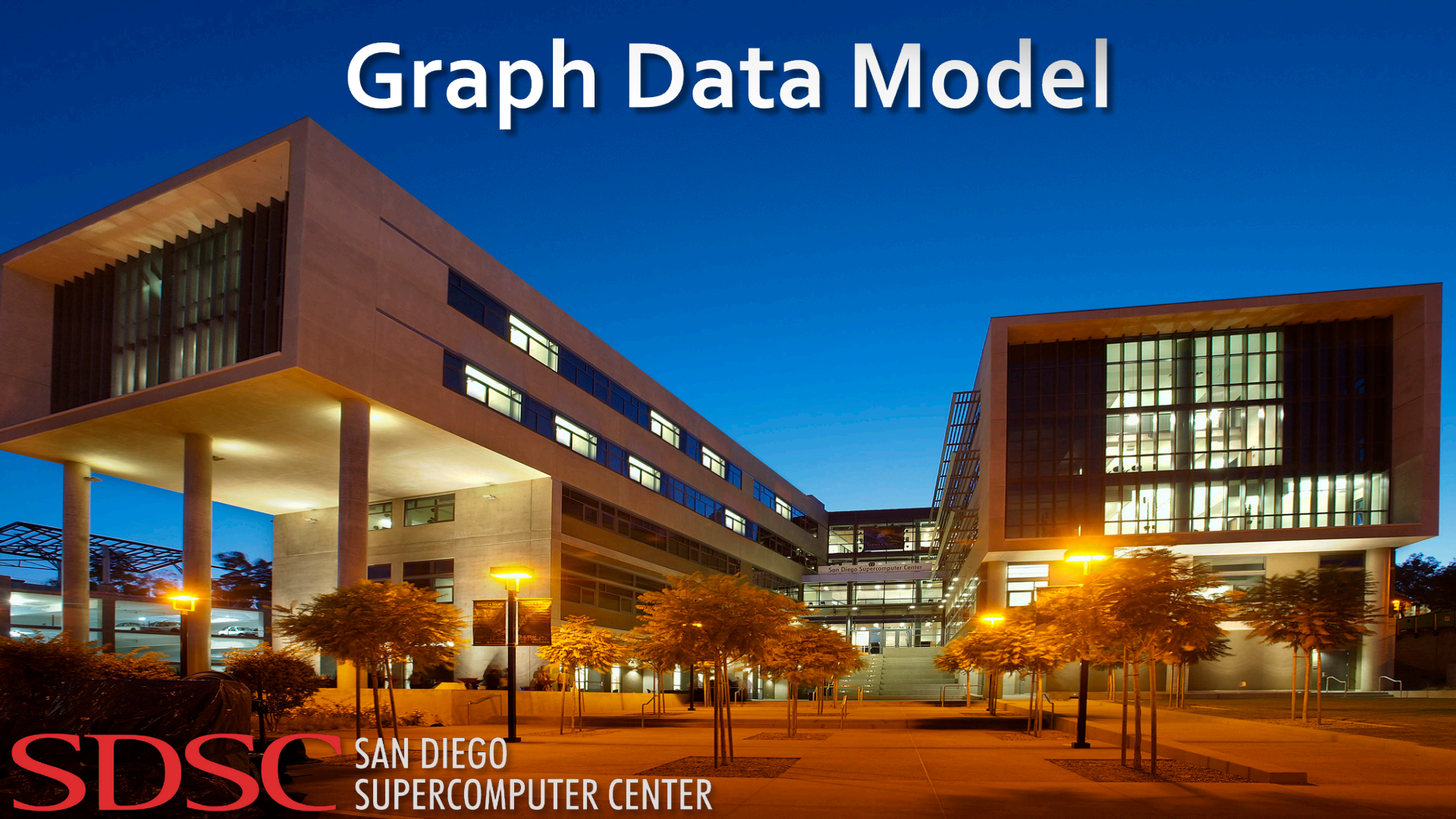
