The Seven Bridges of Königsberg

Tommaso Venturini & Mathieu Jacomy

Although the history of networks as scientific objects begins with Euler's famous walk on Königsberg's bridges (*Solutio problematis ad geometriam situs pertinentis*, 1736), the notion of 'bridge' has rarely been tackled by network theorists. Among the few articles that took bridges seriously, the most famous is probably Mark Granovetter's paper on *The Strength of Weak Ties* (1973). Despite the huge influence of this paper (according to Google Scholars™ it has been cited more than 17.000 times), few have remarked that its most original insights concern precisely the notion of 'bridge' in social network.

Too often, Granovetter's article is reduced to a vindication of weak connections (superficial acquaintance) over strong relations of friendship, kinship or professional cooperation. Reminding sociologists that all relations deserve consideration regardless of their force is certainly an important merit of Granovetter's paper, but not the only one. What is often lost to Granovetter's readers is that, far from being the focus of the article, the strength of social ties is in fact a proxy for something different and far more interesting. So little is the importance that Granovetter attaches to the strength of relations, that he discards the whole question with a few nonchalant lines:

> Most intuitive notions of the "strength" of an interpersonal tie should be satisfied by the following definition: the strength of a tie is a (probably linear) combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize the tie. Each of these is somewhat independent of the other, though the set is obviously highly intracorrelated. Discussion of operational measures of and weights attaching to each of the four elements is postponed to future empirical studies. It is sufficient for the present purpose if most of us can agree, on a rough intuitive basis, whether a given tie is strong, weak, or absent (Granovetter, 1973, p. 1361).

Perfectly satisfied by such 'rough intuitive basis', Granovetter goes on discussing what is really important for his argument: that there might be a fundamental *functional* difference between strong and weak ties.

> Consider, now, any two arbitrarily selected individuals-call them A and B-and the set, S = C, D, E, ... , of all persons with ties to either or both of them. The hypothesis which enables us to relate dyadic ties to larger structures is: the stronger the tie between A and B, the larger the proportion of individuals in S to whom they will both be tied, that is, connected by a weak or strong tie. This overlap in their friendship circles is predicted to be least when their tie is absent, most when it is strong, and intermediate when it is weak. (ibidem, p. 1362)

The strength of social relations, in other word, is not important *per se* but for the *structural function* that different relations may play in relation to the larger network. With an insightful twist, Granovetter realizes that the strength of a relation is a good indicator of another, more important, property: the overlapping of the social circles of the individuals connected by the relation. In the jargon of network analysis, such property is called 'structural equivalence' and (for undirected and non-weighted networks) is defined precisely as in Granovetter's article: "the proportion of individuals in S to whom they [two given individuals] will both be tied".

Granovetter's argument is crystal clear: if A and B are strongly related they probably spend much time together and this, in turn, increases the probability that their acquaintances also come to know each other ("For example, if A and B are together 60% of the time, and A and C 40%, then C, A, and B would be together 24% of the time", ibidem, p. 1362). As a conclusion, the stronger is the relations between A and B, the more structurally equivalent A and B will probably be.

The consequences of such idea are easy to appreciate both for single individuals and for the global network. Whereas strong ties serve to keep us close to people that share our same social milieu, weak ties connect us with farther region of the social space. While strong ties promote homogeneous and isolated communities, weak ties foster heterogeneity and crossbreeding. Or, to use the old tönnesian cliché, strong ties generate *Gemeinshaft*, while weak ties generates *Geselleshaft* (Coser, 1975).

