Social Network Analysis

Chen Avin
School of Electrical and Computer Engineering

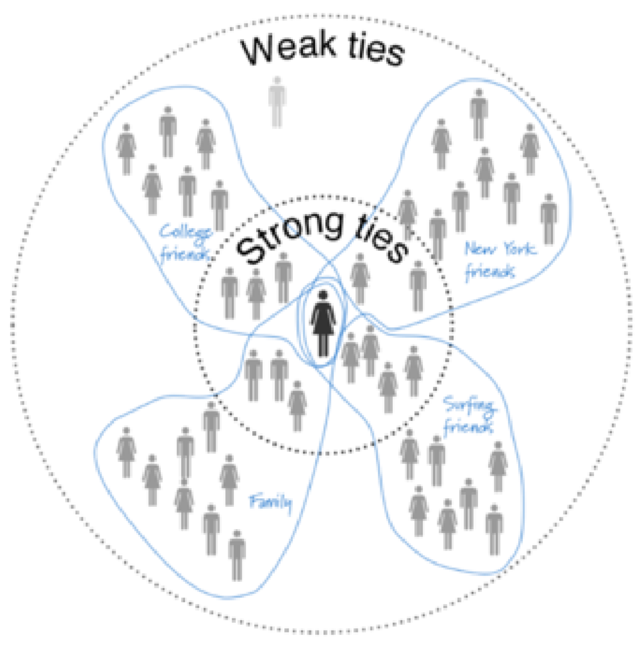
Unit 3
Strong and Weak Ties

(Based on Networks, Crowds, and Markets: Reasoning about a Highly Connected World. By David Easley and Jon Kleinberg. Chapter 3)

How people discover their new job?

Mark Granovetter research in late 1960s

- Many learned information via personal contacts
- But, surprisingly (?) via acquaintances rather than close friend
- Granovetter answer links two different perspectives on distant friendships
  - Structural - how friendship span different portions on the full network
  - Interpersonal - friendship between people being either strong or weak

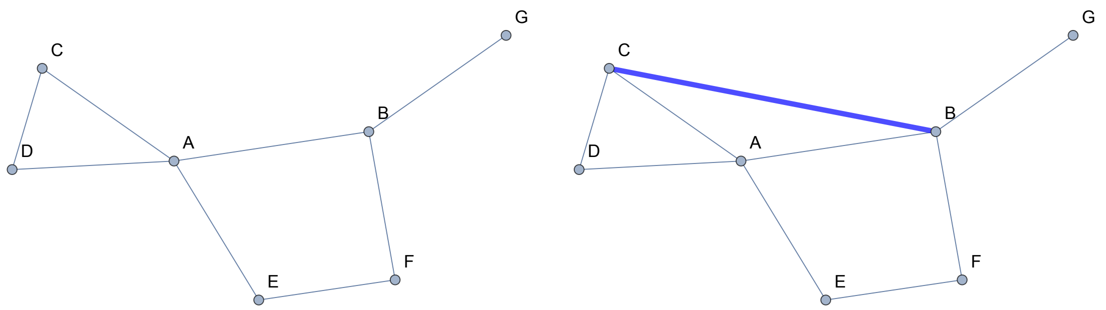


Triadic Closure

- How do nodes arrive and depart and edges form and vanish?
- The Triadic Closure is a very basic principle

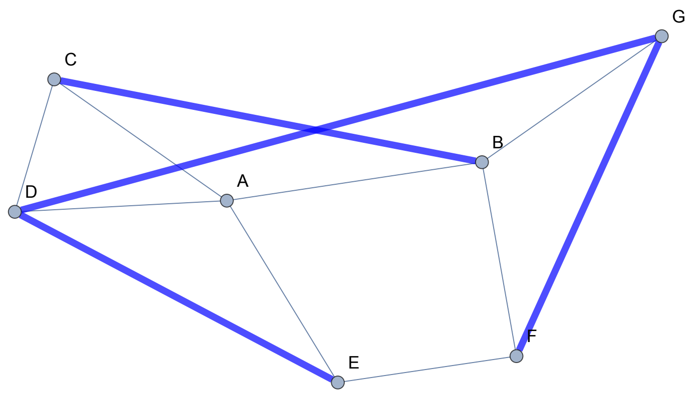
"If two people in a social network have a friend in common, then there is an increased likelihood that they will become friends themselves at some point in the future"

Out[5]=



In[8]:= `HighlightGraph[AfterBase, Style[new, Blue, Thickness[.01]]]`

Out[8]=



Reasons for Triadic Closure

- Opportunity (for B to meet C)
- Trust (between B and C, based on A)
- Incentive (for A to bring B and C together)

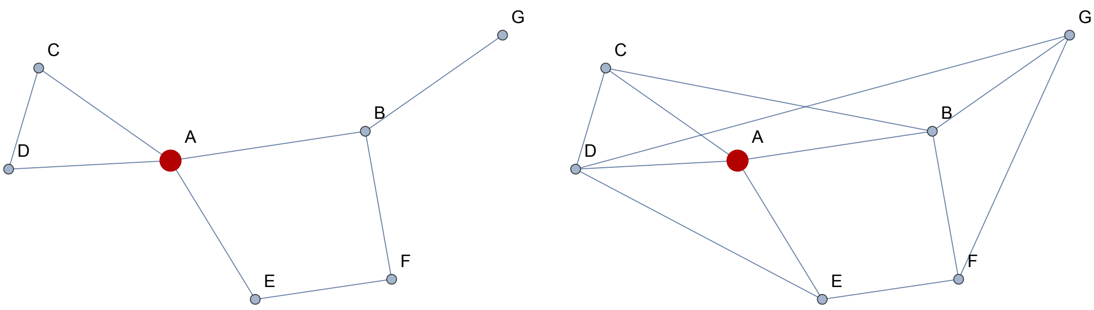
Measuring Triadic Closure - Clustering Coefficient

(Local) clustering coefficient

- (Local) clustering coefficient of the vertex v is the fraction of pairs of neighbors of v that are connected over all pairs of neighbors of v

In[9]:=
```
GraphicsRow[{HighlightGraph[Base, "A", VertexSize → Medium],
  HighlightGraph[AfterBase, "A", VertexSize → Medium]}]
```

Out[9]=



- How many possible pairs of neighbors ?

