

Searching for intermediaries An iterative structural deconstruction of personal networks

Claire Bidart,^{*} Alain Degenne[†], Michel Grossetti[‡]

October 6, 2020

Abstract

Abstract: Sociologists attempt to compare personal networks in order to identify explanatory factors that may indicate differences between them, above and beyond individual life histories. However, the usual indicators distinguish a single structural characteristic. In this article we attempt to find a more satisfactory way of explaining the structure of personal networks. We propose an iterative deconstruction of personal networks focused on betweenness centrality which is recalculated by an algorithm at each stage of the process. When applied to a sample of 287 cases gathered by direct interviews, it allows us to establish the betweennesses that are otherwise concealed by higher values, to determine the intermediaries at each stage, to record their exit rank, to characterise all the alters by several criteria and, finally, to take into consideration the number of stages required to deconstruct the network, thereby creating a new indicator of network stratification. These results are coherent with what we know about the evolutions of personal network structures and social roles along the life cycle. In a way, they reflect the imprints that time and life events have left on Ego's network.

Keywords: personal networks; structural analysis; methodology; indicators; betweenness centrality;

^{*}Aix Marseille Univ, CNRS, LEST, Aix en Provence, France. (claire.bidart@univ-amu.fr)

[†]CNRS, Tours. (anicalain.degenne@numericable.fr)

[‡]CNRS, EHESS, Université de Toulouse UT2J, LISST-Cers. (michel.grossetti@univ-tlse2.fr)

1 Introduction

For social scientists, personal networks are the systems of relationships or contacts maintained by individuals, who have built them up over the course of their lives and their exchanges in various contexts. Thus networks are formed little by little, through meetings in groups or one-to-one, ties added or lost, connections and disconnections. Personal networks bear the imprint of individuals' commitments and affiliations in various circles and with different people, some of whom remain in contact even after the original context disappears (school friends, for example). Personal networks are documented either by direct survey or by observation of the records left by interactions, as in the case of digital networks. The person whose relationships are being listed is generally denoted by the term 'Ego', while the term 'alter' is used to denote all the individuals to whom Ego is connected. When information on the relationships between the alters is available, the structure formed by those relationships can be analysed, as is done in so-called complete network analyses (in which each alter is questioned in their turn). Thus Ego's current network includes alters from various periods, whose ties might vary considerably in duration and degree of connection with each other. Thus the structure of a personal network gives a picture of the connections forged by alters between different contexts (family, university, job, leisure activities, homes etc.). Sociologists interested in personal networks seek to understand how individuals' histories have shaped their networks and, conversely, how their networks have influenced their decisions and life trajectories. They

also attempt to compare those networks in order to identify the factors that might help to explain the differences between them, above and beyond individual histories. In order to compare these networks, the information has to be reduced and systematised. This is done by means of indicators, which are measures showing how a particular characteristic varies. Thus there are measures of size, network composition as assessed by such and such a criterion (men and women, social classes, family or non-family etc.) as well as more structural measures such as network density, modularity, diameter, etc. The indicators that have been developed, debated and stabilised as network analysis has evolved are very effective, as is evidenced by their ability to cut across academic disciplines, types of data, the research problem under investigation, etc. Nevertheless, the difficulty lies in the need to provide as complete a picture as possible of a network's structure while at the same time simplifying it sufficiently to make it comparable with others. The indicators, after all, provide only one measure at a time and any comparison that is made relates to only a single characteristic of the structure. Now two networks of equal density, for example, can be very different when assessed by other characteristics. Attempts can also be made to represent the overall structure of a network visually, or to summarise it with a particular combination of indicators, as we did in constructing a typology inductively (Bidart, Degenne & Grossetti, 2018). On the basis of this typology, six overall structures can be identified that are both contrasting and make sense sociologically. They are: regular dense, centred dense, centred star, pearl collar, segmented and dispersed networks. The typology can be used to carry out comparisons. However, it lacks a little finesse as soon as analysts wish to investigate this overall shape more closely and to make more subtle structural distinctions, and its robustness for others types of corpus remains to be demonstrated. Thus we have attempted to find a more satisfactory way of explaining the overall structure of personal networks while at the same time avoiding excessive reductiveness or dispersion of the indicators. What we are proposing here is an iterative deconstruction of personal networks. The idea is to proceed in successive sequences that little by little decon-

struct the network while retaining the records of the deconstruction, in order both to measure its degree of stratification, to describe the alters who constitute it and to identify explanatory factors which make sense sociologically. Thus it is a processual approach that records the succession of sequences as a part of the result. In this process, we opted to use the notion of betweenness centrality. This is an indicator that is both based on the overall structure and describes nodes. It also enables sociologists, who generally have a great deal of information on nodes, to understand what underlies certain individuals' betweenness. To begin, we explain our choice of this indicator by discussing studies that present the comparative advantages of the various centralities. We then describe the data we used to construct this experimental method, giving a few examples of the graphs produced by it. In the next section, the original algorithm developed to increment the deconstruction process is presented. A number of cases are then presented in order to bring the structural method into dialogue with the sociological data. The descriptions of the alters that emerge from the deconstruction process are examined and then combined with sociological variables whose effects on personal networks are well known (type of tie, age, marital status, social class). We then return to the level of Ego in order to propose as a new indicator of complexity the number of stages required to deconstruct Ego's network, i.e. its degree of stratification.

2 Betweenness

2.1 Different centralities and their meanings

The notion of betweenness plays an important role in network analysis because it is a means of understanding how dense parts of a network are connected to each other. Betweenness centrality is the indicator that serves to identify what keeps a network together, but in a differentiated way (rather than uniformly, as density does). It is what connects the components to each other, creating bridges between parts that would otherwise be isolated. If there are many paths,

then there are many alternative ways of going from one point to another, as in the case of a very dense network; the betweenness of the various nodes is low, there being a high level of overlap and redundancy. If the network is less dense, there are holes within it and there are few nodes whose edges span them, which gives them significant betweenness. Different forms of centrality have been defined. The most frequently cited article in this regard is that by Linton Freeman (1978), which seeks to clarify these notions. He defines three major categories of centrality indexes: degree centrality, closeness centrality and betweenness centrality. Degree centrality measures each alter's popularity by the number of connections he or she has. However, these connections may overlap with the neighbouring ones and everybody may be very well connected. The node is described but not compared with the others. Closeness proximity looks for the shortest path and therefore shows the alters who are closest to all the others. In social networks, this may indicate individuals who tend to seek consensus. Betweenness centrality, finally, is proportional to the scarcity of the individuals who connect one zone of the network to the next. Individuals with a high degree of betweenness centrality can be described as intermediaries.