What really counts, therefore, is not if a relation is strong or weak, but weather it connects (structurally) similar or dissimilar individuals. In following article (1983), Granovetter express this point explicitly:

> I have not argued that all weak ties serve the functions described in SWT-only those acting as bridges between network segments. Weak ties are asserted to be important because their likelihood of being bridges is greater than (and that of strong ties less than) would be expected from their numbers alone. This does not preclude the possibility that most weak ties have no such function. It follows that an important part of further specifying the argument would be more systematic investigation of the origin and development of those ties which bridge as compared to those which do not. (p. 229)

Drawing on Granovetter's idea, authors working on social capital (Gittell & Vidal, 1998; Putnam, 2000) proposed to call *bonding* the ties that "brings closer together people who already know each other" and *bridging* the ties that "brings together people or groups who previously did not know each other" (Gittell & Vidal, 1998, p. 15). The same authors observed that the prevalence of one type of relations or the other has distinct effects at both micro and macro level. At a micro level, individuals possessing larger bridging capital can transmit and access to information more easily, are more comfortable in diverse social settings and cumulate more differentiated roles (according to Boltansky, 1973 they constitute the dominant class). Individual having larger bonding capital, on the contrary, tend to be more influential within a specific group and more strongly affected by such affiliation. At macro level, networks characterized by bounding relations tend to display separate and closely connected groups, whereas networks characterized by bridging relations are generally loosely and homogeneously connected.

For practical and theoretical purposes, the bridging/bounding couple seems to represent a crucial distinction in social networks, one that (as Granovetter boldly affirms) could open the way to the Grail of sociology: the micro-macro link

> I will argue, in this paper, that the analysis of processes in interpersonal networks provides the most fruitful micro-macro bridge. In one way or another, it is through these networks that small-scale interaction becomes translated into large-scale patterns, and that these, in turn, feed back into small groups (1973, p. 1360).

And yet, surprisingly enough, when it comes to provide an operational distinction between different types of relations, Granovetter does not draw on network analysis. Presenting his case study on occupation search, he admits that

> I asked those who found a new job through contacts how often they saw the contact around the time that he passed on job information to them. I will use this as a measure of tie strength (ibidem, p. 1371).

Although Granovetter does realize that *bridging* is the phenomenon he is looking after, two major difficulties prevented him from a direct operationalization of such concept:

> We have had neither *the theory* nor *the measurement and sampling techniques* to move sociometry from the usual small-group level to that of larger structures (ibidem, p. 1360, emphasis added).

Let's start from "the measurement and sampling techniques". In order to compute the bridging force of a given connection, one need, in the first place, to be able to draw the *complete* graph of the social network he/she is investigating. Networks constructed with traditional ego-centered and snowballing techniques are too biased to compute bridging forces. Complete graphs of small social groups will not work either, since such groups are, by definition, dominated by bounding relations. Since the essence of bridges is to connect individuals across distant social regions, they can only be computed in large and complete social graphs. Not necessarily the graph of whole world society (though this would be preferable), but at least one that includes several thousands individuals and the few millions of ties connecting them.

Hopeless until a few years ago, such endeavor seems more and more reasonable as digital media spread through society. Thanks to digital traceability it is now possible to draw large and even huge social networks (Venturini and Latour, 2011). Think for example of what Granovetter could have done if he had the database of Facebook™ or of LinkedIn™ at his disposal, or if he could reconstruct the graph of the mail exchanged through GMail™. And even without these databases (which exist but are still difficult to access), think at the large networks of scientific publications that can be easily extracted by ISI Web of Science™ other bibliographic archives, or to the huge networks of hyperlinks delivered by the simplest webcrawler (the *Issuecrawler.net* for example).

<MATHEMATICIANS MAY START HERE>

As retrieving vast social networks becomes less and less of a problem, the second difficulty identified by Granovetter remains undiminished: we still miss "the theory" necessary to work out a mathematical/topological definition of the bridging force. Note the interest of such a definition is not limited to social networks. Being able to identify bounding and bridging connections has a clear interest for any type of network. In social networks, bounding/bridging measures tell us which relations build social territories and which connections allow items (ideas, pieces of information, opinions, money…) to travel through them. In scientometrics networks, this measures tell us which references define disciplines and paradigms and which breed interdisciplinarity. In ecological networks, they tell us which relations create specific ecological communities and which ones connect them to larger habitats. In web networks, they can tell us which hyperlinks define online communities and which serve as gateways between different communities.

