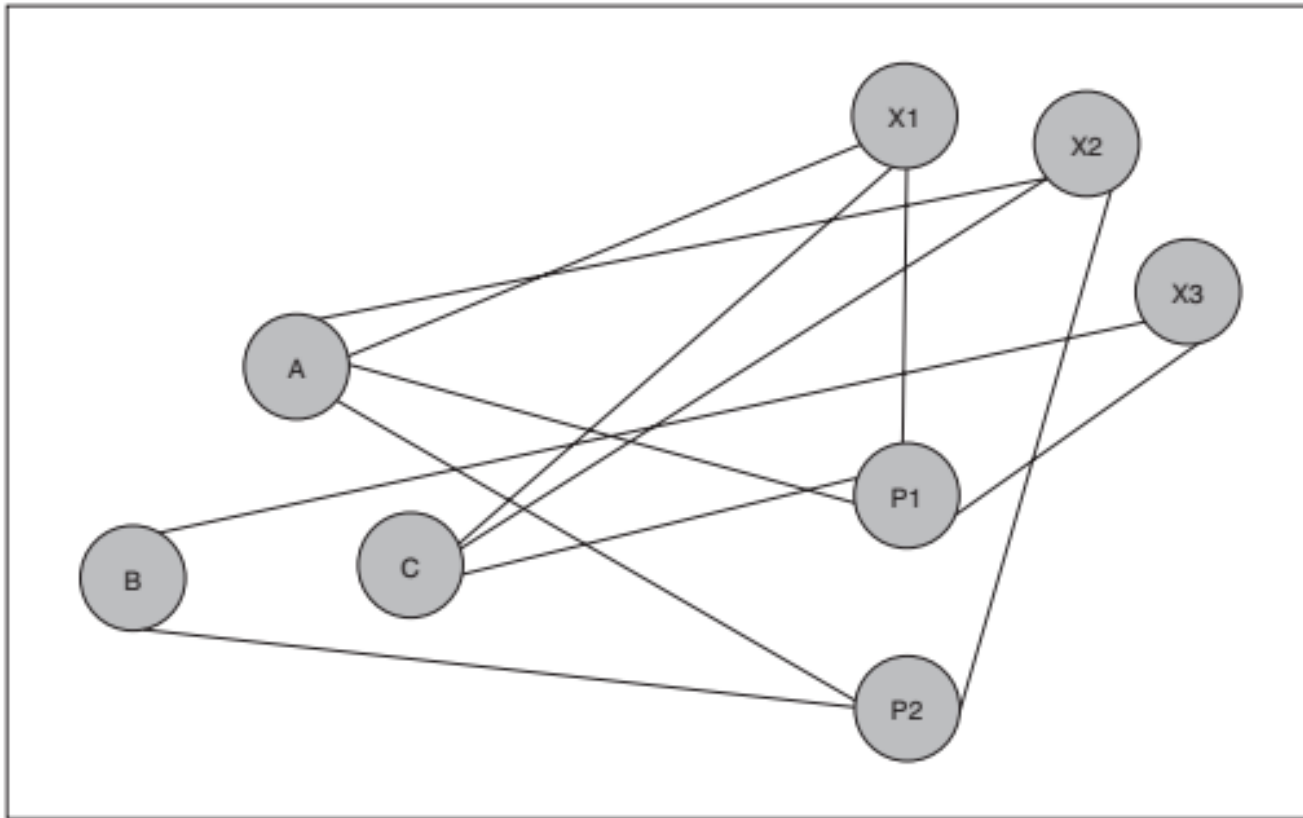


Simrank and Counting triangles using Mapreduce

Simrank

- SimRank is a measure of node similarity between nodes of the same type.
- For example, in a movie database network consisting of nodes representing movie, actors, etc., SimRank can measure similarity between sets of actors, sets of movies and the like.
- It uses an algorithm simulating random surfers similar to the computation of PageRank.