In[10]:=
```
Binomial[Length[AdjacencyList[Base, "A"]], 2]
```

Out[10]=

6

In[11]:=
```
LocalClusteringCoefficient[Base, "A"]
```

Out[11]=

$\frac{1}{6}$

In[12]:=
```
LocalClusteringCoefficient[AfterBase, "A"]
```

Out[12]=

$\frac{1}{2}$

In[13]:=
```
LocalClusteringCoefficient[AfterBase]
```

Out[13]=

$\left\{\frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right\}$

In[14]:=
```
Mean[LocalClusteringCoefficient[AfterBase]]
```

Out[14]=

$\frac{17}{42}$

In[15]:=
```
MeanClusteringCoefficient[AfterBase]
```

Out[15]=

$\frac{17}{42}$

(global) clustering coefficient

- (global) clustering coefficient of G is the fraction of paths of length two in G that are closed over all paths of length two in G

In[16]:=
```
GlobalClusteringCoefficient[AfterBase]
```
Out[16]=

$$\frac{2}{5}$$

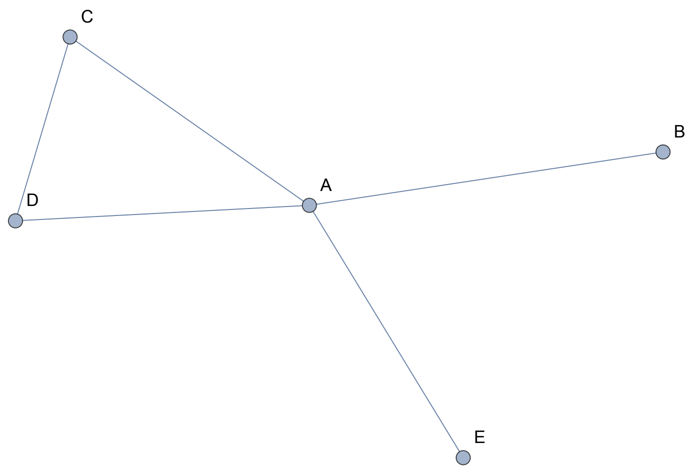
In[17]:=
```
Length[FindCycle[AfterBase, 3, All]] * 3
```
Out[17]=
```
12
```

In[19]:=
```
Binomial[VertexDegree[AfterBase], 2]
```
Out[19]=
```
{6, 6, 3, 6, 3, 3, 3}
```

In[18]:=
```
Total[Binomial[VertexDegree[AfterBase], 2]]
```
Out[18]=
```
30
```

In[20]:=
```
Total@Total@UpperTriangularize[MatrixPower[AdjacencyMatrix[AfterBase], 2], 1]
```
Out[20]=
```
30
```

In[21]:=
```
T = NeighborhoodGraph[Base, "A"]
```
Out[21]=



In[22]:=
```
GlobalClusteringCoefficient[T]
```
Out[22]=

$$\frac{3}{8}$$

In[23]:=
```
Length[FindCycle[T, 3, All]] * 3
```
Out[23]=
```
3
```

In[24]:=
```
VertexDegree[T]
```
Out[24]=
```
{4, 1, 2, 2, 1}
```

In[25]:=
```
Total[Binomial[VertexDegree[T], 2]]
```
Out[25]=
```
8
```

In[26]:=
```
U = UpperTriangularize[MatrixPower[AdjacencyMatrix[T], 2], 1];
```

In[27]:=
```
U // MatrixForm
```
Out[27]//MatrixForm=

$$\begin{pmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

In[28]:=
```
Total[Total[U]]
```
Out[28]=
```
8
```

The Strength of Weak Ties

bridge

- How triadic closure is related to Granovetter findings ?
- An edge *e* joining two nodes A and B in a graph is a bridge if deleting *e* would cause A and B to lie in two different components

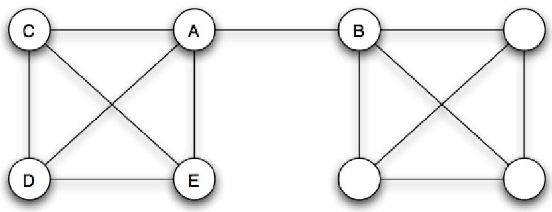


Figure 3.3: The *A*-*B* edge is a *bridge*, meaning that its removal would place *A* and *B* in distinct connected components. Bridges provide nodes with access to parts of the network that are unreachable by other means.

- But, almost no bridges in real networks

local bridge

- An edge *e* joining two nodes A and B in a graph is a local bridge if its endpoints A and B have no friends in common

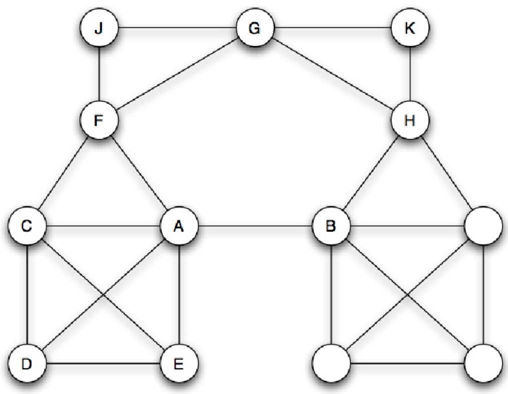
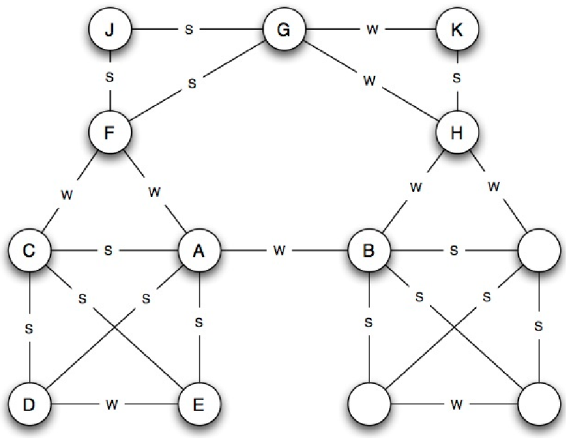


Figure 3.4: The *A*-*B* edge is a local bridge of span 4, since the removal of this edge would increase the distance between *A* and *B* to 4.

- Deleting a local bridge increasing the distance between A and B to **more** than 2.
- The span of a local bridge - the distance after deletion
- Relation to triadic closure - an edge is a local bridge if it is not a side in any triangle
- The role of a (local) bridge - provide endpoints with access to parts of the networks, source of information, they would otherwise be far away from
- A network concept that relates to the job-seeking observation!