In all this contexts (and in many others), it is the very same question that we wish to ask to our connections: do they reinforce the density of a cluster of nodes (bounding) or do they connect two separated clusters (bridging)? Formulated in this way, the bridging/bounding question seems easy to answer. After having identified the clusters of a network, all you would have to do is to observe if the two ends of by an arc belong to the same cluster (bounding) or to two different clusters (bridging). The only problem would be selecting a clustering method (modularity? cliques percolation? minimum cut? plain observation of the density in the spatialized graph?) and then bounding and bridging would compute with no effort.

This approach however has two major disadvantages. First of all, the intra-cluster/inter-cluster approach can only provide a binary distinction unfit to measure subtle variations in the bridging force. For example, it would be impossible, using such approach, to distinguish an arc connecting two adjacent clusters from one bridging two clusters far apart in the network space. Or, to say it with other words, to distinguish the situation where several different bridges connects two clusters (thereby drawing them closer in a spatialized graph) from the situation where only a few connections allow to go from one cluster to the other. Analogously, defining bounding just as an intra-cluster arc prevent from distinguishing the bounding force of arcs within denser or looser clusters.

Its binary nature, however, is not the main drawback of this approach. The biggest problem with the intra-cluster/inter-cluster approach is that, by definition, it can only work if clusters are well defined. Regardless of the clustering method, it is well known that there are graphs that clusterized well and other that do not. When clusters are well defined and clearly separated, it is easy to identify which arcs lie within a single clusters and which arc span across two different clusters. Things, however, gets more complicated when clusters are ill-defined, when they tangle and overlap, when graphs can be cut in many alternative and equally acceptable ways. In these cases, as in the vast majority of real networks, the intra-cluster/inter-cluster approach gives poor results because the clustering itself is not reliable.

The intra-cluster/inter-cluster approach is deeply flawed by its inherent circular logic: it uses cluster to define bridging and bounding ties when it is precisely the balance of bridges and bounds that determines clusterization. Please remark that, far from being a mathematical subtlety, this question is a key problem in social theory. Defining internal (*gemeinshaft*) and external (*geselleshaft*) relations by presupposing the existence and the composition of social groups is absurd as groups are themselves defined by social relations. Maybe Granovetter was right when he claimed that the investigation of bounding and bridging networks could cast new light on the micro-macro link, but not in the way he believed If the interactions determine the structures and the structures determine the interactions, why should we assume that they exist on two different levels, micro and macro? This question is too vast to be treated here, but see Latour et al. (2012) for a longer discussion.

For the moment, let's get back to our networks and let's see if we can find out a way to measure the bridging force that do not draw on clustering. To do so, we will introduce a very special network that will help us illustrate our ideas.

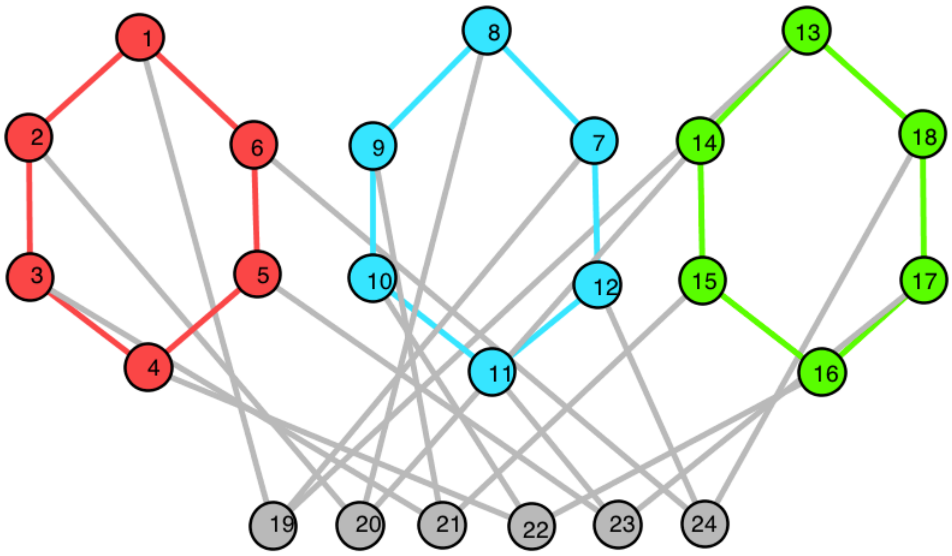

Fig. 1. The bounding-bridging network and how it is constructed.