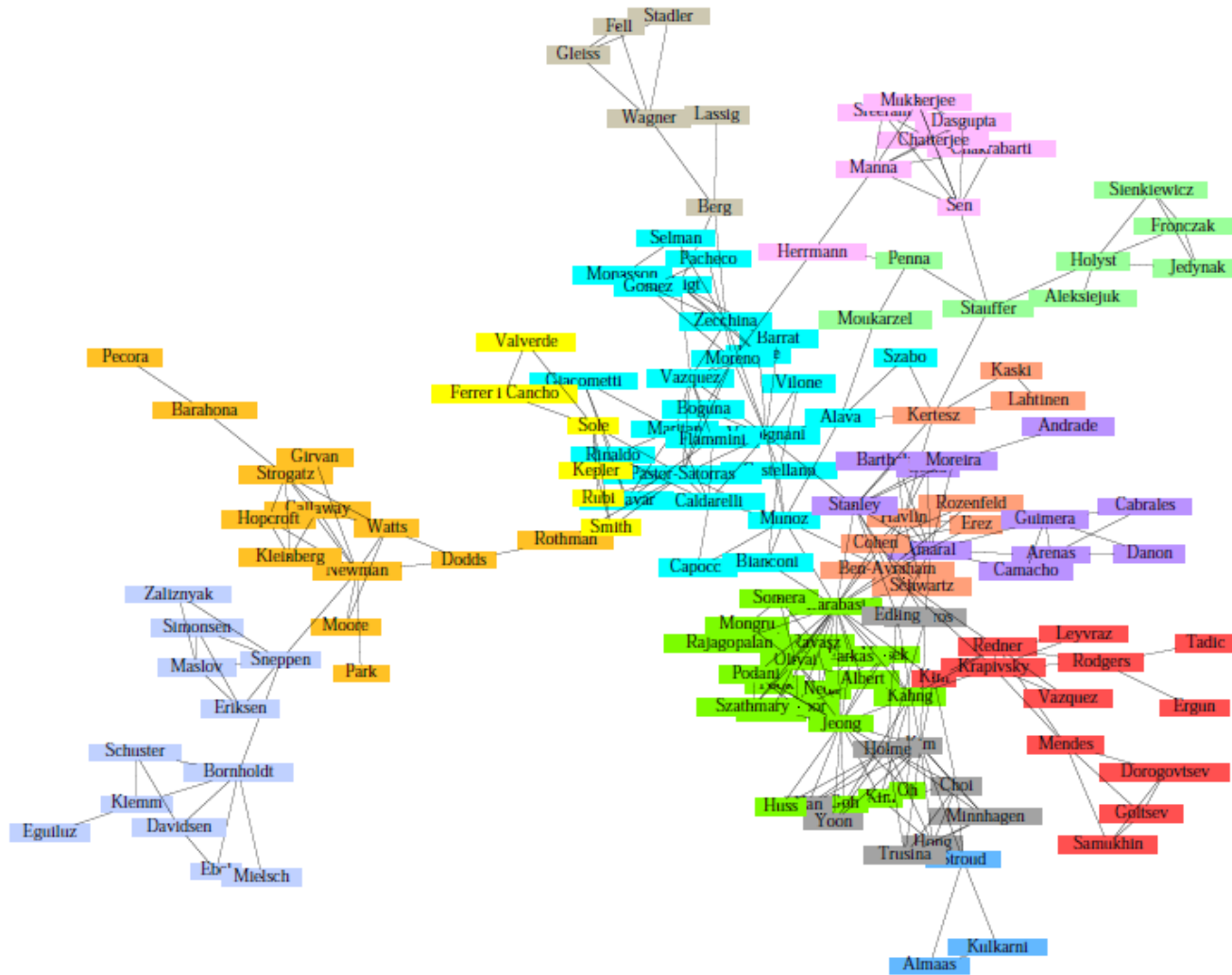
“ two objects are considered
to be similar if they are referenced by similar
objects”

Counting triangles in a social graph using map reduce

- The nodes of a graph are numbered $1, 2, \dots, n$
- A relation E to represent edges
- The Map tasks divide the relation E into as many parts
- one Map task is given the tuple $E(u, v)$ to send to certain Reduce tasks.