The Strong Triadic Closure Property

- Interview subjects didn't say, "I learned about the job from a friend connected by a local bridge."
- Need to distinguish between different levels of strength in the links of a social network.
- Stronger links represent closer friendship and greater frequency of interaction.
- For simplicity two types: strong ties and weak ties (acquaintances).



- From triadic closure we can assume

If a node A has edges to nodes B and C, then the B–C edge is especially likely to form if A's edges to B and C are both strong ties.

- Granovetter suggested a more formal version

We say that a node A violates the Strong Triadic Closure Property if it has strong ties to two other nodes B and C, and there is no edge at all (either a strong or weak tie) between B and C. We say that a node A satisfies the Strong Triadic Closure Property if it does not violate it.

- The above example satisfy this example.
- In real networks is may be too extreme to assume all node obey this property, but still helpful to reason about strong and weak ties

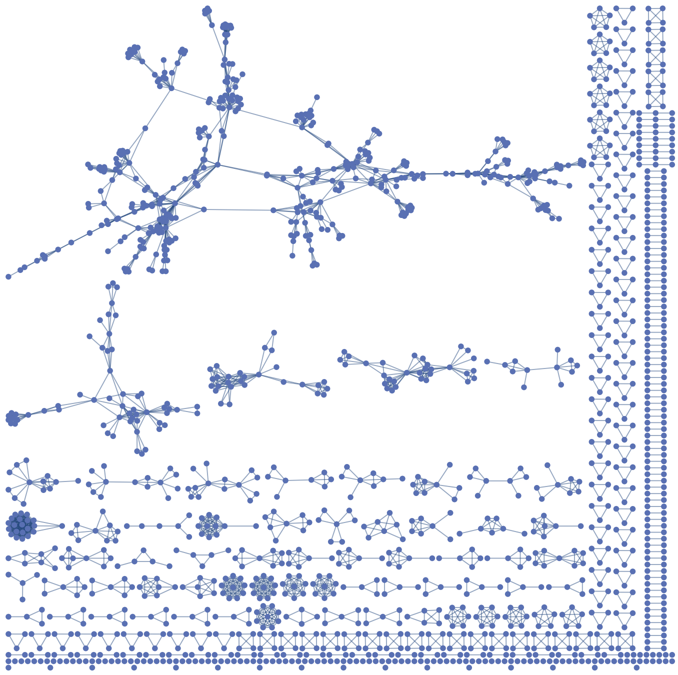
Example - strong and weak ties

In[29]:=
```
ExampleData[
 {"NetworkGraph", "CoauthorshipsInNetworkScience"}, "LongDescription"]
```

Out[29]=
```
Coauthorship network of scientists working on network theory and experiment,
  as compiled by M. Newman in May 2006. The edge weight indicates
  the strength of the collaboration between pairs of individuals.
```

In[30]:=
```
NS = ExampleData[{"NetworkGraph", "CoauthorshipsInNetworkScience"}]
```

Out[30]=



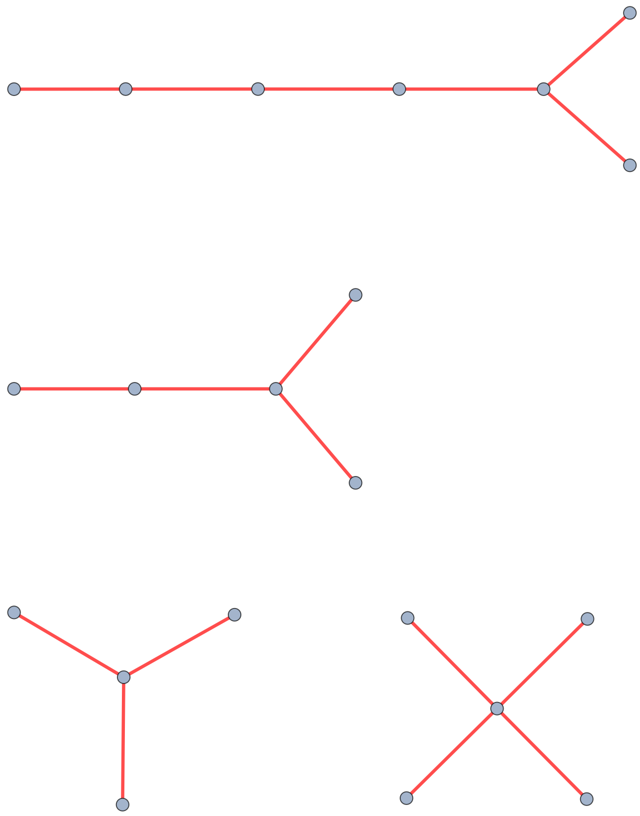
In[31]:=
```
NSnum = AdjacencyGraph[AdjacencyMatrix[NS]];
```

In[32]:=
```
StrongEdges[Mat_, limit_] := Module[{rules, T, T1},
  rules = ArrayRules[Mat];
  T = Table[If[rules〚i〛〚2〛 < limit, -1, rules〚i〛〚1〛], {i, Length[rules]}];
  T1 = DeleteCases[T, -1]]
TopCC[g_, k_] := Subgraph[g, Flatten[Take[WeaklyConnectedComponents[g], k]]];
```

In[34]:=
```
SE = StrongEdges[WeightedAdjacencyMatrix[NS], 2];
SG = TopCC[SimpleGraph[Graph[SE]], 4];
WG = Subgraph[NSnum, VertexList[SG]];
```

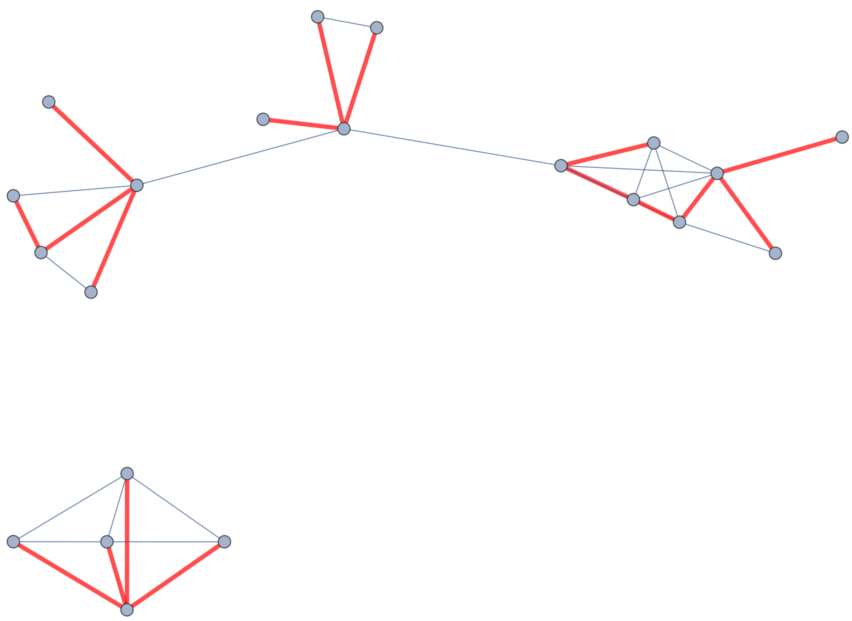
Graph of strong edges - 4 largest connected components

