Economics of Information Networks

Stephen Turnbull

Division of Policy and Planning Sciences

Lecture 3c: December 3, 2020

Abstract

We look at some basic structures of graphs, and applications to social networks, including Granovetter's famous argument about *The Strength of Weak Ties* and approximate network partitioning.

In Part 3, we look at Granovetter's arguments about "the strength of weak ties," and at some quantitative analysis of social network data.

The strength of weak ties

- Granovetter [1973] observed that people frequently learned of opportunities such as job opportunities through personal contacts who were *not* considered friends. That is, the "tie" to these acquaintances was "weak."
- Based on the process of triadic closure, we would expect that
  1. links to "acquaintances" are bridges
  2. friends share much information in common
  3. however, friendly competition is a thing

  concluding that information about such opportunities often would flow from acquaintances naturally.
- Why doesn't triadic closure operate with acquaintances? We propose a *weighted graph* model with strong and weak links. *Strong links* correspond to friends, and *weak links* to acquaintances.

Strong triadic closure

- "Important opportunities" like mid-career job changes are infrequent, so triadic closure converts (local) bridges to non-bridges relatively quickly. We expect information flow via bridges is uncommon.
- The words "friend" and "acquaintance" suggest a solution: a weighted graph, where some links are *strong* and others *weak*.
- *Strong triadic closure* if whenever strong links exist between A and B and between A and C, there is a link (strong *or* weak) between B and C.
- **Theorem:** In a strong triadically closed graph, if A has any other strong links, and is the endpoint of a (local) bridge, the bridge is a *weak* link.
  - Note: the interpretation of a local bridge as an "acquaintance" makes a lot of sense: they are both not part of the (not necessarily strong) triadic closed group of friends and local bridge is a weak link.

Strong triadic closure

- Important opportunities arise infrequently
- Local bridges should convert quickly
- Information shouldn't flow by bridges
- Friends, not acquaintances
- Need to justify acquaintances (weak ties)
- This is why we generalize to a weak-strong graph, weak = acquaintance, strong = friend

A note on "strong"

- "Strong version of a condition" *vs.* usage of "strong" in "strong triadic closure."
  - A *strong condition* usually is more restrictive, and applies to fewer cases.
  - *Strong triadic closure* is the opposite; the condition is less restrictive than in triadic closure (it only applies when the links are both strong, so more graphs are strong triadic closed than are triadic closed).

Graph structure corresponds to real phenomena

- Theorem: *if*

  1. graph is strong triadic closed
  2. node A has a bridge
  3. node A has another strong link

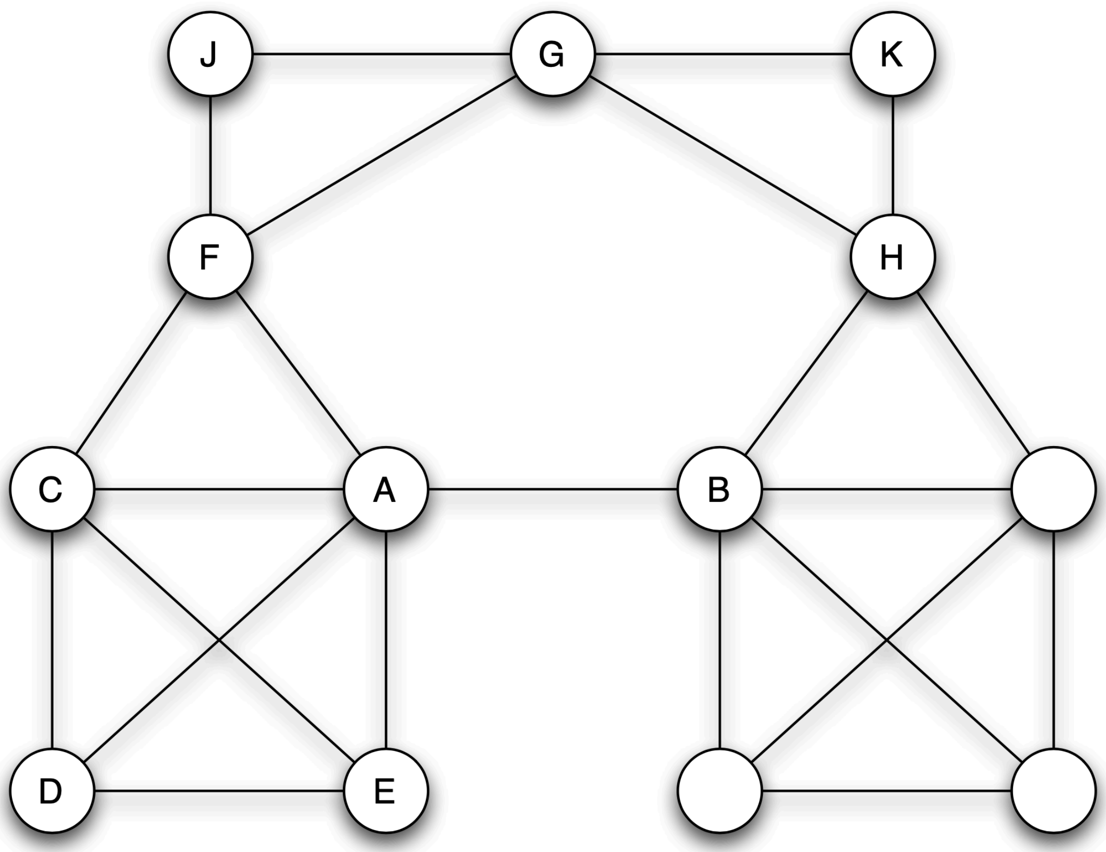
  then the bridge is a weak link.