2.2 The intermediaries

The article by Gould and Fernandez (1989) is generally regarded as the pioneering work in this area. These authors emphasise the role of intermediaries between groups within a network. They identify five types of 'brokers': coordinator, itinerant broker, gate keeper, representative and liaison. In this way, they seek to describe and measure the importance of the various brokerage roles. This is an important approach in organisational analysis and more generally in social exchange theory. Ronald Burt (1992, 2007) highlights actors' capacity to play a personal role in a network and to control their partners' actions. His concept of the structural hole is an attempt to explain the origin of differences in individuals' social capital. In their review of this topic, Everett and Valente (2016) write: 'In all of these papers, the underlying assumption is that actors control resources which are

flowing through the ties that they are incident to.' Here too, betweenness is of greater value than degree centrality. In the chapter they devote to the intermediary's role, de Nooy, Mvar and Batagelj pose the problem in the same terms. For all these authors, an intermediary is a sort of strategist who benefits from his or her control over the betweenness points. However, context effects may mean that an intermediary is a 'freelancer' across different contexts, or a person whose tie with Ego is a multiplex one. He or she may also be the sole 'survivor' of a group linked to another group or to an alter who is sufficient to make the difference. Thus the temporal dimension can be crucial to this positioning. In our case, in which the focus is on personal networks, the hypothesis that an intermediary's status is a privileged one has to be handled with caution, since we are not dealing with networks of individuals connected by a particular type of activity or an organisation but with a relational environment constructed around an individual who generally turns out to be much less strategic and rational. Interpretation of the role of intermediaries is one of the issues at stake in this analysis.

2.3 The measure of betweenness in networks

Betweenness centrality is a particularly interesting indicator since it applies to the nodes while being derived from the structure of a network. It isolates an interesting social property, namely that of acting as intermediary between social groups that would otherwise be isolated. Freeman (1977) attributes the origin of the concept to Bavelas, in an article published in 1948: 'This idea of point centrality was introduced by Bavelas (1948) in his first paper on the subject. He suggested that when a particular person in a group is strategically located on the shortest communication path connecting pairs of others, that person is in a central position. Other members of the network were assumed to be "responsive" to persons in such central positions who could influence the group by "withholding information (or) colouring or distorting it in transmission".' He also cites Anthonisse (1971) as a forerunner. Freeman defines the index of betweenness centrality in the same way, with reference

to the shortest paths between two nodes. Let k be a node on graph G and let i and j be any two nodes on the same graph; let $c_{i,j}$ be the number of geodesics, that is the minimum paths between i and j . Let $c_{i,j,k}$ be the number of these geodesics that pass through k . The betweenness of k is defined by $b_k = \sum_{i,j} (c_{i,j,k}/c_{i,j})$ and $j \neq k$. Since b_k is the sum of relationships calculated for all the pairs i,j , it may be large and in any case not lie between 0 and 1, as is generally desired. Consequently, the betweenness coefficient of a node is used in the form weighted by the number of pairs i,j , where i and j are different from k . $\beta_k = b_k/(n-1)(n-2)$ β_k attains its maximum value of 1 in the case of a star graph and its minimum of 0 in the case of a clique. The advantages of this definition include the fact that it can be applied to disconnected graphs. The betweenness centrality of edges is defined in the same way: Let $e_{i,j,kl}$ be the number of these geodesics that pass through an edge $(k-l)$. The betweenness of $(k-l)$ is defined by $d_{k-l} = \sum_{i,j} (e_{i,j,kl}/e_{i,j})$ $i \neq j \neq k,l$ and the weighted betweenness is defined by $\delta_{kl} = d_{kl}/(n-1)(n-2)$. Other definitions of betweenness centrality have been proposed, whether the aim is to deal with the case of oriented or weighted graphs or to consider not only the geodesics but also all other possible paths in order to take account of the flows within a network.³

2.4 Betweenness in personal networks

Calculating the betweenness centrality of the nodes in this type of personal network does not pose any particular problem, since the standard formula applies. However, there is one particular node whose betweenness centrality is not, on the face of it, calculated, and that is Ego. True, it is high and even maximum, since Ego is linked to all the other nodes and is consequently their closest intermediary; for all that, however, it is not generally equal to 1. Everything depends on the structure of the network and of the ties between the alters. If they are closely connected and the network is dense, Ego's betweenness centrality is affected⁴. This is what Burt (1992) recognised and made much use of in developing his structural holes theory⁵. It finds expression in his network constraint index. In what follows, we will

not analyse Ego's betweenness centrality, which is very strongly negatively correlated with density but cannot be reduced to it. We will focus on the intermediaries among the alters. This article draws on a sample of 287 cases gathered by direct interviews in order to explore the structure of personal networks. The networks in question are small in size (around forty people on average) compared with digital networks but relatively large compared with sociological surveys, which are often centred around 'core networks' or 'support networks'. This will enable us to test a segmentation algorithm that might cause problems if used to analyse very large networks.

3 The data

3.1 The panel survey

We draw on a qualitative longitudinal panel survey of personal networks of young French people, who were followed up as they were entering adulthood. Four waves of interviews were conducted at intervals of three years⁶. In 1995, 87 young people living in the Caen urban area (Normandy, France) were interviewed. One third were in the final year of high school studying for the economic and social baccalauréat, one third in the final year of preparation for the vocational baccalauréat and one third were taking part in a labour market integration programme. Each group contained equal numbers of young men and women. They were aged between 17 and 24 and a large share was drawn from the lower social classes. In the second wave in 1998, 74 young people responded again, 66 responded in 2001, and 60 in 2004. We thus obtained a total of 287 personal networks. In each wave of the survey, the process of constructing each of the personal networks was carried out by means of an original context-focused name generator (Bidart, Charbonneau, 2011). A long series of questions on the various life contexts had to extend as widely as possible across the contexts in which young people live, as well as across previous contexts in which the respondents had known people with whom they were still in contact. In all, 53 contexts in which individuals could potentially be active were systemati-

cally cited in this way. For each context, respondents were asked: "In (this context), have you met people whom you know a little better, with whom you speak a little more?". Two screening questions per context were added: (1) "Do you see any of them outside of (the context)"? (2) "Are any of them important to you?". The individuals mentioned in response to one of these last two questions were regarded as strong ties, the others as weak ties. The socio-graphic characteristics of all these alters were recorded. In the case of the strong ties, the interconnections between them were listed systematically (from Ego's point of view), making it possible to outline the structure of the network. The average size of the resulting networks was 38 alters, but with considerable variation: the smallest network had 6 alters, the largest 134. We showed that social and biographical characteristics influence the size, composition and structure of networks (Bidart, Degenne & Grossetti, 2011).

3.2 Contrasting forms of betweenness

Taking as our corpus the 287 networks generated over time by our survey, we now seek to describe precisely the structures of these networks. The links between Ego and the alters are concealed in our graphs, since they do not provide any additional information; Ego is linked to all the alters by construction and taking those links into account would simply overburden the network for no good end. Nevertheless, in the light of our efforts to interpret the role of intermediaries, it should not be forgotten that Ego is at the centre of the network. The alters here are identified by numbers.