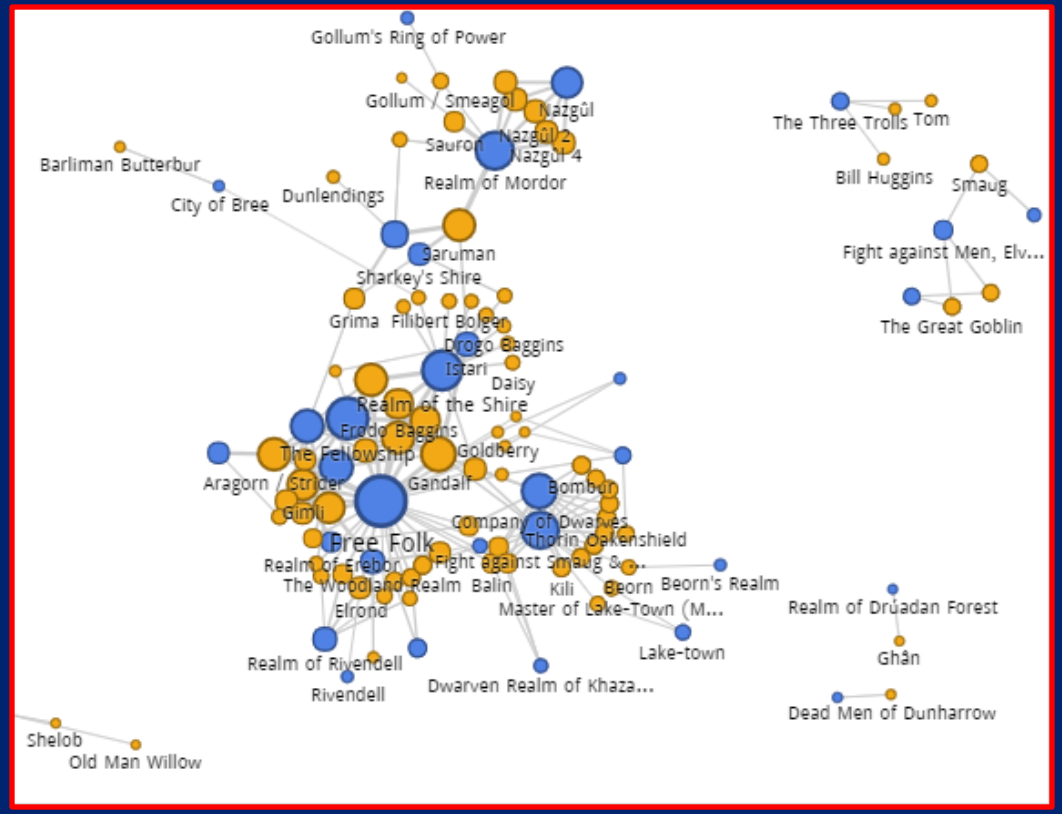
Graph Data Model