In[42]:=
```
HighlightGraph[SG, Style[EdgeList[SG], Thickness[.005], Red]]
```
Out[42]=



Same nodes with weak edges

In[41]:=
```
HighlightGraph[WG, Style[EdgeList[SG], Thickness[.005], Red]]
```
Out[41]=



Local Bridges and Weak Ties

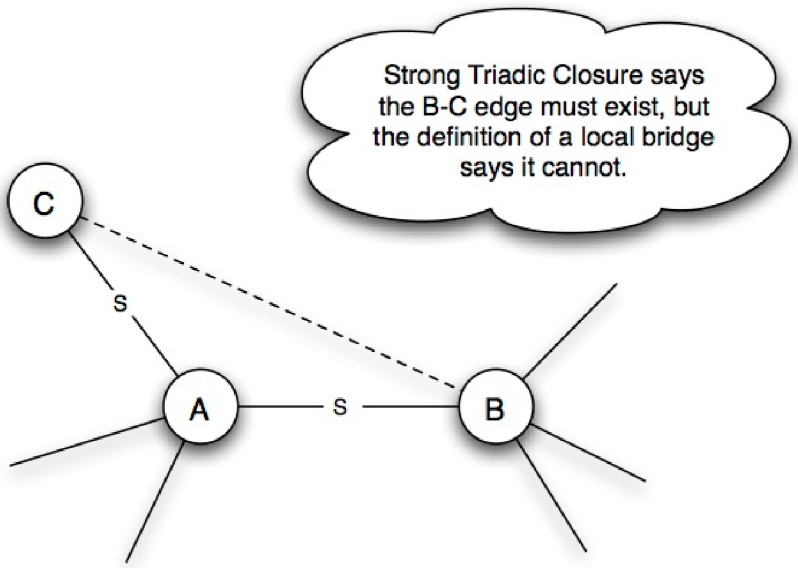
- 1) Local inter-personal distinction of links: weak or strong
- 2) Global, (network) structural notion: local bridge
- Using triadic closure we can connect them:

Claim: If a node *A* in a network satisfies the Strong Triadic Closure Property and is involved in at least two strong ties, then any local bridge it is involved in must be a weak tie.

- So local bridges in a network must be weak ties (under the above assumptions).

Proof of Claim

Claim: If a node *A* in a network satisfies the Strong Triadic Closure Property and is involved in at least two strong ties, then any local bridge it is involved in must be a weak tie.



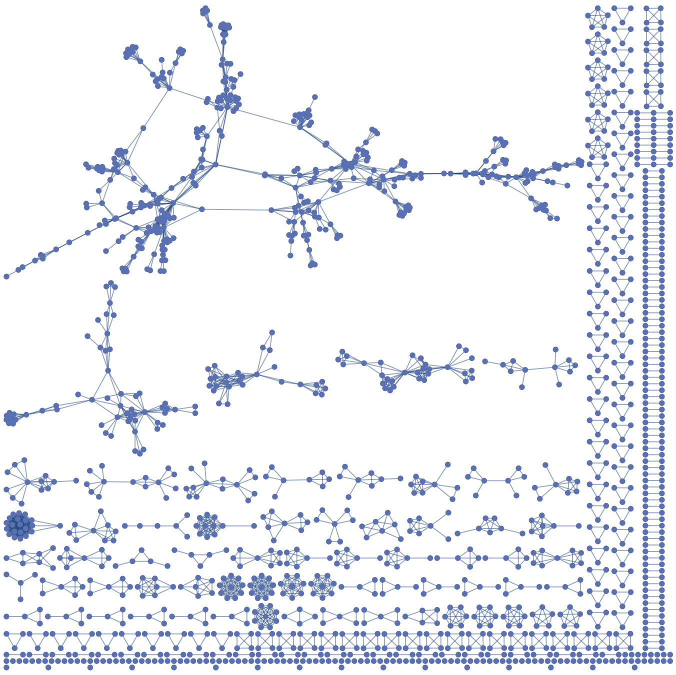
- Proof: Assume by contradiction that *A*-*B* is a local bridge and a strong tie.
- *A* must have another strong tie, say *C*.
- But, now the Strong Triadic Closure now say that an edge *B*-*C* must exist
- So *A*-*B* cannot be local bridge. Contradiction

Finding Jobs... (using weak ties)

- We found the connection between the local property of tie strength and the global property of serving as a local bridge
- New jobs are often rooted in contact with distant acquaintances
- “The argument is that these are the social ties that connect us to new sources of information and new opportunities, and their conceptual “span” in the social network (the local bridge property) is directly related to their weakness as social ties. This dual role as weak connections but also valuable conduits to hard-to-reach parts of the network — this is the surprising strength of weak ties.”

Example: The strength of weak ties

In[43]:=
```
NS
```
Out[43]=

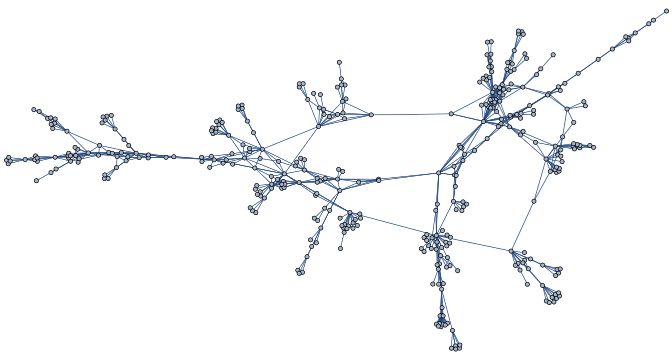


In[44]:=
```
WLCC[g_] := Subgraph[g, First[WeaklyConnectedComponents[g]]];
GC = WLCC[NS];
```