In these personal networks, betweenness centrality can take various positions:

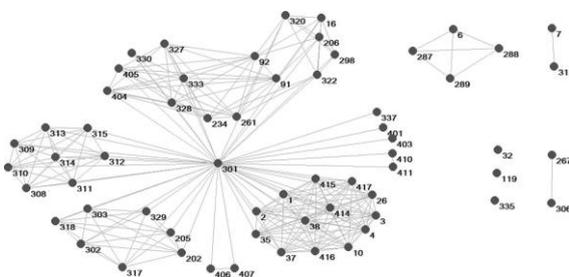


Fig.1. Agnès's network7 in 2004

Agnès is one of the individuals in the panel8; in her network, alter 301 has the highest betweenness centrality. He is the only person who connects various components, including Ego's family, a group of friends in Paris where Agnès lives, a component made up of two partially linked cliques (former neighbours and friends from her home town in Normandy with whom she shared leisure time), a group she met during a trip to Tunisia, as well as three couples and an alter who would otherwise be isolated. Alter 301 is Agnès's husband; he is the only node who, like Agnès herself, connects all these components. On the other hand, her husband is not connected to a small group of 4 high-school friends, two individuals from a former job, two childhood friends and two ex-partners. These alters stem from other parts of Ego's life whom she has no need (or desire) to see with her husband.

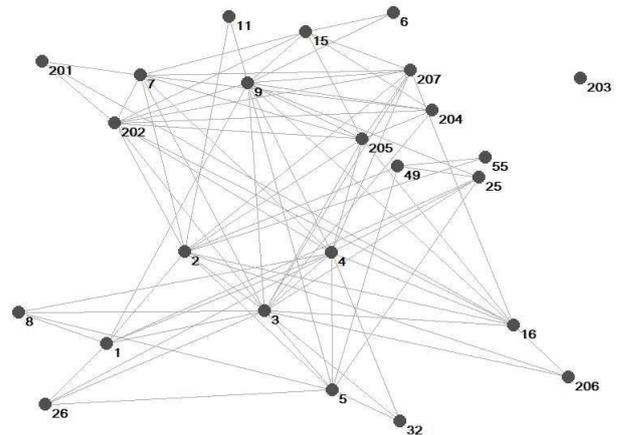


Fig.2. Florence's network in 1998

In this other example, Florence's network, there is no alter who occupies such a central position; rather, several nodes share a moderate level of betweenness that is much less visible in such a denser network. They are her elder sister (3) with whom she often goes out, her father Patrick (5), her childhood friend Meriem (9) and her partner Thomas (2). In this rather dense network, family and friends are partly mingled. Each of these individuals connects a part of the network but without being the sole link.

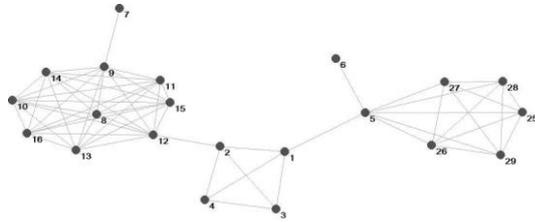


Fig.3. Marc's network in 1995

Marc's network rather resembles a pearl collar. Four different alters have the highest betweenness. They are several intermediaries, each one connecting a different group that would otherwise be isolated. His mother Dominique (5) knows Stéphane (1) with whom Marc went to school. Stéphane introduced some friends to him, including David (2) who, in turn, introduced him to Karl (12), a member of a group of bikers to which some neighbours also belong. Thus Marc is on the periphery of various groups, each of whom accepts him through a different intermediary. These examples shed light on some of the ways in which betweenness is distributed in personal networks.

3.3 Networks with zero betweenness

It is important for what follows to bear in mind what networks are in which all the nodes have zero betweenness. The components of such networks are isolated individuals, isolated dyads or isolated cliques. The betweenness centrality of the nodes in a network can be calculated, whether or not the network is connected. It can be measured for each node in the connected part to which it belongs. Thus all the nodes in a network that is made up of a set of cliques and isolated individuals have zero betweenness centrality, whether their degree centrality is also zero, as it is for isolated individuals, or high, as it is for cliques. This can happen in personal networks, as it does in Simon's case:

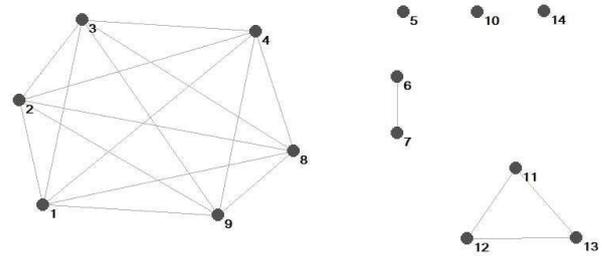


Fig.4. Simon's network in 1995

Here, the betweenness centrality is zero everywhere. Simon's network was, after all, made up of a clique of high-school friends, a family-based triad (his mother, his mother's husband and his brother), a dyad of work colleagues (he repairs tyres) and three isolated individuals: his paternal aunt, who is the only representative of his deceased father's family, a barman and a female classmate living in Israel. None of these different milieus has any contact with the others and nobody acts as intermediary.

These zero betweenness networks are interesting both structurally and sociologically, evincing as they do the absence of a go-between linking universes of varied sizes that remain disconnected from each other. This constitutes the endpoint criterion for the algorithm developed below. The deconstruction of the network is complete when all that remains is final components of zero betweenness.

4 Network deconstruction algorithm

4.1 The standard applications

We are taking as our basis here a network deconstruction algorithm. This is not a very common practice. Nevertheless, several types of applications can be cited. Some proceed by the removal of edges, others - like the one we have adopted - by the removal of nodes. The search for communities is one of the common applications for the indicator of betweenness centrality applied to edges. If an edge is a necessary path between two groups of nodes, it is going to have a high betweenness coefficient; this is the key to the

algorithm developed by Newman (2003) and Newman and Girvan (2003) in order to discover community structure in networks⁹. This algorithm is very interesting, but calculating the betweenness centrality of edges for the purpose of analysing large graphs is a very long-winded process. Consequently, it is often replaced by Louvain's algorithm, which functions by aggregation in order to reveal communities. Some studies have sought rather to reveal the 'skeleton' of a graph. In a 1962 article, Claude Flament (1962) proposed a technique for pruning a graph that he called 'similarity analysis' and which amounts to removing from any triangle the edge that has the highest degree of betweenness. This produces a network that no longer has any cycles of length 3. The algorithm can then be continued by removing the edge that has the highest betweenness in the cycles of length 4, then length 5 and so on. In this way, the network's minimum spanning tree is obtained, from the point of view of the betweenness of the edges (Rosenstiehl, 1967). Although Flament did not formally define betweenness centrality at this time, he had certainly intuited it. Other more recent works operate by counting typical motifs (Milo et al., 2002; Cunningham & al., 2012). The proportion of each type of motifs, most often a triad census (with 0, 1, 2 or 3 edges, directed or not, with various types of nodes...) provides a globalized indicator on the network, without however allowing to synthesize its overall structure which can be very different according to the cases. Security problems in networks may give rise to a desire either to protect the central nodes against attack, as in the case of the Internet or electrical or computer networks, or conversely to remove them in order to disconnect the network and break off communications, as when efforts are being made to combat the spread of infectious disease or control the contents of communications on the Internet. In chemistry, one may want to split complex molecules. In all these cases, it is interesting to know how to deconstruct a network in the most effective way. Five strategies can be found in the articles dealing with these questions¹⁰: RAA: Random attack: the nodes are removed randomly, one by one in succession; IDA: Initial degree: the nodes are removed one by one in the order of their initial degree centrality; IBA: Ini-

tial betweenness: the nodes are removed one by one in the order of their initial betweenness centrality; RDA: Recalculated degree: the nodes with maximum degree centrality are removed, with the degrees being recalculated after each removal; RBA: Recalculated betweenness: the nodes with maximum betweenness centrality are removed, with the betweenness centrality of the remaining nodes being recalculated after each removal. According to simulation studies and studies dealing with actual networks, the RB strategy seems to be the most effective way of undermining the functioning of a network.