Interpretation

- Strong triadic closure implies group of friends is closed
- So bridges allow weak links to non-friends (acquaintences)
- The theorem is purely mathematical, but it implies something that we now know about graph structure from Granovetter, namely that there is reason to expect bridges to weak links.
- This is what mathematics is for: helping us understand *why* certain properties of behavior tend to "cluster."

Discussion of graph with local bridge

- Show that graph is strongly triadically closed, though not triadically closed.
- Point out gatekeepers (several).
- Interpretation as members of a college class, with connections across companies.



"Smoothing" strength

- We needed to distinguish strong and weak links to analyze Granovetter's observation. An obvious generalization that is sometimes useful is to make strength a numerical quantity.
  - Practical measures of link strength in communication networks include number of calls, and cumulative length of calls, in a fixed period.
  - In surveys, we might ask whether a person is *unrelated*, an *acquaintance*, or a *friend*.

“Smoothing” bridges

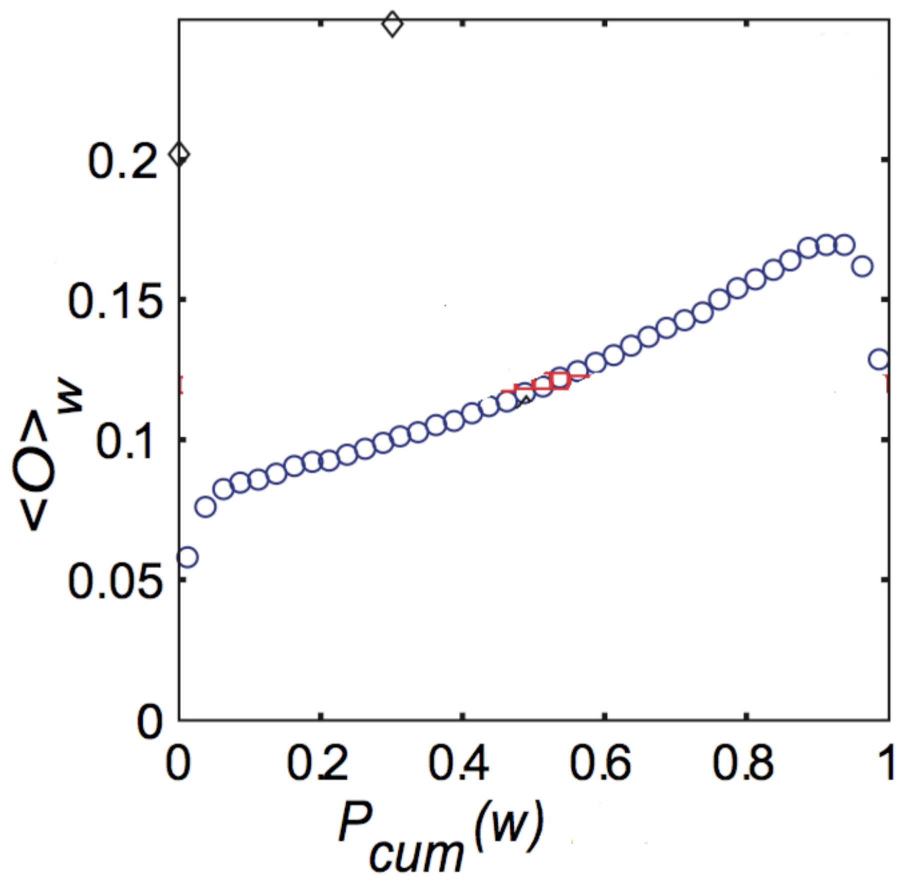
- In many practical applications, only a fraction of edges are local bridges.
- We can make this smoother by defining *neighborhood overlap* of an edge AB as