The bounding-bridging network is constituted by three rings of six nodes each and six more nodes each one connected to one node from each ring. Because of the way it is constructed, it is particularly difficult to extract clusters from the bounding-bridging network. This can be easily showed by spatializing the graph with a force-vector algorithm. No matter the algorithm used, the graph always presents a homogeneous density that makes it impossible to distinguish different clusters.

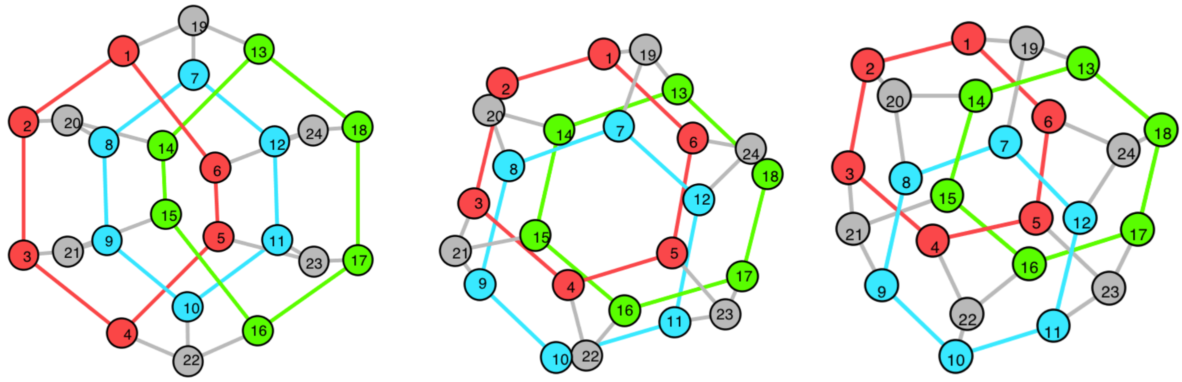

Fig. 2. The bounding-bridging network spatialized in Gephi using (a) ForceAtlas2 (LinLog mode), (b) ForceAtlas, (c) Yifan Hu. The colors are the same as in fig.1.

Analogously, mathematical method to detect clusters fail dramatically on the bounding-bridging network. This is not surprising, as Andreas Noack (2009) demonstrated that visual and mathematical clustering are in fact equivalent, but the results are nevertheless striking. In the next figure, we computed the modularity classes of the graph by running five times the same modularity algorithm. Colors in the image attributed to nodes according to their modularity class.

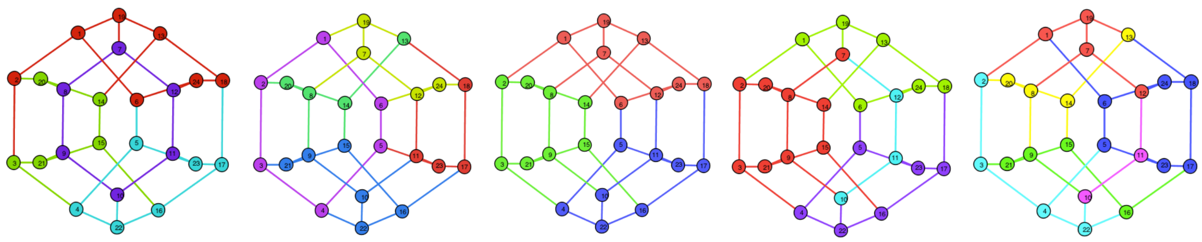

Fig. 3. Clusters computed by 5 subsequent iteration of the same modularity algorithm

Although it is well know that the modularity algorithm is not deterministic, the wide differences in the number and in the composition of the modules proves that modularity cannot be trusted to clusterize this network.