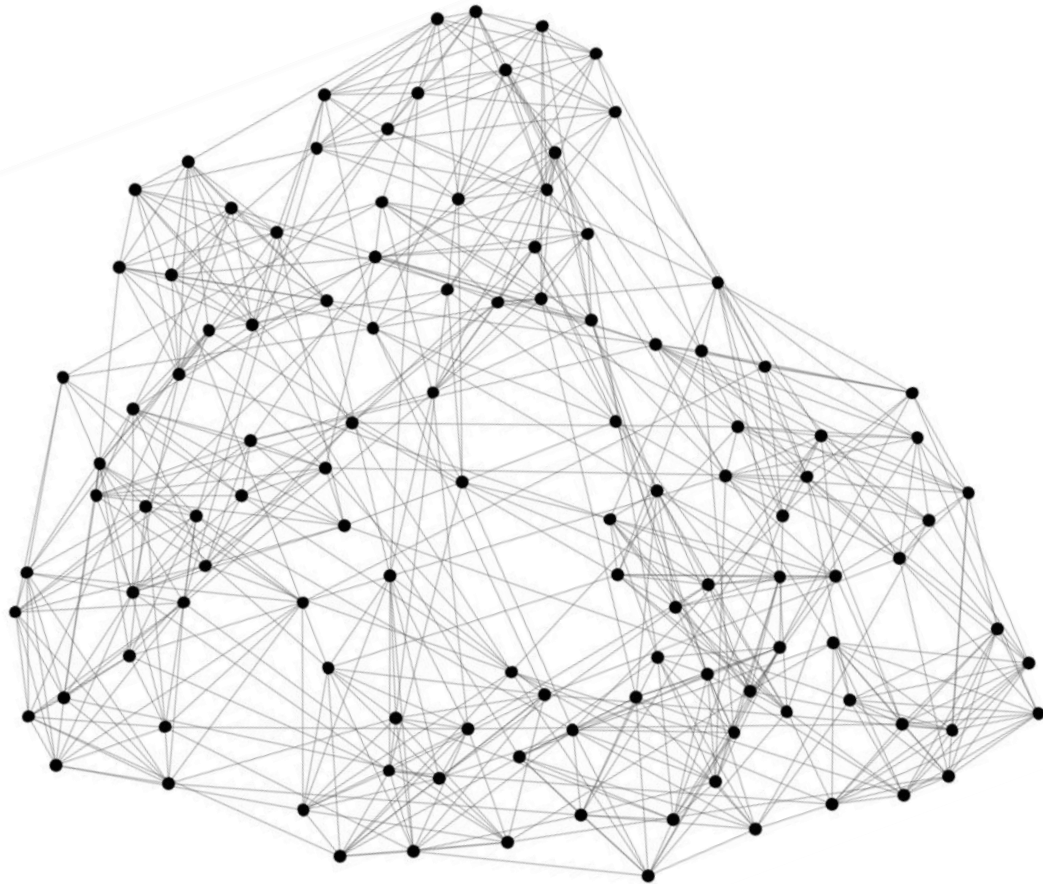
Thank You

Girvan-Newman: Results



Communities in physics collaborations