$$\frac{|N(A) \cap N(B)|}{|N(A) \cup N(B)|}$$

  where $N(X)$ is the set of nodes that are neighbors of X, not including A and B themselves. Note that with this definition, a link has neighborhood overlap of zero exactly when it is a local bridge.
- In data sets where we have a quantitive measure of strength, it often shows a correlation with neighborhood overlap.

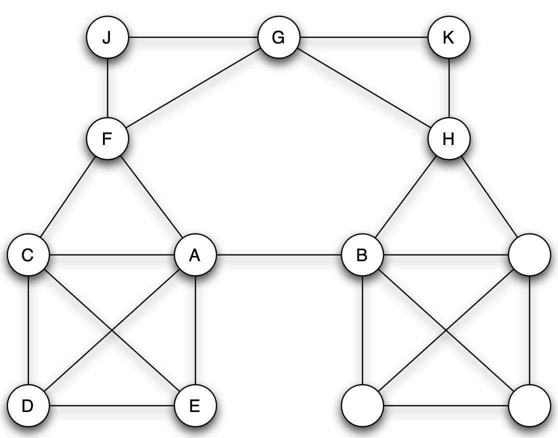
Overlap *vs.* strength in cellphone calls

Onnela *et al.* [2007] worked with cellphone data. An edge exists if two cellphones made calls in both directions in a month. The strength of an edge is measured by its percentile in the distribution of number of minutes used.



Giant component in cellphone data

- Giant components usually don't have many local bridges.
- People usually have many correspondents, and they will typically converse with several. Seems very likely that will be substantial overlap among neighbors (triadic closure).
- However, we expect to see weak ties that link "communities."
  - In the bridge figure we can see three or four such triadically closed "communities."
  - *E.g.* a company with several departments.



Decomposing the giant component

- Onnela *et al.* deleted links one at a time, starting with the strongest.
  - The giant component gradually shrinks as individuals lose all their ties.
- Now try the same process, starting with the weakest.
  - The giant component shrinks more quickly as individuals lose all their ties, but
  - also because it fragments into smaller components when bridges are deleted, and some of these bridges link components of similar size.
  - The fragmentation doesn't happen until late in the strongest first process because these are much less likely to be bridges (or in general have low neighborhood overlap).
- This strongly suggests that weak links join "communities."

Research opportunities

- The theoretical work is not fully worked out.
  - Some advanced mathematics is used, but there remain important concepts (such as "community") that remain undefined.
  - Probably there are elementary theorems (like the one about strong triadic closure and the weakness of bridges) to be proved.
- Experiments like changing the order of link deletion according to link strength often clarify the nature of concepts. They are valuable contributions that non-mathematicians can make..
  - (Homework) What do you think would happen in case of deletion in order of decreasing neighborhood overlap? Increasing order?
- Many opportunities to study real networks, as well as to develop theorems that describe why they behave (show relationships) as they do.

Applications of tie strength

- Besides strong/weak, what can we say?
- Try different ways of measuring strength, and how that affects results.
- Social networks allow examination of relationships in ways that subjects define, so are not artifacts of the research design.
  - Consider office colleagues one of whom is a salesperson, they may call each other a lot. But each may rarely call their wives, because they can't do so at work.
- Relationships like Facebook friend or Twitter follower are declared by users themselves, and have well-defined semantics according to the rules of the social network.
  - Be careful: Social network usage of words like "friend" corresponds imperfectly to the usual usage in everyday language.

