Lecture 3: Strong and Weak Ties

Indiana University,

Emerson Melo

- One of the powerful roles that networks play is to bridge the local and the global-to offer explanations for how simple processes at the level of individual nodes and links can have complex effects that ripple through a population as a whole.
- How information flows through a social network, how different nodes can play structurally distinct roles in this process, and how these structural considerations shape the evolution of the network itself over time.
- We will begin with the famous "*strength of weak ties*" hypothesis from sociology

Emerson Melo (emelo@iu.edu)

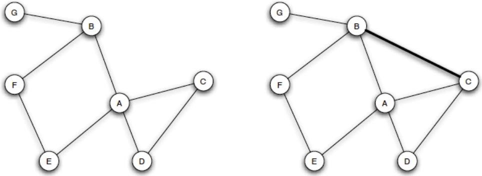
- As part of his Ph.D. thesis research in the late 1960s, Mark Granovetter interviewed people who had recently changed employers to learn how they discovered their new jobs.
- In keeping with earlier research, he found that many people learned information leading to their current jobs through personal contacts.
- But perhaps more strikingly, these personal contacts were often described by interview subjects as acquaintances rather than close friends.

- This is a bit surprising: your close friends presumably have the most motivation to help you when you're between jobs, so why is it so often your more distant acquaintances who are actually to thank for crucial information leading to your new job?
- The answer that Granovetter proposed to this question is striking in the way it links two different perspectives on distant friendships-one *structural*, focusing on the way these friendships span different portions of the full network; and the other *interpersonal*, considering the purely local consequences that follow from a friendship between two people being either strong or weak.

- In this way, the answer transcends the specific setting of job-seeking, and offers a way of thinking about the architecture of social networks more generally.

- In Lecture 2, our discussions of networks treated them largely as static structures-we take a snapshot of the nodes and edges at a particular moment in time, and then ask about paths, components, distances, and so forth.
- While this style of analysis forms the basic foundation for thinking about networks-and indeed, many datasets are inherently static, offering us only a single snapshot of a network-it is also useful to think about how a network evolves over time. In particular, what are the mechanisms by which nodes arrive and depart, and by which edges form and vanish?

- The precise answer will of course vary depending on the type of network we're considering, but one of the most basic principles is the following:
  " *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*"

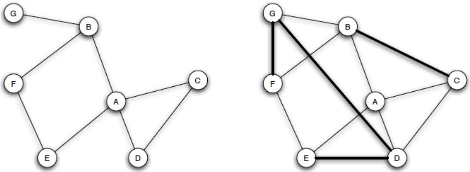


(a) *Before B-C edge forms.* (b) *After B-C edge forms.*

Figure 3.1: The formation of the edge between B and C illustrates the effects of triadic closure, since they have a common neighbor A.

- We refer to this principle as **triadic closure**, and it is illustrated in Figure 3.1: if nodes B and C have a friend A in common, then the formation of an edge between B and C produces a situation in which all three nodes A, B, and C have edges connecting each other-a structure we refer to as a *triangle in the network*.
- The term “triadic closure” comes from the fact that the $B-C$ edge has the effect of “closing” the third side of this triangle.

- If we observe snapshots of a social network at two distinct points in time, then in the later snapshot, we generally find a significant number of new edges that have formed through this triangle-closing operation, between two people who had a common neighbor in the earlier snapshot.



(a) *Before new edges form.* (b) *After new edges form.*

Figure 3.2: If we watch a network for a longer span of time, we can see multiple edges forming — some form through triadic closure while others (such as the D-G edge) form even though the two endpoints have no neighbors in common.

The Clustering Coefficient.

- The basic role of triadic closure in social networks has motivated the formulation of simple social network measures to capture its prevalence.
- One of these is the clustering coefficient.
- The clustering coefficient of a node A is defined as the probability that two randomly selected friends of A are friends with each other.
- In other words, it is the fraction of pairs of A's friends that are connected to each other by edges.

- For example, the clustering coefficient of node A in Figure 3.2(a) is $1/6$ (because there is only the single $C-D$ edge among the six pairs of friends $B-C$, $B-D$, $B-E$, $C-D$, $C-E$, and $D-E$), and it has increased to $1/2$ in the second snapshot of the network in Figure 3.2(b) (because there are now the three edges $B-C$, $C-D$, and $D-E$ among the same six pairs).