In[46]:=
```
WeightedGraphQ[SG]
```
Out[46]=
```
False
```

In[53]:=
```
GCEL = EdgeList[GC];
GCEW = PropertyValue[{NS, #}, EdgeWeight] & /@ EdgeList[GC];
```

In[49]:=
```
WG = Graph[GC, EdgeWeight → GCEW]
```
Out[49]=



In[50]:=
```
WeightedGraphQ[WG]
```
Out[50]=
```
True
```

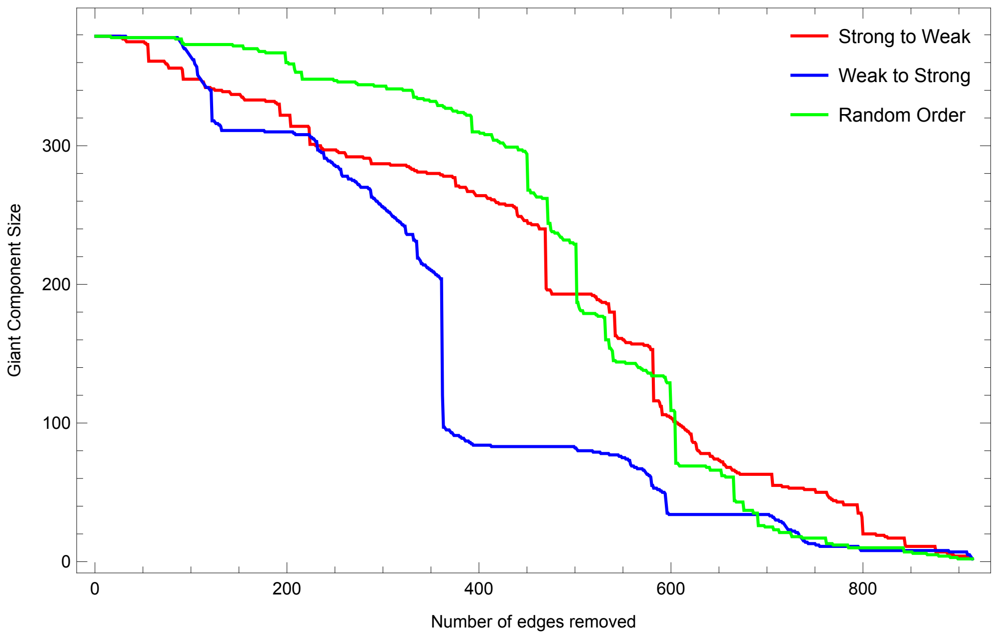
How do removing edges one-by-one reduce the Giant Component size?

Let’s see - Edge removal and giant component size

In[55]:=
```
EdgeWeightList[g_] := PropertyValue[{g, #}, EdgeWeight] & /@ EdgeList[g]
SortEL = GCEL⟦Ordering[EdgeWeightList[WG], All, Greater]⟧;
RandomEL = RandomSample[GCEL, EdgeCount[WG]];
```

In[59]:=
```
ListLinePlot[
 {Table[Length@First@WeaklyConnectedComponents[Graph[Take[SortEL, i]]],
   {i, Length[SortEL], 1, -1}],
  Table[Length@First@WeaklyConnectedComponents[Graph[Take[SortEL, -i]]],
   {i, Length[SortEL], 1, -1}],
  Table[Length@First@WeaklyConnectedComponents[Graph[Take[RandomEL, -i]]],
   {i, Length[RandomEL], 1, -1}]}, PlotStyle → {Red, Blue, Green}, PlotLegends →
  Placed[{"Strong to Weak", "Weak to Strong", "Random Order"}, {Right, Top}],
 FrameLabel → {"Number of edges removed", "Giant Component Size" },
 Frame → True]
```

Out[59]=



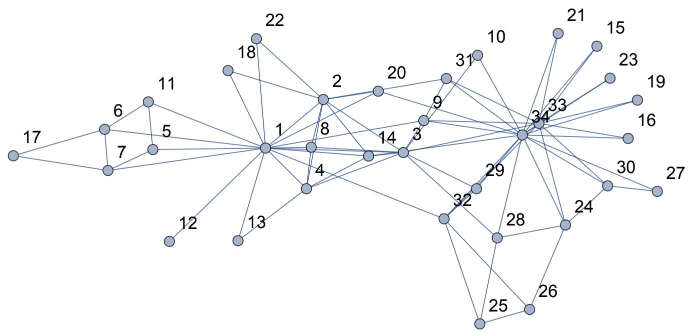
Neighborhood Overlap

- Generalizing the Notions of Weak Ties and Local Bridges
- Neighborhood overlap of *A-B:*
- $\text{NO}(A, B) = \dfrac{\text{number of nodes who are neighbors of both } A \text{ or } B}{\text{number of nodes who are neighbors of at least one of } A \text{ or } B}$
- When 0 it is a local bridge
- When “close” to 0, it is “close” to local bridge (weaker tie???)

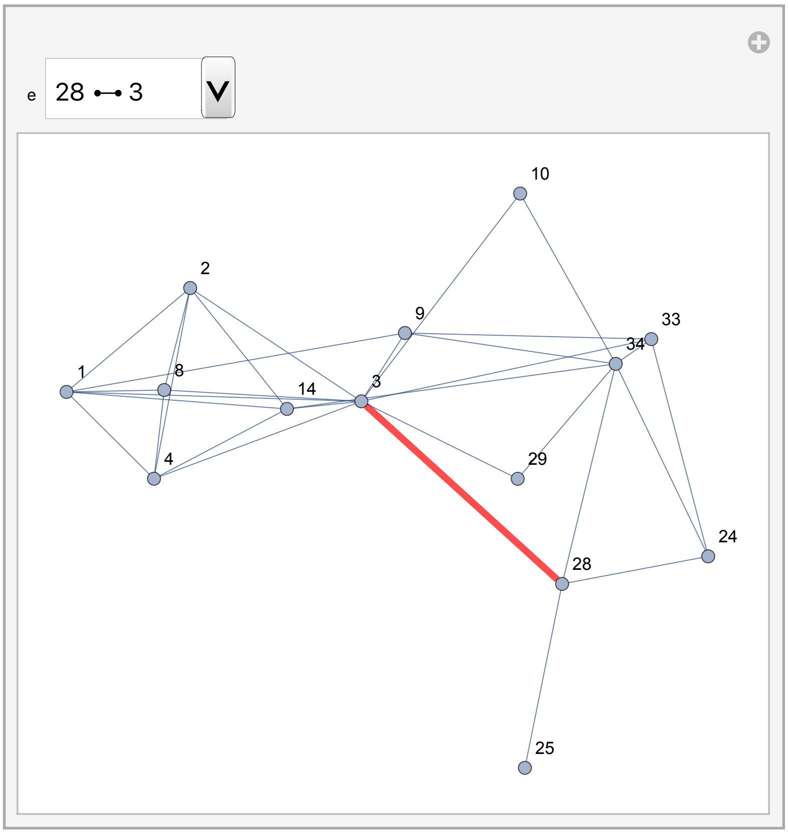
In[60]:=
```
Z = ExampleData[{"NetworkGraph", "ZacharyKarateClub"}];
```