The bounding-bridging network simply does not allow clustering. As a consequence, it is impossible to use the intra-cluster/inter-cluster approach to distinguish the bounding and bridging arcs of this network. Does this means that there are no bounds or bridge in the graph? Intuitively, it is difficult to resign to such an idea, especially by looking at fig.1. and considering how the graph was built. Isn't it patent that the colored arcs are bounds (as the connects nodes within the same circle) and the grey arcs are bridge (as they reach beyond the circles)? Intuition, however, can be mistaken, so let's look for sounder evidence that there is a fundamental structural difference between colored and grey arcs. What we are looking for is a measure that could be computed for each arc and that would differentiate colored and grey arcs of the bounding-bridging network.

Strangely enough, up until today, network theory has devoted little interest to measuring arc properties. The proof is that no one of such measure is available in the standard Gephi package. This is all the more surprising, as many different measures are proposed for nodes instead: degree; pagerank; authority/hub; clustering coefficient; betweeness centrality; closeness centrality; eigenvector centrality; eccentricity (just to name the ones available in Gephi). Unfortunately these measures can be used to discriminate bridges and bounds as arc that would connect nodes having different values for one of these measures. The reason is simple: because of the way it has been constructed, in the bounding-bridging network all nodes have the exact same values all these measures (with two remarkable exception that we will discuss later). Since all nodes are connected to the same number of arcs and since no cliques are present in the graph, all nodes have the same degree (= 3), the same pagerank (0.042), the same authority/hub value (0.042), the same clustering coefficient (0), the same eigenvector centrality (1) and the same eccentricity (5).

Only two measure give different results for the nodes of the bounding-bridging network: betweeness centrality (2.87 for nodes within the 3 circles, 2.783 for nodes beyond the 3 circles) and closeness centrality (22.333 for nodes within the 3 circles, 18 for nodes beyond the 3 circles). These two measures of centrality are commonly used to identify nodes that plays some sort of bridging function, according to the intuitive idea that nodes that is more central should connect different regions of the graph. Such idea proves to be remarkably false in this network as nodes within circles connected to 2 colored (bounding) arcs and 1 grey (bridging) arc appears to be more central than nodes outside circles connected to 3 grey (bridging) arcs.

It seems therefore that we cannot trust any of the standard network measure to compute bridging force in our testbed graph. So resistant is the bounding-bridging network to our efforts to distinguish its bridges and bounds that we are tempted to conclude that the example is ill chosen. Perhaps, we should simply resign to the idea that there are no bridges or bounds in poorly clusterized graphs. Perhaps, the intra-cluster/inter-cluster approach is indeed the only possible. Perhaps, it is the way we constructed our testbed graph that tricked us into believing that is was made of bridges and bounds, when in fact nothing can tell apart grey and colored arcs except for the colors that we imposes to them. Or is there?

Looking closer at our graph, we believe we can identify at least one difference between grey and colored arcs. Let's state it this way: a spider moving randomly through the bounding-bridging network has higher changes to return to the same bound then to the same bridge. Such difference derives, of course, from the way the network is built. When crossing a colored arc, the spider arrives on a colored node that is connected to two arcs of the same color and one grey arc. At its next move, the spider has therefore twice as many probabilities to remain within the same ring that to leave it and, while it remains within the ring, it has higher probability to return to the same colored arcs it already crossed. Grey arcs, instead, allow our spider to leave a colored ring and reach a grey node. Grey nodes are connected to three grey arc each one leading to a different ring. At its next move the spiders has therefore twice as many probabilities to arrive to a new ring than to get back to original one. Entering a different a ring, the spider has higher probability to remain in it than to leave (as explained above), thereby reducing its possibilities to return to the original grey arc.