- In general, the clustering coefficient of a node ranges from 0 (when none of the node's friends are friends with each other) to 1 (when all of the node's friends are friends with each other), and the more strongly triadic closure is operating in the neighborhood of the node, the higher the clustering coefficient will tend to be.

Reasons for Triadic Closure.

- Triadic closure is intuitively very natural, and essentially everyone can find examples from their own experience.
- Moreover, experience suggests some of the basic reasons why it operates.
- One reason why B and C are more likely to become friends, when they have a common friend A, is simply based on the opportunity for B and C to meet: if A spends time with both B and C, then there is an increased chance that they will end up knowing each other and potentially becoming friends.

- A second, related reason is that in the process of forming a friendship, the fact that each of B and C is friends with A(provided they are mutually aware of this) gives them a basis for trusting each other that an arbitrary pair of unconnected people might lack.

- A third reason is based on the incentive A may have to bring B and C together: if A is friends with B and C, then it becomes a source of latent stress in these relationships if B and C are not friends with each other.
- This premise is based in theories dating back to early work in social psychology; it also has empirical reflections that show up in natural but troubling ways in public-health data.
- For example, Bearman and Moody have found that teenage girls who have a low clustering coefficient in their network of friends are significantly more likely to contemplate suicide than those whose clustering coefficient is high.

The Strength of Weak Ties.

- So how does all this relate to Mark Granovetter's interview subjects, telling him with such regularity that their best job leads came from acquaintances rather than close friends?
- In fact, triadic closure turns out to be one of the crucial ideas needed to unravel what's going on.

Bridges and Local Bridges.

- Let's start by positing that information about good jobs is something that is relatively scarce; hearing about a promising job opportunity from someone suggests that they have access to a source of useful information that you don't.
- Now consider this observation in the context of the simple social network drawn in Figure 3.3.

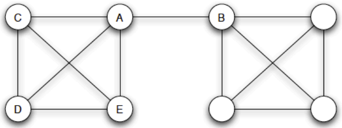


Figure 3.3: The A-B edge is a *bridge*, meaning that its removal would place A and B in distinct connected components. Bridges provide nodes with access to parts of the network that are unreachable by other means.

- The person labeled A has four friends in this picture, but one of her friendships is qualitatively different from the others: A's links to C, D, and E connect her to a tightly-knit group of friends who all know each other, while the link to B seems to reach into a different part of the network.
- We could speculate, then, that the structural peculiarity of the link to B will translate into differences in the role it plays in A's everyday life: while the tightly-knit group of nodes A, C, D, and E will all tend to be exposed to similar opinions and similar sources of information, A's link to B offers her access to things she otherwise wouldn't necessarily hear about.

- To make precise the sense in which the $A - B$ link is unusual, we introduce the following definition.
- We say that an edge joining two nodes A and B in a graph is a *bridge* if deleting the edge would cause A and B to lie in two different components.
- In other words, this edge is literally the only route between its endpoints, the nodes A and B.

- Now, if our discussion in Lecture 2 about giant components and small-world properties taught us anything, it's that bridges are presumably extremely rare in real social networks.
- You may have a friend from a very different background, and it may seem that your friendship is the only thing that bridges your world and his, but one expects in reality that there will be other, hard-to-discover, multi-step paths that also span these worlds.

- In other words, if we were to look at Figure 3.3 as it is embedded in a larger, ambient social network, we would likely see a picture that looks like Figure 3.4.

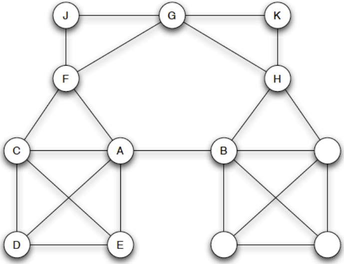


Figure 3.4: The A-B edge is a local bridge of span 4, since the removal of this edge would increase the distance between A and B to 4.

- We say that an edge joining two nodes A and B in a graph is a *local bridge* if its endpoints A and B have no friends in common-in other words, if deleting the edge would increase the distance between A and B to a value strictly more than two.
- We say that the span of a local bridge is the distance its endpoints would be from each other if the edge were deleted.