SDSC SAN DIEGO SUPERCOMPUTER CENTER

Social Network

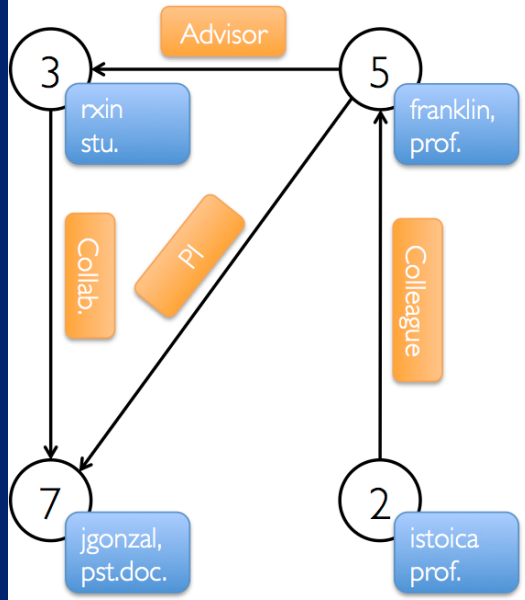


After this video you will be able to

- Identify graph data in practical problems
- Describe path operations, neighborhood operations, and connectivity operations in graph data

Data + Connectivity

Property Graph



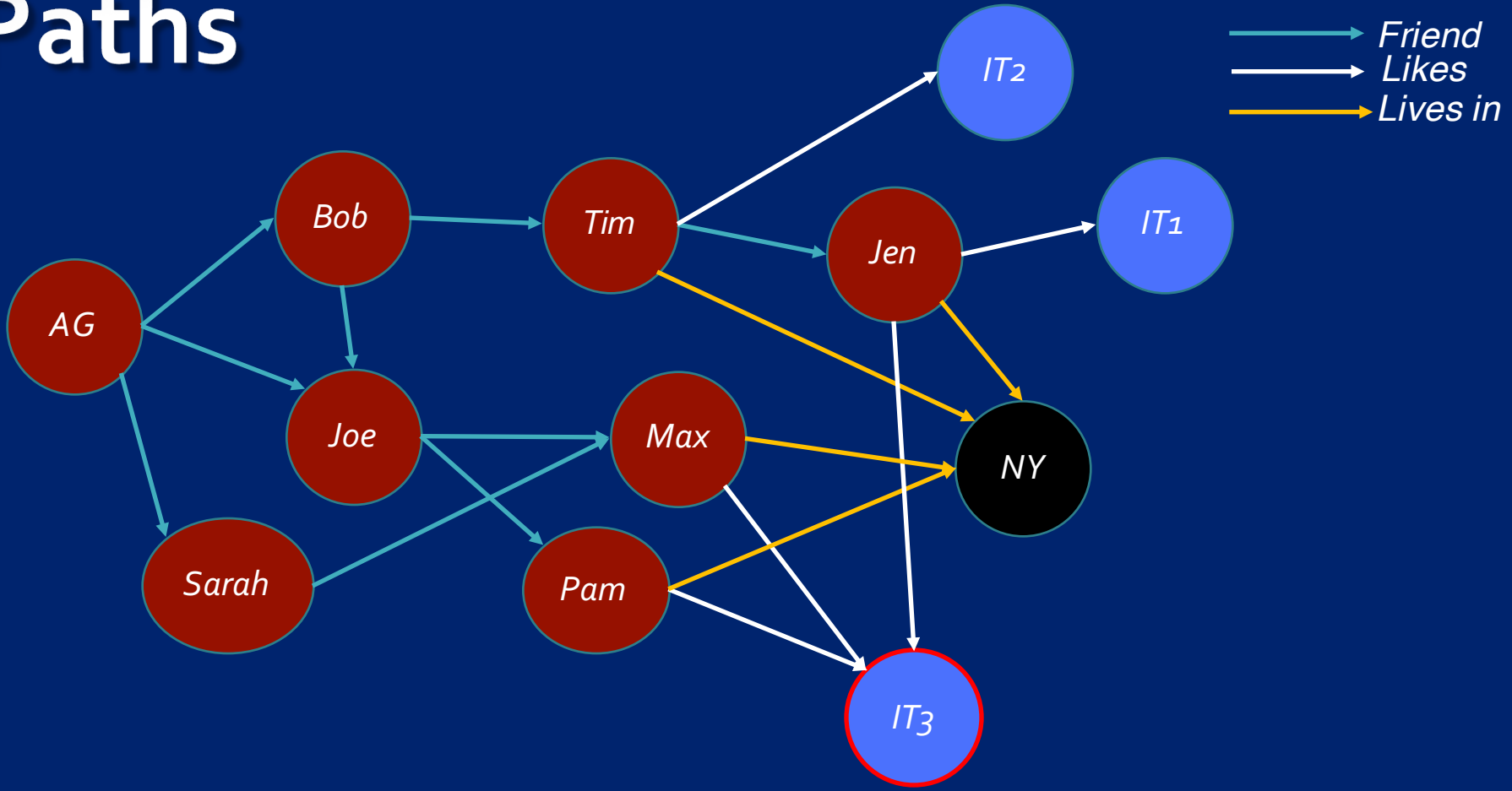
Vertex Table

Id	Property (V)
3	(rxin, student)
7	(jgonzal, postdoc)
5	(franklin, professor)
2	(istoica, professor)