Friends on Facebook

- Facebook friends aren't necessarily real-life friends, but this is easier to deal with than the reification of "friend" as measured by number of phone calls.
- On the other hand, Facebook friendship is necessarily symmetric, which is not always true in real life.
- With Facebook, you can get an exact list of friends from the Facebook API. Not necessarily true of people trying to list real-life friends.
  - However, in the case of the karate club, there's a small universe known to the research so can be more accurate (*e.g.*, give a list).

Behavior on social networks

- Existence of flamewars on email lists and newsgroups starting from the 1980s.
  - Even virtual violence (deleting files, for example).
- Similar existence of flaming, trolling, doxxing on modern social networks.
- Social networks perhaps even more flammable than mailing lists, since it is (at least, can be) more public.
- Are they really different?
- Are behavioral differences due to changes in individuals' thinking and behavior, or is it emerging from network structure?
  - *E.g.*, text-based communication does not express subtleties of emotion (even with smilies and emoji).

A remark on Twitter feeds

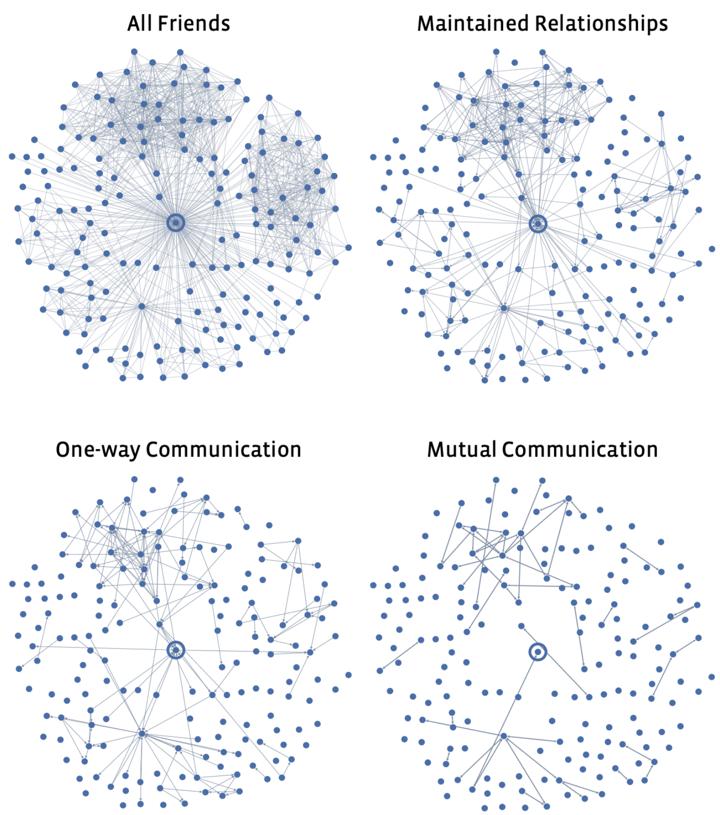
- Compare personal experience. The "algorithm" is part of the network structure.
  - People I follow don't seem to post much. Is that because they aren't posting much or because Twitter isn't putting them in my timeline because they don't seem to post things I'm interested in?
  - I don't like Trump, and I don't like my Republican Senator. But I follow my Senators and Representative to be well-informed in future elections. However, I see only nasty opposition in the replies to the Republican Senator. Is that all there is, or has Twitter figured out I don't like him and so only shows the nasty replies?

Interaction on Facebook

- Some researchers looked at intensity of interaction among Facebook users.
- When people have Facebook friends, do they actually interact with them?
- Some people use Facebook to coordinate a group. Interaction looks small, although the group make interact offline frequently and intensely.
- Others post comments each way (*e.g.*, a family spread out geographically.
- Researchers defined three degrees of communication between friends:
  - reciprocal communication (each member sent messages to the other),
  - one-way communication (only one member has sent a message to the other), and
  - maintained relationship (one way), clicking on content or multiple profile visits.

Apparent nesting of relationships by strength

- We look at the *network of neighbors of one member*. The graph shows the various graphs associated with a particular member.
- According to the definitions in this research, *every* reciprocal communication link is also a one-way communication link.