We can therefore define the bridging force of an arc as the inverse of the probability of returning to that arc or, more precisely, *the force of bridging of an arc A is the inverse probability of crossing A twice (or more) in a random S-steps walk starting from crossing A in one direction and in the other*. Please remark that our measure does not use the simple probability of encountering A in a random walk (a sort of inverse Pagerank for arcs), but the probability to cross one arc twice after having crosses it one (in one direction *and* in the other). An arc can be unlikely reached in a random walk and still not being a bridge. In figure 4, for an example, arc L is very difficult to reach for a random spider, but there is no sense in which it can be called a bridge. According to our definition, L is not a bridge because if crossed from left to right (from p to q) is very likely to be crossed again (as in q there are no other option for the spider then to cross back L).

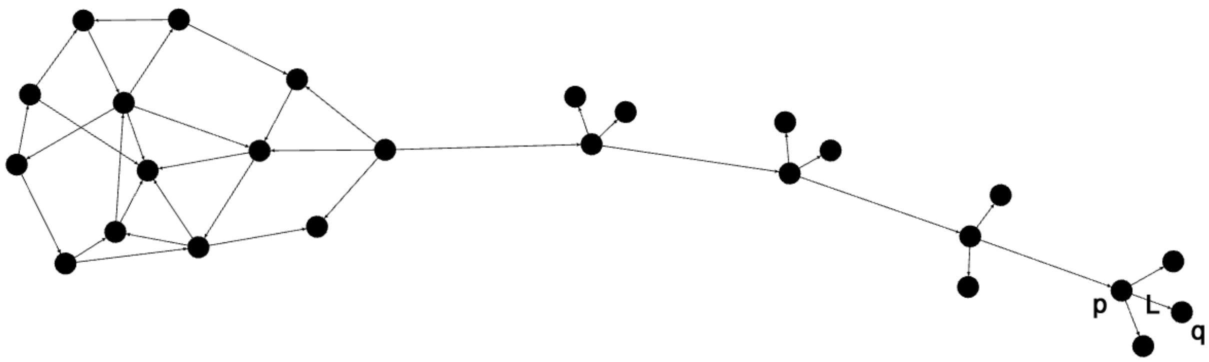

Fig. 4. An arc unlikely to be reached which is not a bridge.

We think that our measure defines effectively what the bridging force of arcs, unfortunately we are fully aware that such measure is impossible to compute on large (and even medium networks). Computing the inverse probability of returning to an arc in a S-steps random path starting from it (in one direction and the other) requires computing all the possible S-steps random paths starting from each arc of the network, which in turn requires huge resources of computation and memory.

What we are looking for, therefore, is a reasonable approximation to our measure…

References

Boltanski, L. (1973). L'espace positionnel : multiplicité des positions institutionnelles et habitus de classe. *Revue française de sociologie*, 14.

Coser, R (1975). "The Complexity of Roles as Seedbed of Individual Autonomy." In L. Coser (Ed.), *The Idea of Social Structure: Essays in Honor of Robert Merton*. New York: Harcourt Brace Jovanovich.

Gittell, R., & Vidal, A. (1998). *Community Organizing: Building Social Capital as a Development Strategy*. Thousand Oaks, CA: Sage.

Granovetter, M. (1973). The Strength of Weak Ties. *American Journal of Sociology*, *78*(6), 1360-1380.

Granovetter, M. (1983). The Strength of Weak Ties: A Network Theory Revisited. *Sociological Theory*, *1*, 201.

Latour, B., Jensen, P., Venturini, T., Grauwin, S., & Boullier, D. (n.d.). “The Whole is Always Smaller Than Its Parts” —How Digital Navigation May Modify Social Theory. *British Journal of Sociology*

Noack, A. (2009). Modularity clustering is force-directed layout. *Physical Review*, 79(2).

Putnam, R. (2000). *Bowling alone : the collapse and revival of American community*. New York: Simon & Schuster.

Venturini, T., & Latour, B. (2010). The Social Fabric: Digital Traces and Quali-quantitative Methods. *Proceedings of Future En Seine 2009*.