In[61]:=
```
Graph[Z, VertexLabels → "Name"]
```

Out[61]=



In[64]:=
```
Manipulate[ShowOverlap[Z, e], {e, EdgeList[Z]}]
```
Out[64]=



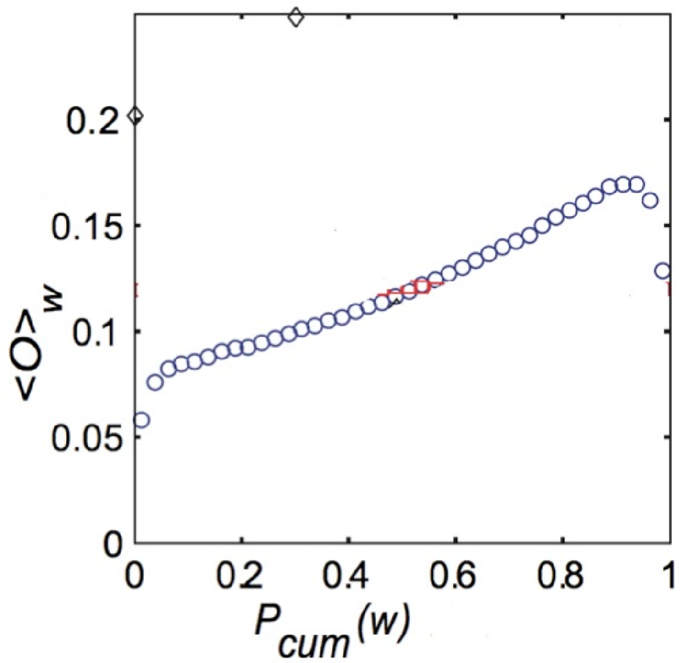
In[65]:=
```
NeighborhoodOL[Z, 1 ⟷ 11]
```
Out[65]=

$$\frac{2}{17}$$

- Some "Wrong Examples", Isolated edges

Empirical Results on Tie Strength and Neighborhood Overlap

- Phone calls duration dataset



A plot of the neighborhood overlap of edges as a function of their percentile in the sorted order of all edges by tie strength. The fact that overlap increases with increasing tie strength is consistent with the theoretical predictions (Image from Onenla et al.)

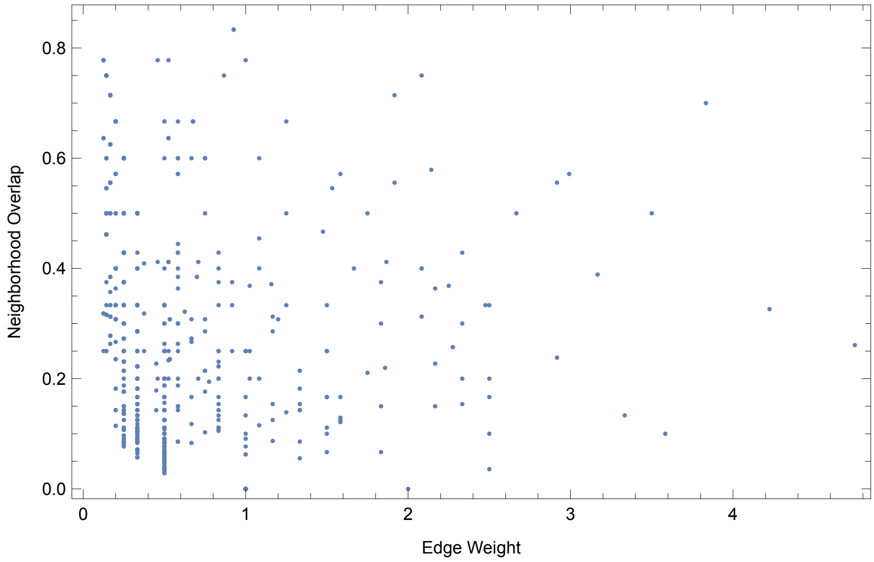
- In real data???

In[66]:=
```
EdgeOverList[g_] := NeighborhoodOL[g, #] & /@ EdgeList[g];
EdgeOverList[WG];
EdgeWeightList[WG];
```

In[67]:=
```
XYData = Transpose[{EdgeWeightList[WG], EdgeOverList[WG]}];
```

In[68]:=
```
ListPlot[XYData, PlotRange → All, Frame → True,
 FrameLabel → {"Edge Weight", "Neighborhood Overlap"}]
```

Out[68]=



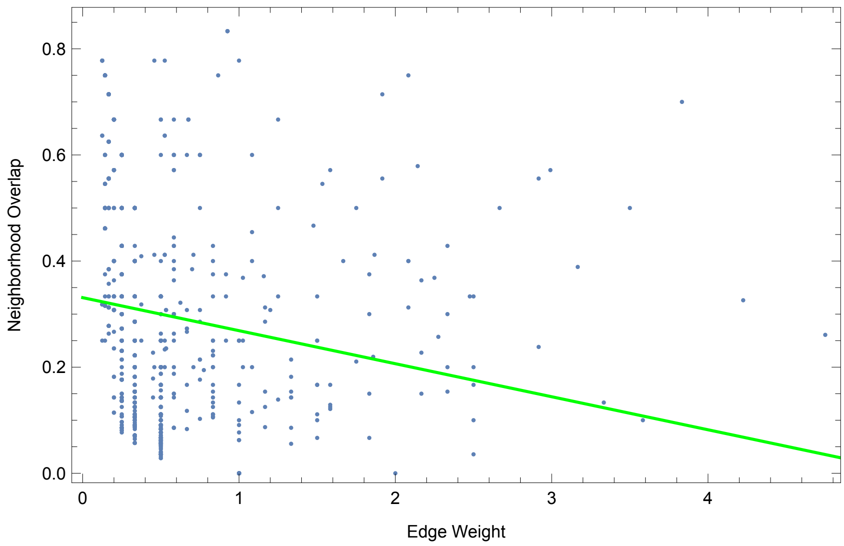
In[69]:=
```
line = Fit[XYData, {1, x}, x]
```