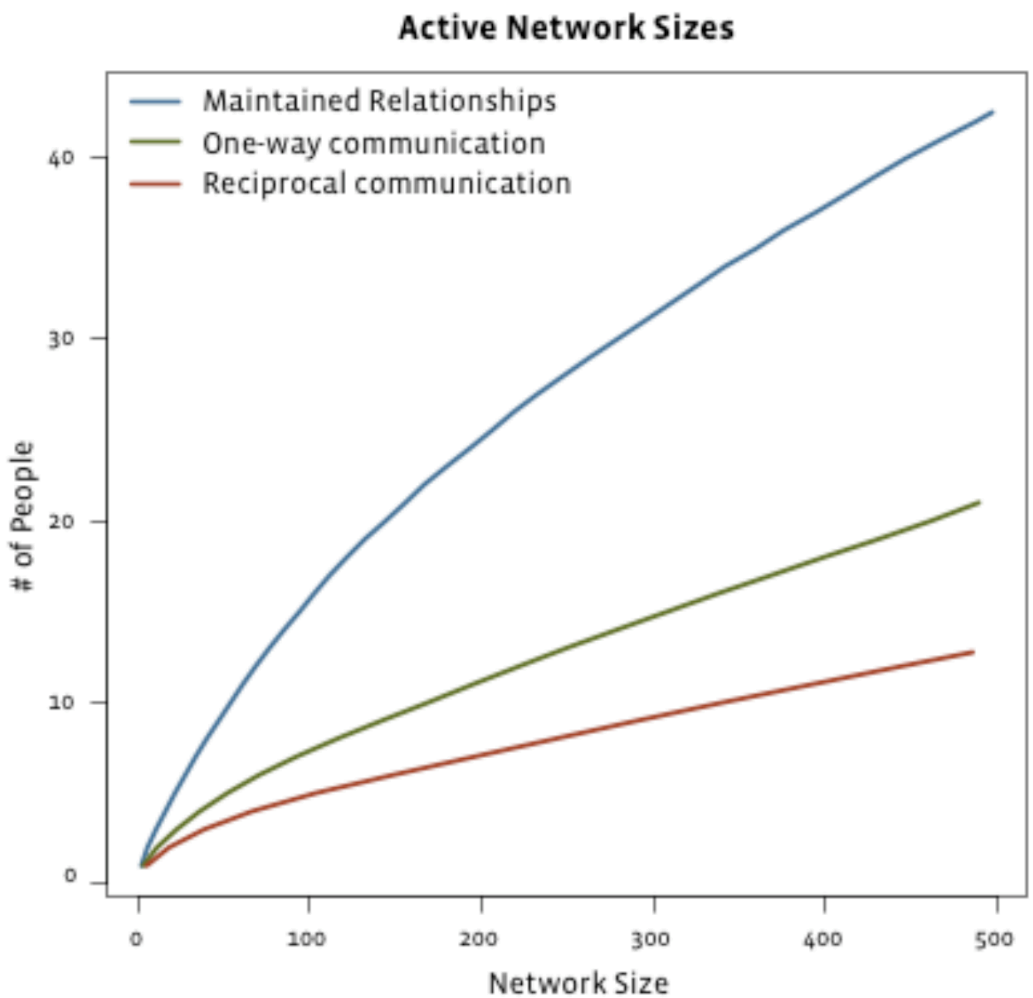


Analysis of nesting of relationships

- For this member, it's easy to visually confirm that the one-way communication links are pretty much all maintained relationship links. I did not expect this. In using Twitter, there are people whose profiles I check *because* I don't "at" them, get "at"-ed by them, *etc.*
- Thinning takes place by "area," not uniformly in the graph. Suggests multiple mechanisms for finding new relationships.
  - Can this be measured statistically? If so, look at distribution of number of mechanisms.