4.2 Our choice of algorithm

We examined these various strategies (except for random attack, which cannot be reproduced). We also tested the recalculated closeness centrality algorithm.

- Initial Degree (ID): Except for the isolated points on the initial graph, all the nodes have non-zero degree centrality. Consequently, a threshold has to be chosen in order to remove some of them. This choice is either arbitrary or reasoned and in all cases it remains relative. Even though certain thresholds may empirically produce some interesting results, this arbitrariness or dependence on the type of network limits the value of this strategy. Moreover, as mentioned above, degree centrality is focused on the node alone and does not differentiate on the basis of neighbourhood.
- Initial betweenness: The most reasonable choice is to remove the nodes with maximum betweenness. If the network is formed from the outset from components connected to each other by a few intermediaries, the result will be close to that obtained using the recalculated betweenness method (and in the particular case of a network made up of components entirely unconnected to each other, it will actually be identical). However, in certain cases in which the nodes' betweenness centrality varies little, this may lead to a large number of nodes being removed indiscriminately. The threshold question also remains unresolved. There are links between closeness centrality and betweenness centrality (Brandes, Borgatti & Freeman, 2016). We will not dwell on this because it is betweenness that concerns us here and also because using recalculated closeness to analyse a

network does not offer a procedure with an automatic endpoint.

4.3 The recalculated betweenness algorithm (RB)

One technique commonly used in network analyses is based on the use of thresholds. It is technically much easier to put into practice, but we will see from examples that the algorithm we are proposing to use has certain advantages over a method based solely on initial thresholds. An initial illustration of our proposal can be found by taking a graph in the form of a simple chain (fig.5):



Fig.5. A chain of intermediaries

The betweenness centrality indexes are given between brackets. If we select the most central, the first will be 4, followed by 3 and 5. If the most central (4) is removed and the betweenness centralities are recalculated, we have:



Fig. 6. A chain of recalculated intermediaries after the removal of the most central one.

The most central now are 2 and 6, which turn out to be local intermediaries, whereas they were not the highest intermediaries at the overall level. Thus the iterative structural deconstruction approach has the property of revealing local intermediaries, which observation of the betweenness values does not do a priori. This is what we are looking for in the networks in our database. We included this algorithm in Appendix.

4.4 The process at work

Let us decompose the process gradually here, using Marc's network as already shown above. The process can be followed on the graphs below, on which the intermediaries are identified by diamonds and the other nodes by circles. The colours correspond to the stages of the process in which they are identified.

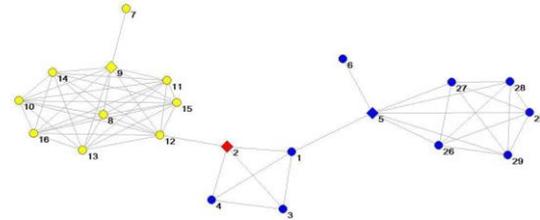


Fig.7. The deconstruction of Marc's network, stage 1 (initial)

The centralities during the initial stage 1 are, in descending order: 2 (0.526), 12 (0.521), 1 (0.479), 5 (0.468) and 9 (0.1). All the others are zero. So we begin by removing alter 2, who is David. He is both the one who was introduced to Marc by his childhood friend Stéphane and the one who introduced Karl to him. Thus he emerges as the most central intermediary between childhood and the most recent group of bikers.

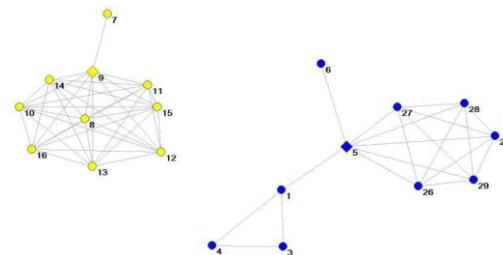


Fig.8. The deconstruction of Marc's network, stage 2

No alter with zero betweenness (final or isolated clique) emerges in this stage. Even though the left-hand component is now detached, it does not (yet) have zero betweenness because of the presence of node 7. So the betweenness centralities are recalculated. In stage 2, the values are as follows: 5 (0.134), 1 (0.082), 9 (0.047). It should be noted that alter 12, who was second during the initial stage, has 'lost' his betweenness, which was dependent on David's, which has been removed. This alter is Karl who, having lost the connection, now finds himself confined to his bikers' clique. Marc's mother, Dominique (5), is now 'promoted' to first place; her betweenness is raised by this removal and she also moves ahead of Stéphane (1). Thus in stage 3 this alter 5, Dominique, who acted as an intermediary between the

family and friends via the oldest of them, Stéphane, is removed. The family clique on the right-hand side becomes isolated.

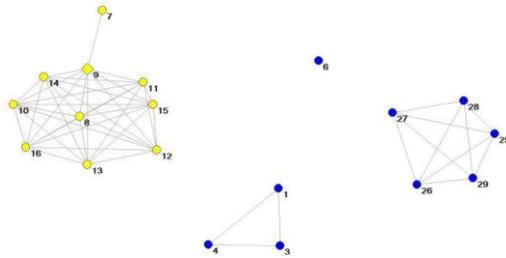


Fig.9. The deconstruction of Marc's network, stage 3

Now, as a result of this operation, node 1 has also lost its betweenness. Stéphane remains in a final clique. Only node 9 now has a non-zero betweenness of 0.052. This is a neighbour who knows Marc's employer (node 7). Here too, he acts as an intermediary between two universes, that of the bikers and the service station where Marc works. This neighbour is removed in order to end up with a 'final' network of cliques and isolated individuals with zero betweenness.