Edge Table

SrcId	DstId	Property (E)
3	7	Collaborator
5	3	Advisor
2	5	Colleague
5	7	PI

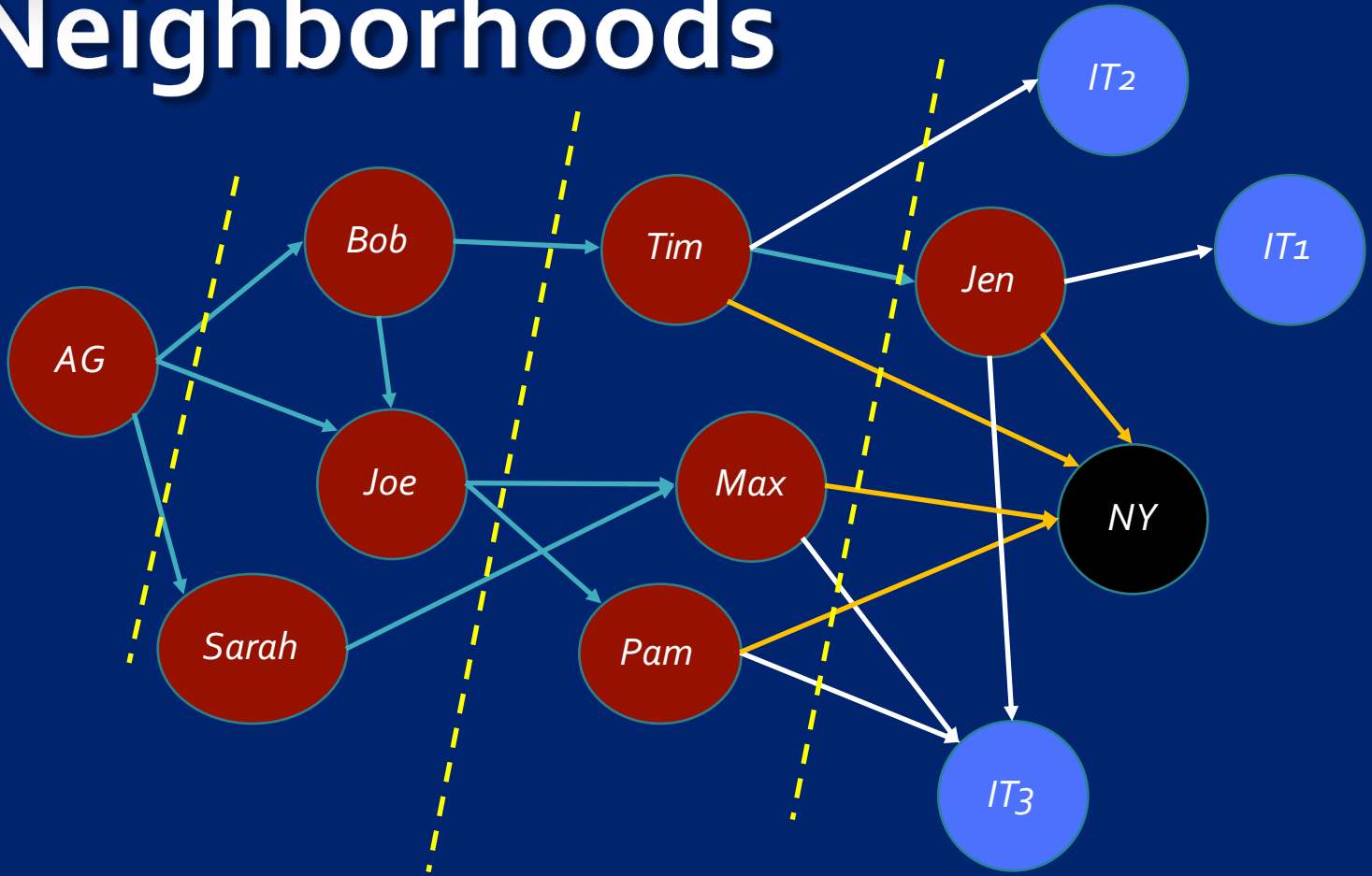
Paths


Friend
Likes
Lives in
IT2
Bob
Tim
Jen
IT1
AG
Joe
Max
NY
Sarah
Pam
IT3

"Optimal Path" Operations

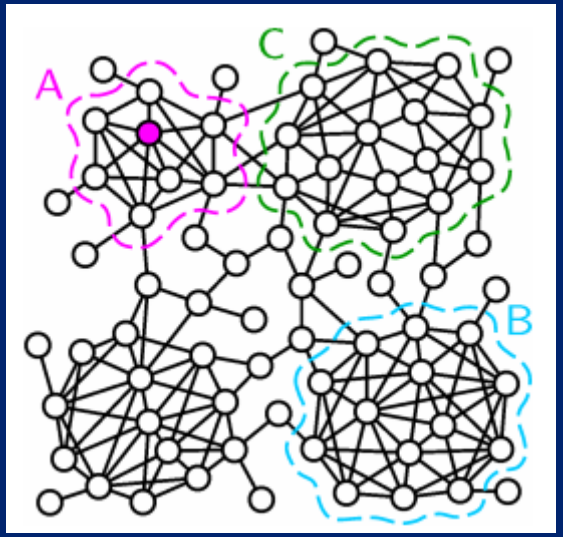
- Find the shortest path between two nodes
- Find an optimal round-trip path that must include some specific nodes
- Find "best compromise" paths between two nodes
  - Pareto-optimality

Neighborhoods



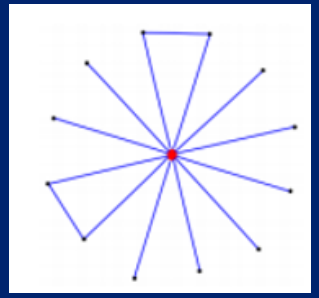
Communities