- Thus, in Figure 3.4, the $A-B$ edge is a local bridge with span four; we can also check that no other edge in this graph is a local bridge, since for every other edge in the graph, the endpoints would still be at distance two if the edge were deleted.

- Notice that the definition of a local bridge already makes an implicit connection with triadic closure, in that the two notions form conceptual opposites: *an edge is a local bridge precisely when it does not form a side of any triangle in the graph*.

- Local bridges, especially those with reasonably large span, still play roughly the same role that bridges do, though in a less extreme way.
- They provide their endpoints with access to parts of the network, and hence sources of information, that they would otherwise be far away from.
- And so this is a first network context in which to interpret Granovetter's observation about job-seeking: we might expect that if a node like A is going to get truly new information, the kind that leads to a new job, it might come unusually often (though certainly not always) from a friend connected by a local bridge.

The Strong Triadic Closure Property

- Granovetter's interview subjects didn't say, "*I learned about the job from a friend connected by a local bridge*."
- If we believe that local bridges were overrepresented in the set of people providing job leads, how does this relate to the observation that distant acquaintances were overrepresented as well?
- To talk about this in any detail, we need to be able to distinguish between different levels of *strength* in the links of a social network.

- Intuitively the idea is that stronger links represent closer friendship and greater frequency of interaction.
- In general, links can have a wide range of possible strengths, but for conceptual simplicity-and to match the friend/acquaintance dichotomy that we're trying to explain-we'll categorize all links in the social network as belonging to one of two types.
- strong ties (the stronger links, corresponding to friends), and weak ties (the weaker links, corresponding to acquaintances).

- Once we have decided on a classification of links into strong and weak ties, we can take a social network and annotate each edge with a designation of it as either strong or weak.
- For example, assuming we asked the nodes in the social network of Figure 3.4 to report which of their network neighbors were close friends and which were acquaintances, we could get an annotated network as in Figure 3.5.

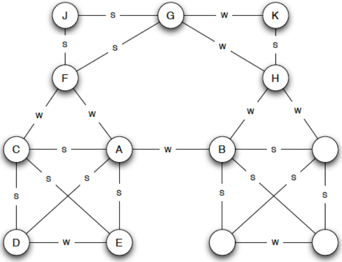


Figure 3.5: Each edge of the social network from Figure 3.4 is labeled here as either a *strong tie* (S) or a *weak tie* (W), to indicate the strength of the relationship. The labeling in the figure satisfies the Strong Triadic Closure Property at each node: if the node has strong ties to two neighbors, then these neighbors must have at least a weak tie between them.

- It is useful to go back and think about triadic closure in terms of this division of edges into strong and weak ties.
- If we recall the arguments supporting triadic closure, based on opportunity, trust, and incentive, they all act more powerfully when the edges involved are strong ties than when they are weak ties.

- This suggests the following qualitative assumption:
  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.

- To enable some more concrete analysis, Granovetter suggested a more formal (and somewhat more extreme version) of this, as follows.
  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.

- You can check that no node in Figure 3.5 violates the Strong Triadic Closure Property, and hence all nodes satisfy the Property.
- On the other hand, if the $A - F$ edge were to be a strong tie rather than a weak tie, then nodes A and F would both violate the Strong Triadic Closure Property: Node A would now have strong ties to nodes E and F without there being an $E - F$ edge, and node F would have strong ties to both A and G without there being an $A - G$ edge.
- As a further check on the definition, notice that with the labeling of edges as in Figure 3.5, node H satisfies the Strong Triadic Closure Property: H couldn't possibly violate the Property since it only has a strong tie to one other node.

- Clearly the Strong Triadic Closure Property is too extreme for us to expect it hold across all nodes of a large social network.
- But it is a useful step as an abstraction to reality, making it possible to reason further about the structural consequences of strong and weak ties.

Local Bridges and Weak Ties.

- We now have a purely local, interpersonal distinction between kinds of links-whether they are weak ties or strong ties-as well as a global, structural notion-whether they are local bridges or not.
- On the surface, there is no direct connection between the two notions, but in fact using triadic closure we can establish a connection, in the following 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.