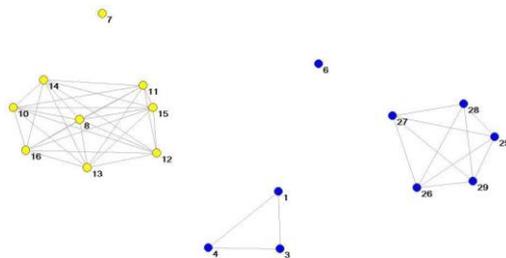


Fig.10. The deconstruction of Marc's network, stage 4

The network is now deconstructed, the process is finished. This sequence of operations can be described in the shape of a tree, with the nodes on the trunk representing the intermediaries and the leaves representing the isolated individuals or final cliques that are disconnected by their removal. To each of these elements can also be added the stage at which it is removed from the network. In the initial stage depicted here only Ego has been removed.

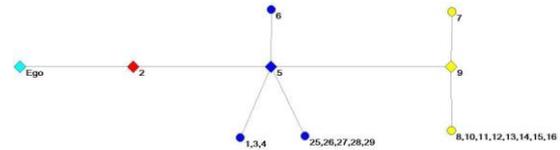


Fig.11. The tree of the deconstruction of Marc's network

4.5 Rapid deconstruction

In 2004, Agnès was living with Olivier and, as we saw above, her network was centred on him, with just a few childhood friends, work colleagues and a former partner not connected to him.

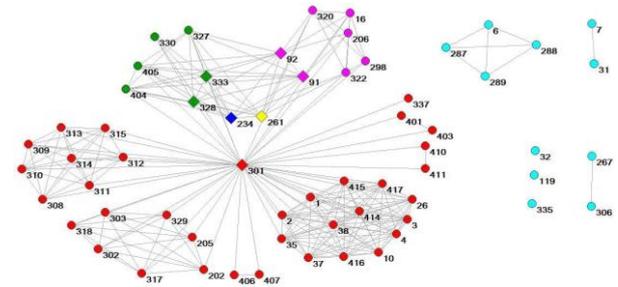


Fig.12. Agnès' network in 2004, with colours indicating the deconstruction stages

When Olivier (301), the most central alter, is removed, the three cliques and the couples at the bottom become detached and have then zero betweenness.

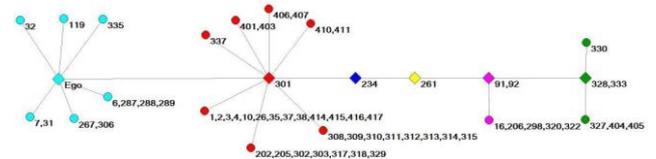


Fig.13. The tree of the deconstruction of Agnès' network

Here, in the initial stage, Ego has been removed along with the isolated individuals, dyads and small cliques whom, as we have seen, Agnès was not keen to connect to her spouse. It can be seen subsequently that Olivier's removal has the greatest impact in terms of disconnected and final elements, of various sizes. The next two nodes (234 and 261) have

high levels of betweenness but they do not disconnect anything since this part of the network is dense and there are other possible paths. Once Olivier is removed, they have high betweenness but do not articulate the various components of the network. They are a pair of friends who, with others, connect two groups of neighbours and friends from Agnès's home town. The blue clique is then removed because it has one more alter than the brown one; the betweenness of its nodes (91 and 92) is slightly higher than that of the brown nodes 328 and 333. These nodes did not have any visible betweenness before, so the process reveals local centralities. This deconstruction is massive in stage 2 and is completed in 5 stages. Thus large networks can be decomposed very rapidly. The largest networks are not necessarily the most stratified and therefore the algorithm being proposed here offers a perspective on personal networks that is not necessarily intuitive.

We will try to identify later which members of the network are intermediaries and which are in the final cliques. We will also consider the link between their status and their characteristics. The number of stages required to decompose the network completely is one of the indicators produced by this method. We use the term 'stratification' to denote this characteristic of the network on the grounds that each stage corresponds to a stratum of the network's structural organisation.

4.6 A slow deconstruction

In 1998, during the second wave of the survey, Florence was 21 and living with a partner. However, her network was not centred on her partner (this is more likely to be the case in the higher social categories, which is where she has her origins).

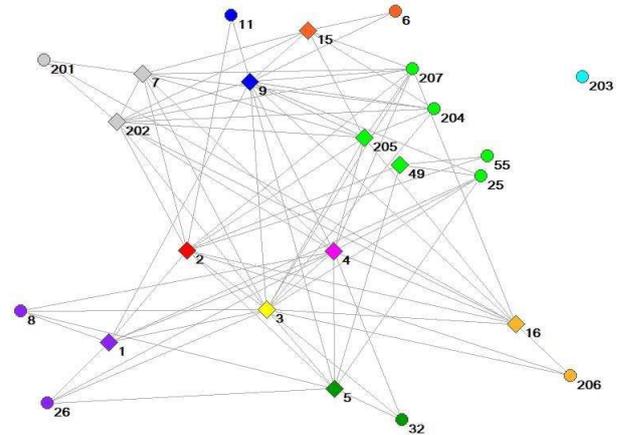


Fig.14. Florence's network in 1998, with colours indicating the deconstruction stages

This is one of the networks with the highest degree of stratification, as measured by the number of stages in our algorithm. This is evident from the graphic below, which comprises 11 stages. During stage 1, one isolated individual appears besides Ego.



Fig.15. The tree of the deconstruction of Florence's network

During the second stage, as well as during the 4th and 5th stages, intermediaries are identified but no components are disconnected as a result. In each of the other stages, only isolated individuals are concerned, and the final stage disconnects four isolated individuals. Thus this is a process that inches its way along slowly, with very dispersed intermediaries and, surprising though it may seem in a dense network, articulators who disconnect only isolated individuals, one by one. Deconstruction is certainly a process that differs from searching out the overall shape of a network or using a single, static indicator. It enables us gradually to identify the intermediaries by casting light on their roles within the network.

Thus besides the fact that, as we have seen, it is cited as the most effective by the authors of articles on network attacks, the recalculated betweenness strategy has several advantages for network analysis:

- It does not require a threshold to be chosen for removal of the nodes; those that have the high-

est betweenness centrality are removed in succession and the betweennesses are recalculated after each removal.

- It provides an automatic endpoint criterion: when there are only independent cliques left, all the betweenness centralities are zero.

- The nodes with the highest betweenness centrality at one stage in the process and which, for that reason, are removed from the graph, are called 'intermediaries'. There may be nodes of equal betweenness that are removed at the same time; they are known as 'intermediary blocks'. An additional particularity of some nodes is that they connect parts of the graph that without them would be separated from each other. These nodes or groups of nodes can be called 'articulators'.

- One of the results obtained is that the nodes can be divided into two distinct groups: on the one hand, the intermediaries and, on the other, those elements of the components that become disconnected when the articulators are removed (isolated nodes, isolated dyads or isolated cliques), which we describe as the 'final' nodes, dyads or cliques.

- With the nodes divided in this way, a tree summarising the calculation process can be constructed, the leaves of which are final components and the other nodes the removed intermediaries.

- The number of stages constitutes an indicator of network stratification, which we will call RB (recalculated betweenness) stratification.

- The size and order of exit of the final objects provide indications as to the alters concerned.