Friendship strength

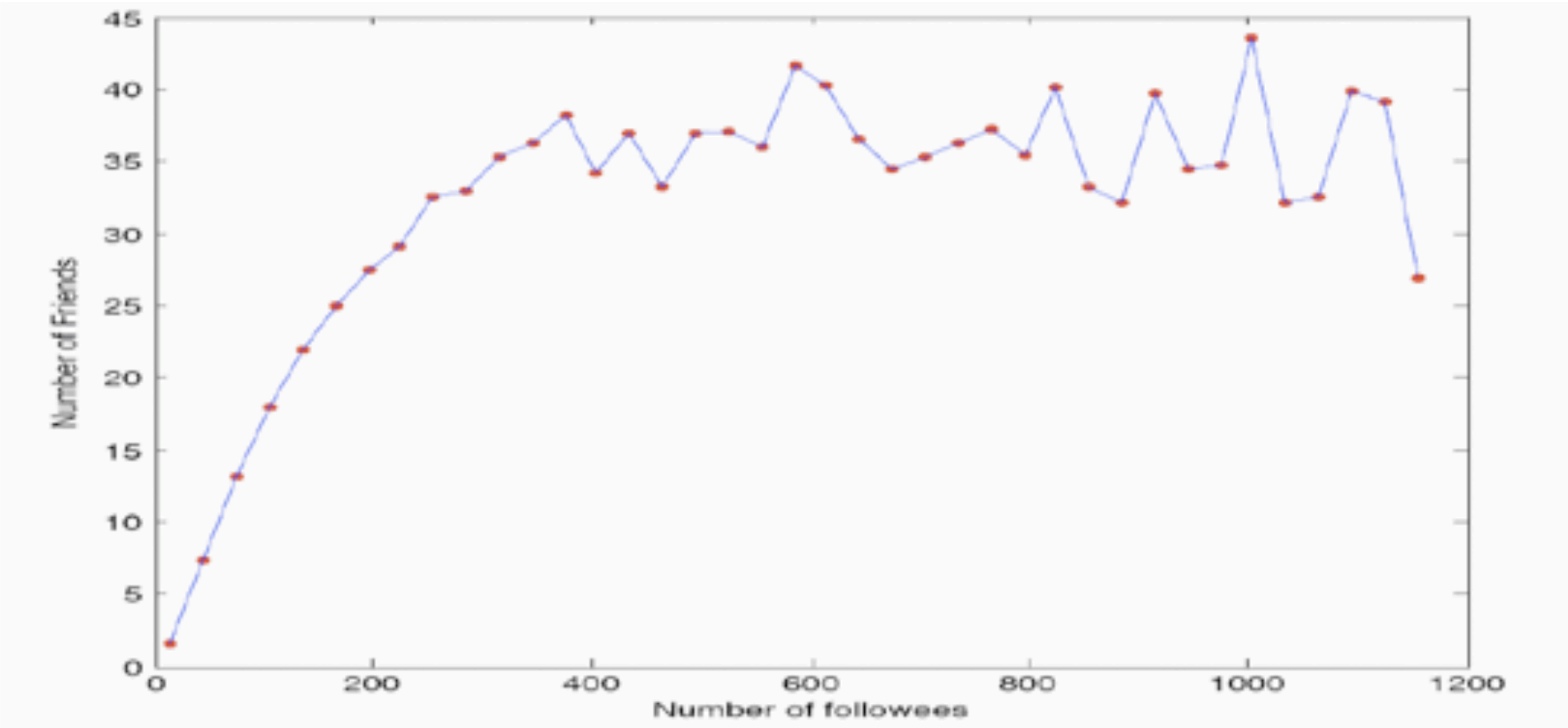
- The graph shows the average count of each kind of link over the members with a particular number of friends.



Analysis of friendship strength

- We see that on average the quantitative relationship is verified.
  - This does *not* mean that the nesting relationship is verified.
  - Reciprocal links are logically a subset of the one-way links, but
  - I would like to see some statistics on the percentage of maintained links that are also one-way and reciprocal links.
- The book points out that even people with 500 friends on Facebook on average limit themselves to about 40 maintained links, 20 one-way links, and 10 mutual links.

Increasing friendship strength on Twitter



Analysis of Twitter friendship strength

- Even people with 500 friends on Facebook on average limit themselves to about 40 maintained links, 20 one-way links, and 10 mutual links.
- I'm more interested that the asymptotes do not seem to be horizontal: people with more friends have more strong relationships of each type.
- This is different for Twitter: the asymptote is horizontal, there seems to be a cap on strong ties.
  - Not clear what this says: could be a limitation of Twitter clients.
  - Could be a limitation of the methodology (no measure of attention paid to tweets, which corresponds to maintained relationships).
  - It could be a different kind of social network, different style of communication: *needs more research*.