- In other words, assuming the Strong Triadic Closure Property and a sufficient number of strong ties, the local bridges in a network are necessarily weak ties.

- Take some network, and consider a node A that satisfies the Strong Triadic Closure Property and is involved in at least two strong ties.
- Now suppose A is involved in a local bridge-say, to a node B-that is a strong tie.
- We want to argue that this is impossible, and the crux of the argument is depicted in Figure 3.6.
- First, since A is involved in at least two strong ties, and the edge to B is only one of them, it must have a strong tie to some other node, which we'll call C.
- Now let's ask: is there an edge connecting B and C?
- Since the edge from A to B is a local bridge, A and B must have no friends in common, and so the $B-C$ edge must not exist.

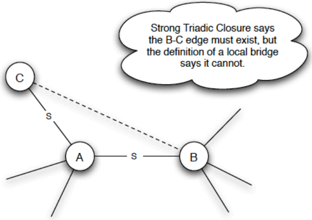


Figure 3.6: If a node satifies Strong Triadic Closure and is involved in at least two strong ties, then any local bridge it is involved in must be a weak tie. The figure illustrates the reason why: if the A-B edge is a strong tie, then there must also be an edge between B and C, meaning that the A-B edge cannot be a local bridge.

- But this contradicts Strong Triadic Closure, which says that since the $A-B$ and $A-C$ edges are both strong ties, the $B-C$ edge must exist.
- This contradiction shows that our initial premise, the existence of a local bridge that is a strong tie, cannot hold, finishing the argument.

Generalizing the Notions of Weak Ties and Local Bridges.

- An "almost" local bridge can be defined as the neighborhood overlap of an edge connecting A and B to be the ratio

$$\frac{\text{number of nodes who are neighbors of both A and B}}{\text{number of nodes who are neighbors of at least one of A or B}},$$

where in the denominator we don't count A or B themselves (even though A is a neighbor of B and B is a neighbor of A).

- As an example of how this definition works, consider the edge $A - F$ in Figure 3.4.
- The denominator of the neighborhood overlap for $A - F$ is determined by the nodes B, C, D, E, G, and J, since these are the ones that are a neighbor of at least one of A or F.
- Of these, only C is a neighbor of both A and F, so the neighborhood overlap is $1/6$.

- The key feature of this definition is that this ratio in question is 0 precisely when the numerator is 0, and hence when the edge is a local bridge.
- So the notion of a local bridge is contained within this definition-local bridges are the edges of neighborhood overlap 0-and hence we can think of edges with very small neighborhood overlap as being "almost" local bridges. (Since intuitively, edges with very small neighborhood overlap consist of nodes that travel in "social circles" having almost no one in common.)

Empirical Results on Tie Strength and Neighborhood Overlap.

- Using these definitions, we can formulate some fundamental quantitative questions based on Granovetter's theoretical predictions.
- First, we can ask how the neighborhood overlap of an edge depends on its strength; the strength of weak ties predicts that neighborhood overlap should grow as tie strength grows.

- Figure 3.7 shows the neighborhood overlap of edges as a function of their percentile in the sorted order of all edges by tie strength.
- Thus, as we go to the right on the x-axis, we get edges of greater and greater strength, and because the curve rises in a strikingly linear fashion, we also get edges of greater and greater neighborhood overlap.
- The relationship between these quantities thus aligns well with the theoretical prediction.

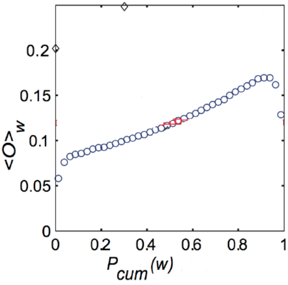


Figure 3.7: A plot of the neighborhood overlap of edges as a function of their percentile in the sorted order of all edges by tie strength. The fact that overlap increases with increasing tie strength is consistent with the theoretical predictions from Section 3.2. (Image from [334].)

Tie Strength, Social Media, and Passive Engagement

- As an increasing amount of social interaction moves on-line, the way in which we maintain and access our social networks begins to change as well.
- For example, as is well-known to users of social-networking tools, people maintain large explicit lists of friends in their profiles on these sites-in contrast to the ways in which such friendship circles were once much more implicit, and in fact relatively difficult for individuals even to enumerate or mentally access.
- What effect does this have on social network structure more broadly?