Out[69]=

0.331022 - 0.0622572 x

In[70]:=
```
Show[{ListPlot[XYData, PlotRange → All], Plot[line, {x, 0, 5}, PlotStyle → Green]},
 Frame → True, FrameLabel → {"Edge Weight", "Neighborhood Overlap"}]
```

Out[70]=



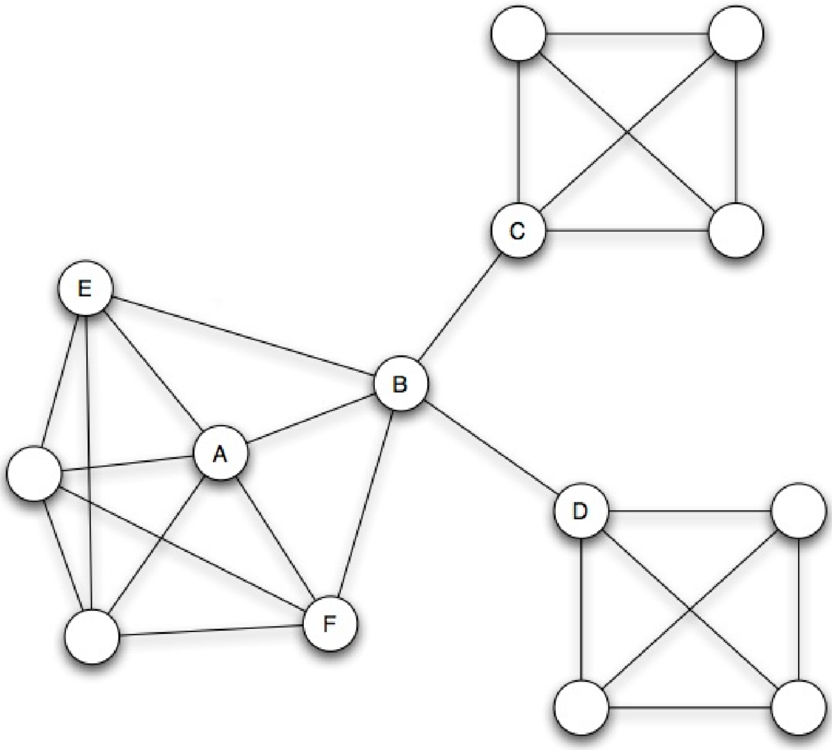
In[71]:=
```
WT = Tally[Sort[EdgeWeightList[WG]]]
```

Out[71]=
```
{{0.125, 10}, {0.142857, 47}, {0.166667, 58}, {0.2, 54}, {0.25, 148},
 {0.333333, 194}, {0.375, 3}, {0.45, 3}, {0.458333, 3}, {0.5, 173},
 {0.525, 8}, {0.533333, 4}, {0.583333, 25}, {0.625, 1}, {0.666667, 8},
 {0.67619, 3}, {0.7, 1}, {0.708333, 2}, {0.75, 10}, {0.775, 1}, {0.833333, 20},
 {0.866667, 1}, {0.916667, 3}, {0.92619, 3}, {1, 47}, {1.025, 3},
 {1.08333, 5}, {1.15833, 1}, {1.16667, 5}, {1.2, 1}, {1.25, 4}, {1.33333, 7},
 {1.47619, 1}, {1.5, 8}, {1.53333, 1}, {1.58333, 5}, {1.66667, 1}, {1.75, 2},
 {1.83333, 4}, {1.85833, 1}, {1.86667, 1}, {1.91667, 2}, {2, 1}, {2.08333, 4},
 {2.14286, 1}, {2.16667, 3}, {2.25, 1}, {2.275, 1}, {2.33333, 4}, {2.47619, 1},
 {2.5, 5}, {2.66667, 1}, {2.91667, 2}, {2.99167, 1}, {3.16667, 1},
 {3.33333, 1}, {3.5, 1}, {3.58333, 1}, {3.83333, 1}, {4.225, 1}, {4.75, 1}}
```

Bonus for this week

- Generate the figure “Edge removal and giant component size” for Marvel (weight ≥ 30) and Network of Thrones
- Generate the figure “neighborhood overlap as function weight” of for the Marvel (weighted) network for weight ≥ 30,and Network of Thrones. Does it works?

Closure, Structural Holes, and Social Capital



- Role of edges
- Role of nodes
- Nodes between groups and nodes in the middle of a group (A, B)
- General view: tightly-knit groups and the weak ties that link them
- **Embeddedness**
  - Node A has Hugh Clustering Coefficient
  - Embeddedness of an edge = the number of common neighbors the two endpoints have (recall neighborhood overlap)
  - For A all edged have high embeddedness...
  - high embeddedness --> trust
  - high embeddedness --> disallow misbehavior / potential sanctions
  - Low embeddedness --> riskier interactions / contradictory norms
- **Structural holes**
  - B’s location at the end of multiple local bridges also is a fundamental advantages

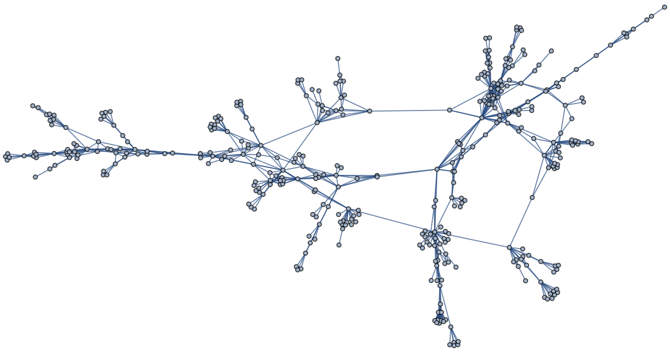
  - Research connects Individual’s success within a company to their access to local bridges
  - B span a **structural hole** (informal)
  - B’s advantages
    - Early access to information
    - Creativity - synthesis of multiple ideas
    - Social “gate- keeping”
  - Interests of node B and of the organization as a whole may not be aligned
- **Closure and Bridging as Forms of Social Capital**
- Social Capital: “the ability of actors to secure benefits by virtue of membership in social networks or other social structures”
- Hard to formalize
- physical capital, human capital, economic capital, cultural capital, etc.
- Group / individual property
- Is it the social interactions among the group’s members — or it is also based on the interactions of the group with the outside world???
- Tension between *closure* (embeddedness, triadic closure) and *brokerage* (bridging, weak ties)