- The stage of the algorithm at which an intermediary comes into play is an interesting indicator, since those who appear at the beginning (i.e. when they have a high betweenness at the outset) generally have an overall cohesive effect on the structure, whereas those that emerge at the end of the algorithm are often more local in character. Above all, the algorithm makes it possible to uncover these mid-ranking intermediaries (between 3 and 6). In all cases, they make the connections between small, separate worlds as well as, often, between separate life periods.

5 The intermediaries

5.1 Distribution of the alters among the categories derived from the deconstruction

This process of deconstruction also makes it possible to identify some of the alters' characteristics. Table 1 shows the way in which they are distributed among the various categories at the end of the network deconstruction process. We use the term articulators to denote those intermediaries whose removal disconnects other alters by adding at least one component to the network. They are either single alters or blocks of several alters with identical betweenness. Other intermediaries, such as alters 234 and 261 in Agnès's network, are not articulators. The non-intermediaries are the components with zero betweenness (isolated individuals, dyads or cliques); they were either present initially (rank 1) or revealed by the deconstruction ('final') and of a rank greater than 1.

Table 1. Distribution of types of alters after deconstruction

Type of alter after deconstruction	N	%
Single articulator	579	8.6
Articulator block	746	11.1
Non-articulating intermediaries	146	2.2
Total intermediaries	1471	21.9
Non-intermediary isolated individual	1229	18.3
Element in a non-intermediary dyad	1098	16.3
Element in a non-intermediary clique	2918	43.4
Total non-intermediaries	5245	78.1
Total	6716	100

As might be expected, the intermediaries are in the minority. Among them, in the vast majority of cases they are articulators. The cases in which the removal of the most central node in terms of betweenness does not disconnect any element in the network are relatively few in number.

5.2 The social characteristics of the various types of alters: the example of the life cycle

The sociological validity of this structural characterisation can be verified by testing the effects of sociographic variables that are known to be generally discriminating for personal networks. Let us take the example of age and the life cycle. By combining the types of alters (intermediaries and non-intermediaries) with the type of tie (Ego's family, partner, partner's family, non-family) and Ego's age, a process of change can be observed. With the passing years, the intermediaries, who at the outset are mainly family members and friends of the same age (who account for the vast majority of the alters classed as 'non-family'), come to include partners (97% of whom are intermediaries in the last age group) and their families, a sign of the growing importance of 'constructed' families. At the same time, the share of non-family ties falls steadily. The tendency to incorporate partners into the 'heart' of a network, or rather into its most widely shared part, can be perceived by taking into account how long a couple has been together (Table 2). The partners can be observed moving gradually from a position that is frequently marginal (isolated individual or an element in a clique) to a very central position as the sole articulator.

Table 2. Partners as intermediaries by length of time the couple has been together

	Time together				Total
	< 2 years	2-3 years	4-10 years	> 11 years	
Sole articulator	15.6%	80.0%	75.3%	90.0%	67.6%
Articulator block	12.5%	4.4%	5.9%	0%	6.0%
Non-articulator intermediary	6.2%	2.2%	9.4%	10.0%	7.1%
Isolated individual	43.8%	6%	1.2%	0%	9.9%
Element in a non-intermediary dyad	6.2%	2.2%	1.2%	0%	2.2%
Element in a non-intermediary clique	15.6%	4.4%	7.1%	0%	7.1%
Total	100%	100%	100%	100%	100%

5.3 Ranks in the deconstruction process

The algorithm also records the stage at which an intermediary is removed and thus assigns an exit ranking to each removal. Rank 1 is assigned to Ego (by

definition the highest for betweenness) as well as to initially isolated alters who are connected only to Ego and to the initial cliques, which straightaway have zero betweenness.

Table 3. Intermediaries by exit ranking assigned by the algorithm and types of ties

Type and rank	non-family	Ego's family	in-laws	partner	Total
	Rank 2 intermediaries	25.7%	39.0%	0.3%	35.0%
Rank 3 to 6 intermediaries	34.9%	55.0%	7.2%	2.8%	100%
Intermediaries with a ranking higher than 6	63.2%	25.7%	9.0%	2.1%	100%
Non-intermediaries of rank 1 or 2	73.5%	20.2%	5.4%	0.9%	100%
Non-intermediaries of rank 3 to 6	55.7%	36.6%	7.4%	0.4%	100%
Non-intermediaries with ranking higher than 6	67.1%	20.4%	12.1%	0.4%	100%
Total	58.5%	32.4%	6.4%	2.7%	100%

Non-family alters are by far the most numerous in the alters' set and account for more than half of intermediaries. They are followed by family members, most of whom are parents and siblings. They also constitute more than half of those in ranks 3 to 6. There are more non-family members in the higher ranks (so they are disconnected later in the process), while parents are more frequently found in ranks 3 and 4. Brother and sisters, grandparents, uncles and aunts and cousins are less frequently represented in the first ranks. Finally, and above all, the vast majority of partners (smaller in number of course) are placed in the second rank, so just behind Ego. The non-intermediaries, who do not connect the parts of the network, are either isolated individuals (although they may constitute an individualised echo of a former context in Ego's life, such as the only friend left over from a now-abandoned leisure activity) or members of a structurally undifferentiated group (a clique). The vast majority in all ranks are non-family alters. When they do belong to Ego's family, they usually appear in the middle ranks and when they are in-laws they appear in the high ranks, after a certain number of stages. A priori, therefore, they are more local. Finally, partners are very seldom non-intermediaries. These characterisations correspond fairly closely to social roles. Partners assume their role as articulators at the outset, while Ego's family are represented in the intermediate rankings and friends in the high rankings. Seldom intermediaries at the outset are non-family members, then family members in the intermediate rankings and in-laws in

the high rankings.

5 - The number of stages in the deconstruction process: stratification

At the level of Ego's network, the method reveals an interesting parameter, namely the number of stages required to deconstruct the network, which we have called stratification. This index has a lognormal distribution (Fig. 16). For the calculations that follow as well, we will take the neperian logarithm of the number of stages.

Fig. 16. Distribution of the number of stages (mean= 4.36; standard deviation = 1.953a; N = 287). This indicator is positively correlated with network size and betweenness centralization and negatively correlated with density. This means that the larger a network is and the more dispersed the betweenness centralities are, the more likely it is to be stratified. The denser it is, the less stratified it will be, in particular because density prevents betweenness, as it were, by providing alternative paths.

Table 4. Correlations between stratification and the other parameters of network structure

		betweenness centralization	density	Network size
stratification	Pearson correlation	.174**	-.176**	.503**
	Sig. (bilateral)	.003	.003	.000
	N	287	287	287

We call stratification l_r the standardised residual of the linear regression of stratification l on the three variables with which it is correlated (betweenness centralization, density and network size). Thus this indicator has no linear correlation with these three standard indicators. Rather, it represent what our method provides in the way of new, different information compared with the indicators that characterise the overall shape of the network. If we analyse how it varies according to various Ego's characteristics, two interesting correlations can be observed. The first concerns age: stratification declines with Ego's age. This may mean that betweenness is more differentiated in Ego's youth and becomes more homogenised later, probably around spouse and family. It may also mean that, on reaching adulthood, Ego loses some of the intermediaries in his network; in particular, his childhood friends no longer see his parents, his brothers and sisters no longer see his friends,

and so on.