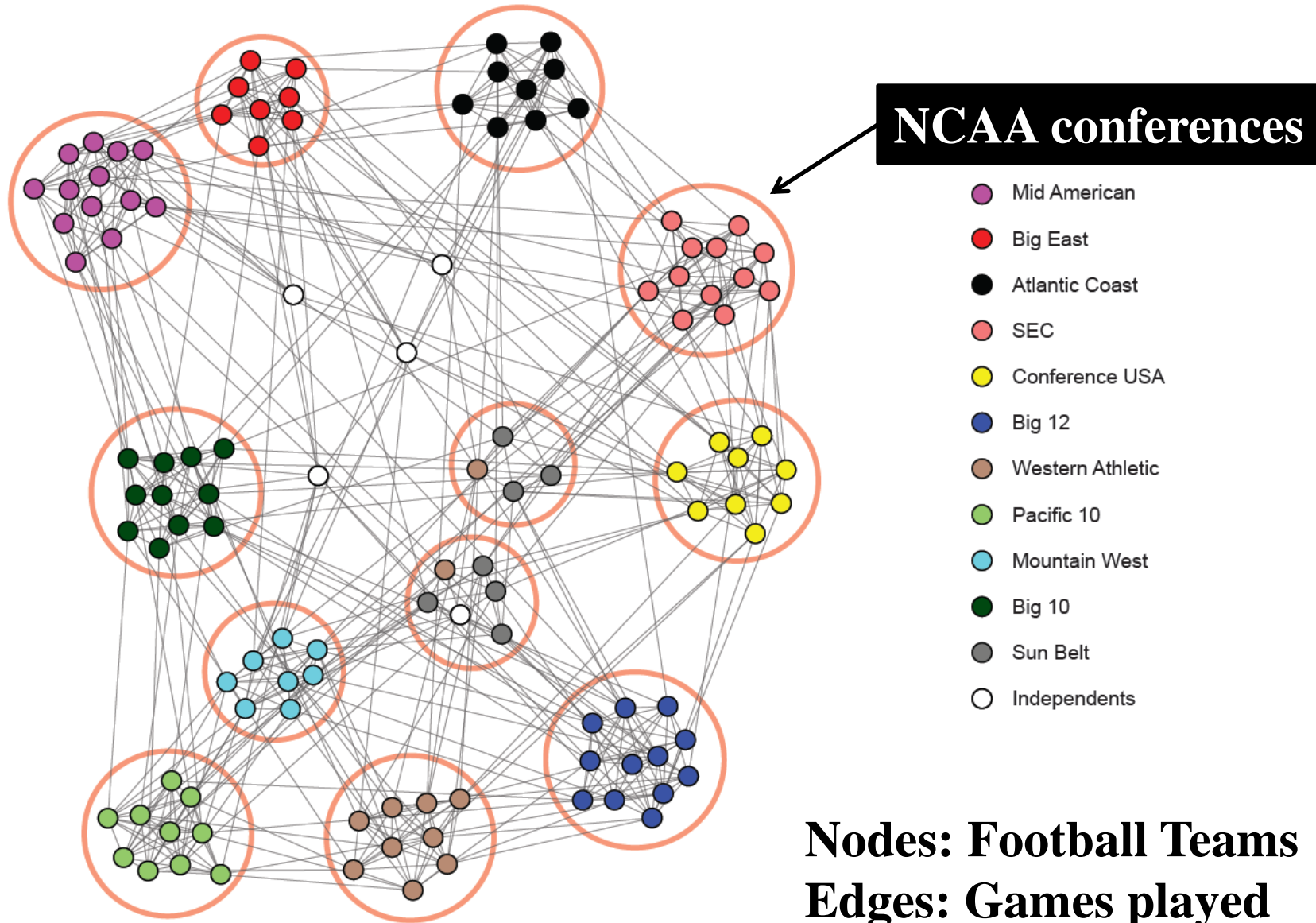
Identifying Communities



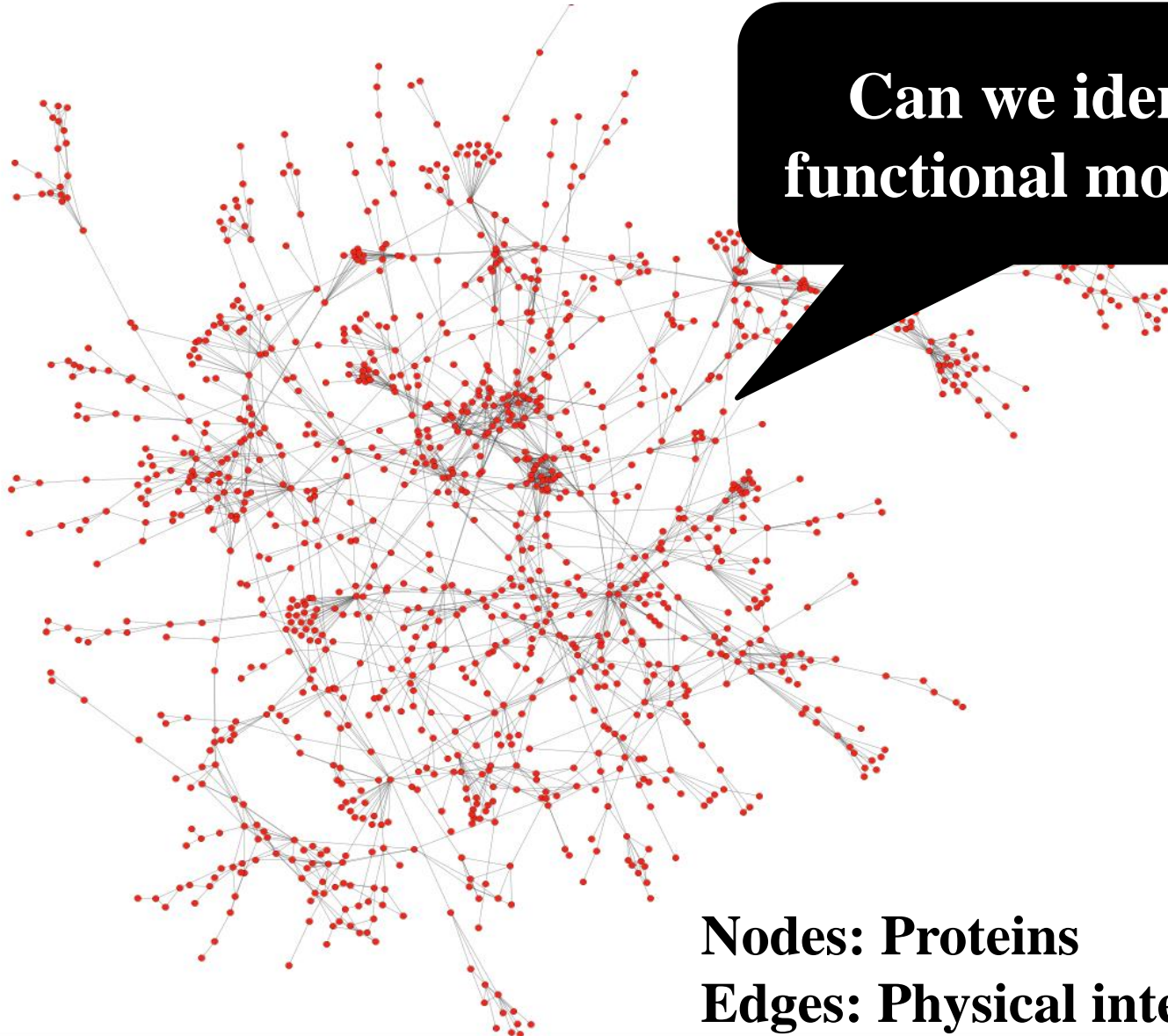