Example

Co-authorship Graph

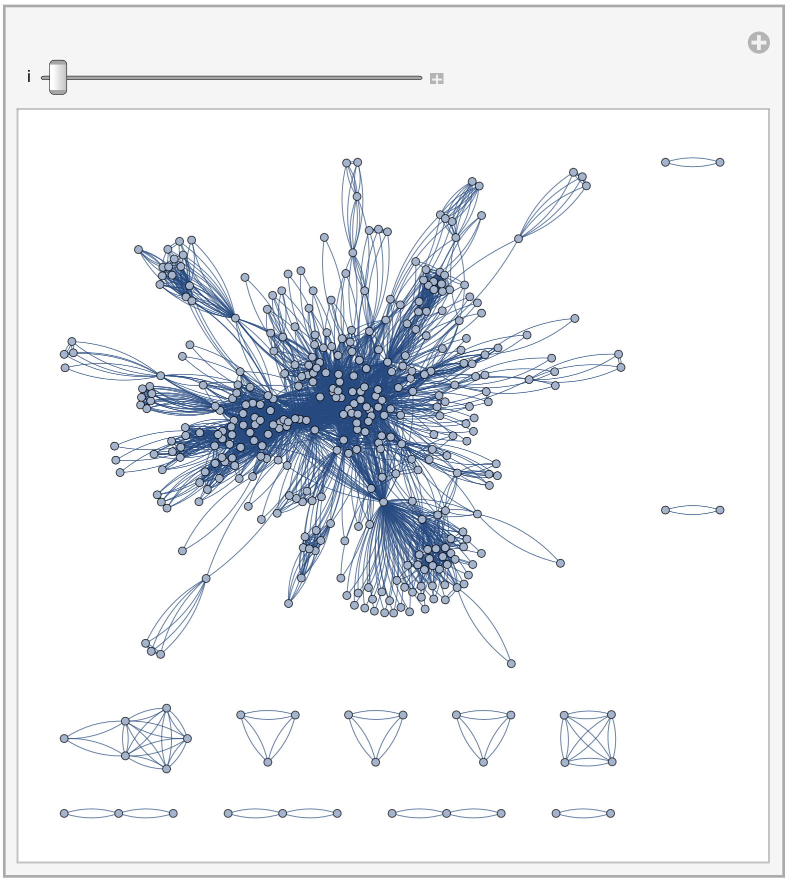
In[72]:=
```
WG = Graph[GC, EdgeWeight → GCEW]
```

Out[72]=



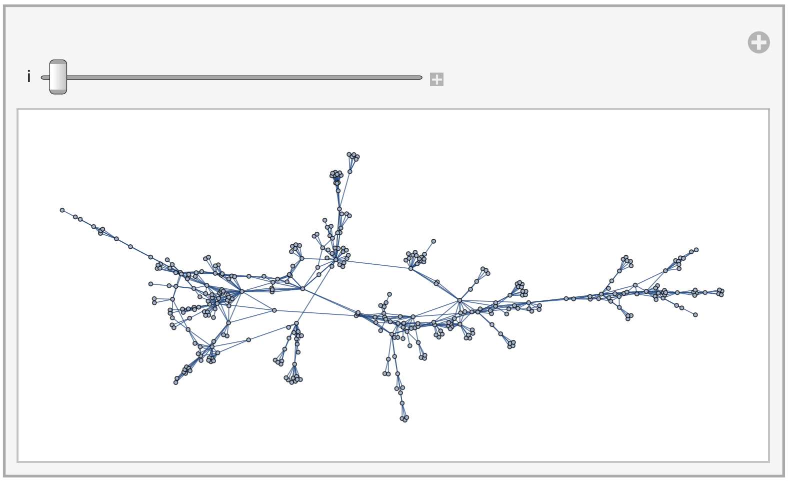
In[73]:=
```
Manipulate[Graph[Take[SortEL, -(Length[SortEL] - i)]], {i, 1, 500}]
```

Out[73]=

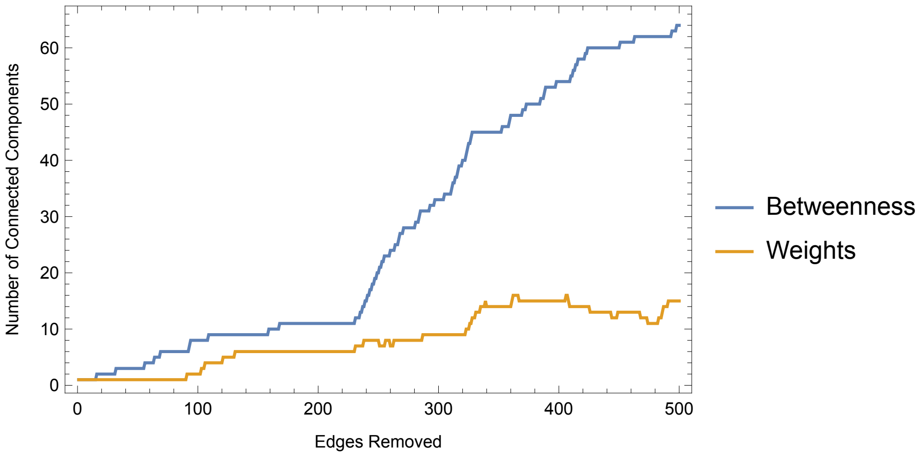


In[74]:=
```
BetweenEL = GCEL⟦Ordering[EdgeBetweennessCentrality[WG], All, Greater]⟧;
```

In[75]:=
```
Manipulate[Graph[Take[BetweenEL, -(Length[BetweenEL] - i)]], {i, 1, 500}]
```
Out[75]=



In[76]:=
```
ListLinePlot[{Table[Length[ConnectedComponents[
      Graph[Take[BetweenEL, -(Length[BetweenEL] - i)]]]], {i, 1, 500}],
    Table[Length[ConnectedComponents[Graph[Take[SortEL, -(Length[SortEL] - i)]]]],
     {i, 1, 500}]}, PlotLegends → {"Betweenness", "Weights"},
  FrameLabel → {"Edges Removed", "Number of Connected Components"},
  Frame → True]
```
Out[76]=



Talk about the missing edges assignment