Table 5. Stratification by Ego's age

age_ego_cl5	mean stratification l_r	No.
17 to 20	.4008041	59
21 to 24	.0758299	102
25 to 28	-.1473632	87
29 to 33	-.4759358	39
Total	0.000	287

p-value: 0.000

The second concerns social origin. Network stratification is higher in the higher social classes and it falls the further down the social scale one goes. This would tend to confirm what the measures and typologies already tell us, namely that the higher social classes have networks that are less dense and richer in structural holes than those at the bottom end of the social scale. Thus it can be added that their networks are also significantly more stratified; they have more differentiated intermediaries and the process of deconstructing them is slower.

Table 6. Stratification by Ego's social origin

Ego's social origin	mean stratification l_r	No.
higher	.2565652	43
middle	.0927736	83
lower	-.1163510	161
Total	0.000	287

p-value: 0.055

The ranks at which the intermediaries were removed also provide an interesting interpretation. By simply combining the rank at which the intermediaries appeared with Ego's social origin, it can be seen that the intermediaries in ranks 2 to 3 are overrepresented in the lower social classes and that the ranks greater than 6 are overrepresented in the higher social classes. Thus the latter tend to have more stratified networks.

Table 7. Intermediaries' rank by Ego's social origin

		social origin of ego			Total	No
		higher	middle	lower		
rank	2 to 3	16.4%	27.5%	56.1%	100%	672
	4 to 5	13.4%	35.8%	50.8%	100%	528
	6 & +	27.3%	36.2%	36.5%	100%	271
	Total	17.3%	32.1%	50.6%	100%	1471

p-value: 0.000

Thus a network's level of stratification corresponds to certain social characteristics. If we take up the case of Florence's complex network, it becomes clear that each stage of the deconstruction corresponds to different periods of her life and the intermediaries linked to them. First of all, in her initial network, Séverine is the only isolated alter as well as being the only recent relation. They met during a Biology course six months ago; Séverine holds the number 1 rank with Ego.

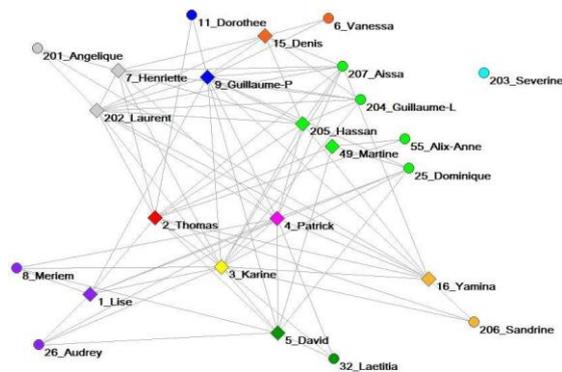


Fig.17. Florence's network in 2004, with names

The first intermediary, ranking second in the deconstruction, is Florence's partner, Thomas. Florence attended high school with Thomas' brother Vincent, and Thomas was in her older sister Karine's class. Vincent has been Karine's partner for some years. Vincent and Thomas were in a theatre troupe together, and now they are co-workers. But, Florence no longer counts Vincent among her strong ties, she finds him irritating and sees him only with Thomas. Thomas was also a childhood friend of Guillaume, Florence's friend from high school. This intertwining of family and old relationships makes Thomas an intermediary between several contexts, which perhaps even contributed to their love affair because, as Florence says, "there are a lot of ties between us". At the third rank we find Meriem, who creates links between several periods. Florence has "always known" Meriem, the daughter of her sister's former nanny. Meriem helped her a lot at the beginning of high school, when Florence was in trouble. Meriem was also in the theatre troupe with Vincent

and Thomas, she lives near Florence and now she is friends with Thomas' ex-partner. She therefore is a constant through all the periods of Florence's life, until today, "In fact, Meriem knows everyone. I introduce her to other people more easily... She has her own friends now because she works at a day camp. But she keeps seeing everyone, because we actually know the same people, we knew the same people at the same time." With Meriem the algorithm removed Guillaume, Thomas' childhood friend, who is therefore, him, going through the periods of Thomas' life. In rows 4 and 5 are Florence's sisters, first Karine whom Florence admires and with whom she shares many activities and friends, then Audrey her little sister. Let us note that Karine exits after Meriem, whereas her initial betweenness centrality was higher. Our process of deconstruction thus enhances Meriem's betweenness once Thomas is removed, as if to time travel back to before his appearance in Florence's life. Indeed, Karine's betweenness was partly based on that of Thomas, as well as that of Vincent who disappeared from the current network, but who left the traces of his connections. Then, in row 6 Florence's father appears, his removal also isolates her cousin. Indeed, her father is connected to two more people (his paternal cousin and Thomas' mother) than her mother who ends up further away. In row 7 appears Denis, Karine's and Thomas' mate, who when removed isolates another of Karine's friends. Next, her mother's removal disconnects a friend of her mother and her grandmother. In the following ranks, friends from high school whom she's known for 3 to 6 years appear as intermediaries connected by two or three, in groups that begin to mix high school and University contexts. Rank 11, the final stage, brings together a very heterogeneous set of Florence's friends from high school, Karine's friends, Thomas's family and friends of the parents with whom Florence recently went on holiday. They are the last local intermediaries. As we can see, this process of progressive network deconstruction highlights Ego's life stages and the resulting relational strata. It also makes it possible to identify the betweenness of the partner who crosses different groups in the present, as well as that of certain alters who cross different temporal strata, such as Meriem. In

this respect, the stratification indicator captures the complexity of the network.

6 Conclusion

Implementing an iterative process for the structural deconstruction of personal networks enables us to avoid reducing their overall structure to a general shape or splitting it up into isolated indicators. In order to give a better account of the structure in a dimension that is highly relevant to sociologists, namely betweenness centrality, we have developed a recalculated betweenness algorithm that offers a number of advantages. It proceeds by removing the alters with maximum betweenness, recalculating the new betweennesses and repeating the operation until all the betweennesses are removed in order to obtain a network of remarkable final shapes (isolated individuals and cliques). It makes it possible to identify the betweennesses that are otherwise concealed by higher values, to pick out the intermediaries at each stage, to record their exit rank, to characterise all the alters by several criteria and, finally, to take into consideration the number of stages required to deconstruct the network, thereby creating a new indicator of network stratification. We chose to use the term stratification to denote the number of stages of the algorithm that are required to exhaust the betweennesses. It does, after all, reflect the diversity of the intermediaries and, in a way, the imprints that time and life events have left on Ego's network. The decline in the degree of stratification with age reflects the fact that the importance of intermediaries in the relational structure weakens over time. However, the fact that, on average, stratification is higher in the higher social classes than in those at the bottom of the social scale indicates that the network's history is generally more present there and better preserved. The ability to preserve old ties had already been identified for the higher social classes, along with the capacity to make more new acquaintances. We can now add that they are also able to position intermediaries between their different life contexts and to retain them. Our analysis of this stratification offers an original perspective that is not provided by indicators of shape such as

density or modularity. It might have been expected that people from the higher social classes would have networks resulting more directly from a strategy that tends to segment and separate the milieus they frequent, as Ronald Burt's structural hole theory would suggest. However, we are not dealing here with the situations he investigates. Our networks do not, on the face of it, reflect competition but relate histories. Our intermediaries are not brokers. They are ties which, by choice or out of sheer inertia, remain present in the network by linking together social milieus from different periods in Ego's life. Thus the deconstruction algorithm developed here induces us to immerse ourselves in the history of the networks and to shed light on their accumulated, shared and preserved strata. It does not suggest any simple explanations because it represents the complexity of those histories. It is a challenge for analysts.