**node groups?
(communities,
modules,**

**Nodes: Football
Teams**

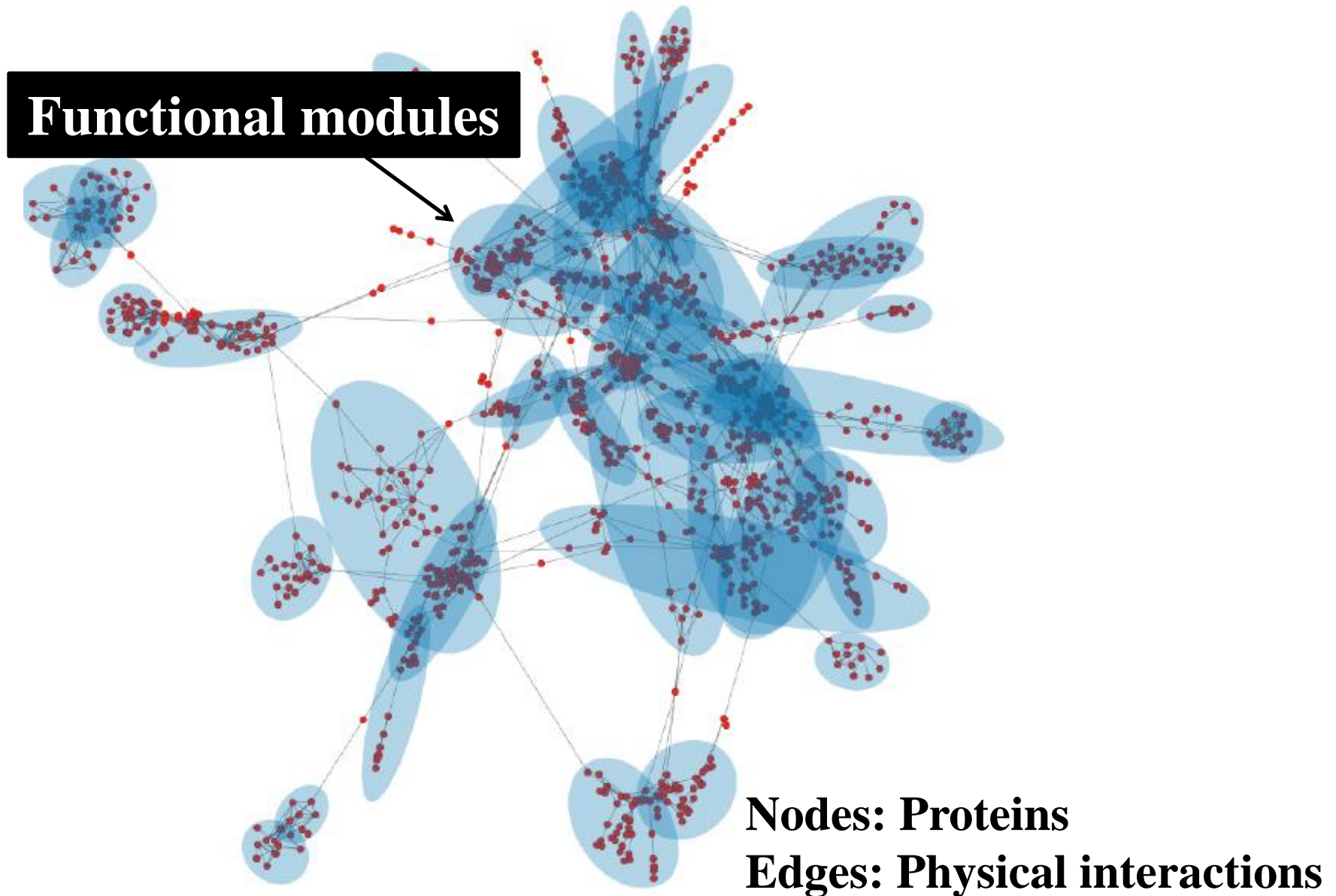