- **A subgraph of a graph that has many more edges *within* the subgraph compared to edges to nodes outside the subgraph**
- **Operations**
  - Dense subgraph finding
  - Optimization of clusteredness

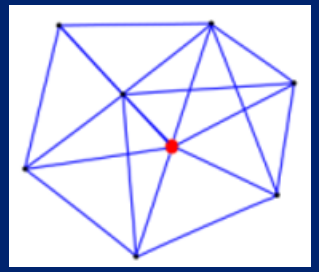


Anomalous Neighborhoods

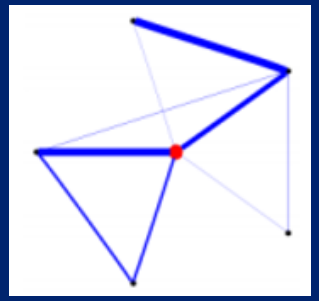
(Akoglu et al 2010)



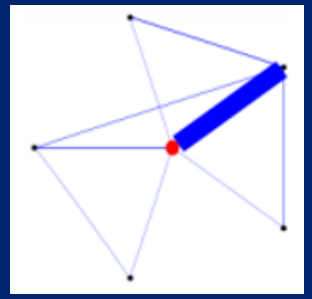
- Near star



- Near clique



- Heavy vicinity



- Predominant Edge

Connectivity Operations

- **Connectedness**
  - Every node is reachable from each node in the undirected version of the graph