- Understanding the changes arising from these forms of technological mediation is a challenge that was already being articulated in the early 1990s by researchers, as the Internet began making remote interaction possible for a broad public; these issues have of course grown steadily more pervasive between then and now.
- Tie strength can provide an important perspective on such questions, providing a language for asking how on-line social activity is distributed across different kinds of links-and in particular, how it is distributed across links of different strengths.

- When we see people maintaining hundreds of friendship links on a social-networking site, we can ask how many of these correspond to strong ties that involve frequent contact, and how many of these correspond to weak ties that are activated relatively rarely.
- Researchers have begun to address such questions of tie strength using data from some of the most active social media sites.

- At Facebook, Cameron Marlow and his colleagues analyzed the friendship links reported in each user's profile, asking to what extent each link was actually used for social interaction, beyond simply being reported in the profile.
- In other words, where are the strong ties among a user's friends?
- To make this precise using the data they had available, they defined three categories of links based on usage over a one-month observation period.

- A link represents reciprocal (mutual) communication, if the user both sent messages to the friend at the other end of the link, and also received messages from them during the observation period.
- A link represents one-way communication if the user sent one or more messages to the friend at the other end of the link (whether or not these messages were reciprocated).
- A link represents a maintained relationship if the user followed information about the friend at the other end of the link, whether or not actual communication took place; "following information" here means either clicking on content via Facebook's News Feed service (providing information about the friend) or visiting the friend's profile more than once.

- Notice that these three categories are not mutually exclusive ? indeed, the links classified as reciprocal communication always belong to the set of links classified as one-way communication.
- This stratification of links by their use lets us understand how a large set of declared friendships on a site like Facebook translates into an actual pattern of more active social interaction, corresponding approximately to the use of stronger ties.

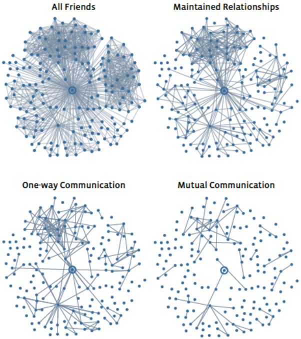


Figure 3.8: Four different views of a Facebook user's network neighborhood, showing the structure of links coresponding respectively to all declared friendships, maintained relationships, one-way communication, and reciprocal (i.e. mutual) communication. (Image from [286].)

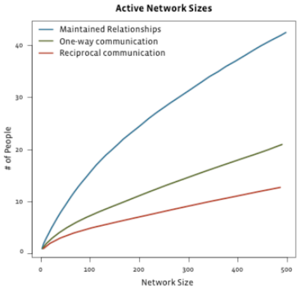


Figure 3.9: The number of links corresponding to maintained relationships, one-way communication, and reciprocal communication as a function of the total neighborhood size for users on Facebook. (Image from [286].)

- On the x-axis is the total number of friends a user declares, and the curves then show the (smaller) numbers of other link types as a function of this total.
- There are several interesting conclusions to be drawn from this.
- First, it confirms that even for users who report very large numbers of friends on their profile pages (on the order of 500), the number with whom they actually communicate is generally between 10 and 20, and the number they follow even passively (e.g. by reading about them) is under 50.

- But beyond this observation, Marlow and his colleagues draw a further conclusion about the power of media like Facebook to enable this kind of passive engagement, in which one keeps up with friends by reading news about them even in the absence of communication.
- But beyond this observation, Marlow and his colleagues draw a further conclusion about the power of media like Facebook to enable this kind of *passive engagement*, in which one keeps up with friends by reading news about them even in the absence of communication.

- They argue that this passive network occupies an interesting middle ground between the strongest ties maintained by regular communication and the weakest ties from one?s distant past, preserved only in lists on social-networking profile pages.
- They write, "*The stark contrast between reciprocal and passive networks shows the effect of technologies such as News Feed. If these people were required to talk on the phone to each other, we might see something like the reciprocal network, where everyone is connected to a small number of individuals. Moving to an environment where everyone is passively engaged with each other, some event, such as a new baby or engagement can propagate very quickly through this highly connected network.*"