NCAA Football Network



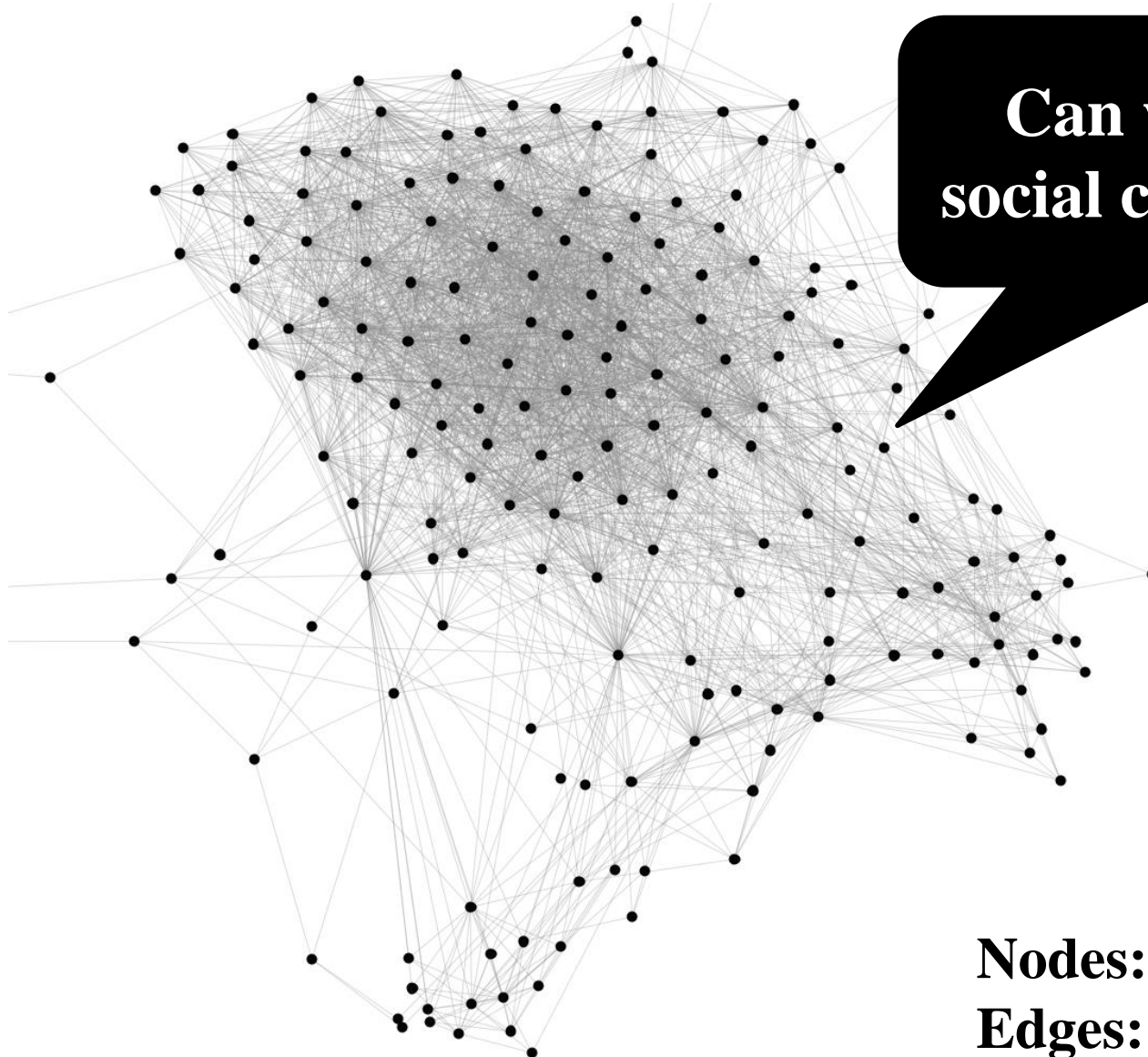
Protein-Protein Interactions



Protein-Protein Interactions



Facebook Network



**Can we identify
social communities?**

Nodes: Facebook Users
Edges: Friendships