References

- [1] Adilson E. Motter, Ying-Cheng Lai, (2002), Cascade-based attacks on complex networks, *Physical Review*, E66, 065102
- [2] Albert, R., Jeong H., Barabasi A. L., (2000), Error and attack tolerance of complex networks, *Nature*, 406, 378-82.
- [3] Anthonisse, J.M. (1971). The rush in a directed graph. Stichting Mathematisch Centrum. *Mathematische Besliskunde*. Stichting Mathematisch Centrum.
- [4] Bavelas A., (1948), A mathematical model for group structure, *Applied Anthropology*, 7:16-30.
- [5] Berge C., (1962), *The Theory of Graphs*, Methuen & Co, London, translated from the french edition published in 1958 by Dunod. (p.203)
- [6] Bidart C., Charbonneau J., 2011, "How to Generate Personal Networks: Issues and Tools for a Sociological Perspective", *Field Methods*, Vol. 23, no3, pp. 266-286.

- [7] Bidart C., Degenne A., Grossetti M., 2011, *La vie en réseau. Dynamique des relations sociales*, Paris, PUF.
- [8] Bidart C., Degenne A., Grossetti M., 2018, Personal Networks Typologies: A Structural Approach, *Social Networks.*, 54, 1-11.
- [9] Borgatti S.P., Everett M.G., (2006), A Graph-theoretic perspective on centrality, *Social Networks*, 28, 466-484.
- [10] Brandes U., (2008), On variants of shortest-path Betweenness centrality and their generic computation, *Social Networks* 30 (2), 136-145.
- [11] Brandes U., Borgatti S.P., Freeman L.C., (2016), Maintaining the duality of closeness and betweenness centrality, *Social Networks*, 44,153-159
- [12] Burt R. S., (1992), *Structural Holes: The Social Structure of Competition*, Harvard University Press, Cambridge.
- [13] Burt R. S., (2007), *Brokerage and Closure: An introduction to Social Capital*, Oxford University Press, New York.
- [14] Cunningham P., Harrigan M., Wu G., O'Callaghan D., 2013, Characterizing ego- networks using motifs, *Network Science*, 1(2), 170-190.
- [15] De Nooy W., Mrvar A., Batagelj V., (2005), *Exploratory Social Network Analysis with Pajek*, Cambridge University Press.
- [16] Everett M. G., Borgatti S. P., (2005), Ego network betweenness, *Social Networks* 27 31-38.
- [17] Everett M. G., Valente T.H., (2016), Bridging, brokerage and betweenness, *Social Networks*, 44, 202-208.
- [18] Everett, M. G., & Borgatti, S. P., (1999). The centrality of groups and classes. *Journal of Mathematical Sociology.* 23(3): 181-201.
- [19] Flament C. (1962), *L'analyse de similitude*, Cahiers du Centre de recherche opérationnelle, Bruxelles, 4-2, 63-97
- [20] Freeman L., (1977), A set of measures of centrality based on betweenness, *Sociometry*, 40-1,35-41.
- [21] Freeman L., (1978), Centrality in Social Networks, *Conceptual Clarification*, *Social Networks* 1, 215-239.
- [22] Freeman L. C., (1982), Centered graphs and the structure of Ego networks, *Mathematical Social Sciences* 3, 291-304.
- [23] Gould R. V., Fernandez R. M., (1989), Structures of mediation : a formal approach to Brokerage in transaction Networks, *Sociological Methodology*, 19, 89-126.
- [24] Milo R., Shen-Orr S., Itzkovitz S. Kashtan N., Chklovskii D. Alon U., 2002, Network Motifs: Simple Building Blocks of complex Networks, *Science*, 298, October 2002, 824-827.
- [25] Newman M. E. J., 2003, The structure and function of complex networks, 2003, *Review of the Society for Industrial and Applied Mathematics (SIAM)*, vol 45, no 2, 167-256.
- [26] Newman M.E.J., Girvan M., (2003). Finding and evaluating community structure in networks, *Stat. Mech.*, 11
- [27] Petter Holme, Beom Jun Kim, Chang No Yoon, Seung Kee Han, (2002) , Attack Vulnerability of complex Networks, PACS numbers 89.75.Fb, 89.75.Hc, 89.65.-s
- [28] Ravasz E., Barabasi A. L., 2008, Hierarchical organization in complex networks, PACS numbers: 89.75.-k, 89.20.Hh, 05.65.+b Références
- [29] Rosenstiehl P., (1967), L'arbre minimum d'un graphe. in "Theory of Graphs, International Symposium," pp. 357-368.

[30] Gordon & Breach, New York. Sur S., Ganguly N., Mukherjee A., 2016, Brokerage-based attack on real world temporal networks, *Network Science* 4(4),446-459.

[31] Yong Li, Wenguo Li, Yi Tan, Fang Liu, Yijia Cao, Kwang Y. Lee, (2017), Hierarchical Decomposition for Betweenness Centrality Measure of Complex Networks, *Scientific Reports*, 7,46491; doi:10.1038/srep46491. Zhe-Ming LU, Xin-Feng Li, (2016), Attack Vulnerability of Network Controllability, *PLoS ONE* 11(9): e0162289. doi: 10.1371/journal.pone.0162289

The intermediaries	11
5.1 Distribution of the alters among the categories derived from the deconstruction	11
5.2 The social characteristics of the various types of alters: the example of the life cycle	12
5.3 Ranks in the deconstruction process.....	12
Conclusion	15

Contents

1 Introduction	1
2 Betweenness	2
2.1 Different centralities and their meanings	2
2.2 The intermediaries	3
2.3 The measure of betweenness in networks.....	3
2.4 Betweenness in personal networks	4
3 The data	4
3.1 The panel survey	4
3.2 Contrasting forms of betweenness	5
3.3 Networks with zero betweenness	6
4 Network deconstruction algorithm	6
4.1 The standard applications	6
4.2 Our choice of algorithm	7
4.3 The recalculated betweenness algorithm (RB)	8
4.4 The process at work	8
4.5 Rapid deconstruction	9
4.6 A slow deconstruction.....	10