Psychological Aspects of Social Communities

Adrien Friggeri
ENS de Lyon/Inria
Université de Lyon
France
adrien@friggeri.net

Renaud Lambiotte
Deparment of Mathematics/Naxys
University of Namur
Belgium
renaud.lambiotte@fundp.ac.be

Michal Kosinski
The Psychometrics Centre
University of Cambridge
United Kingdom
mk583@cam.ac.uk

Eric Fleury
ENS de Lyon/Inria
Université de Lyon
France
eric.fleury@inria.fr

***Abstract*—Social Network Analysis has often focused on the structure of the network without taking into account the characteristics of the individual involved. In this work, we aim at identifying how individual differences in psychological traits affect the community structure of social networks. Instead of choosing to study only either structural or psychological properties of an individul, our aim is to exhibit in which way the psychological attributes of interacting individuals impacts the social network topology. Using psychological data from the myPersonality application and social data from Facebook, we confront the personality traits of the subjects to metrics obtained after applying the C^3 community detection algorithm to the social neighborhood of the subjects. We observe that introverts tend to have less communities and hide into large communities, whereas extroverts tend to act as bridges between more communities, which are on average smaller and of varying cohesion.**

INTRODUCTION

In the analysis of social networks, particular attention has been dedicated to the structural properties of the direct neighborhood of an individual. This ego-centered approach has been used broadly in psychology and sociology to help at better understanding the relationship between an individual and its proximate social circle, and how individuals are integrated in social life [1]. The position of a person in a network and, complementary, the shape of its ego-network is the source of its social capital. By definition, social capital is considered to be "the sum of the resources, actual or virtual, that accrue to an individual or group by virtue of possessing a durable network of more or less institutionalized relationships of mutual acquaintances and recognition" [2]. A dense, interconnected network of often strong ties is associated to the notion of bonding social, enabling the flow of information within the network and containing an element of trust [3]. In comparison, open networks, containing many intransitive triads, are an indicator of bridging social capital, as an individual bridges structural holes between disconnected others, thereby facilitating knowledge sharing across the system [3].

In parallel to these works, much research has been done to explore empirically, to quantify and to model the organization of large-scale social networks. In particular, there has been a rising interest in the study of communities in social networks, although both the precise definition of a community [4], [5] and the methods to find communities in social network [6] have been subject of debate. On its more basic level, a social community can be understood as a group of nodes presenting a high internal interconnectedness and a relative isolation from the rest of the network.

An important aspect missing in the latter structural studies is a characterization of the variability and differences among individuals, and the effect of non-structural attributes on link formation. A well-known example is homophily [7] stating that similarity, e.g. in terms of status or interests, fosters connection, as similar people tend to select each other, communicate more frequently and develop stronger social interactions. In this paper, we aim at going one step further in this direction, and at identifying how individual differences in psychology affect the structure of ego-networks. Rather than neglecting either structural or psychological properties of an individual, we seek to understand how social network topology is shaped by the psychological attributes of interacting individuals. By doing so, we take an individualized approach to the study of social networks and view the actor as an individual who actively transforms the structure of his or her social network differently depending on his own specific properties.

This paper is organized as follows. In Section I, we will present myPersonality, an online application which allows users to take personality and ability tests and provides unique structural information on those users by relying on the Facebook social network. We then present in Section II the notion of social cohesion, a purely structural metric which captures the extent to which a group of people form a community [8] and presents deep links to sociological constructs related to the way information flows on the network [9]. Finally, in Section III we will study the impact of psychological traits on social structuring and reveal, among other things, that extraversion correlates positively with the number of friends and the number of communities of the subject, and negatively with the size of their communities and the extent to which their social neighborhood is partitioned. Although indirectly related to psychology, we will also discuss how how the age of the subjects affects the structure of their ego-network.

I. SOCIO-PSYCHOLOGICAL APPROACH

In this section, we present the data we will analyze and its source. Since our interest lie in the study of the way psychological traits are linked to topological features of the social network, we will analyze the results of an online psychological test called myPersonality, in relation to the social entourage of the test subjects on Facebook. We will

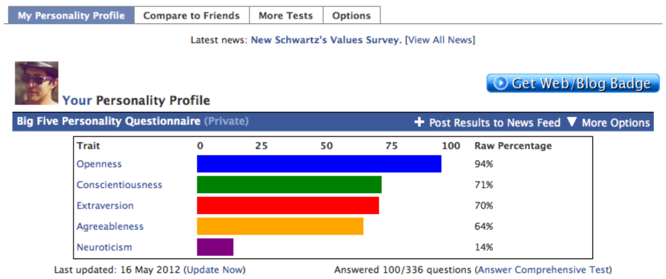
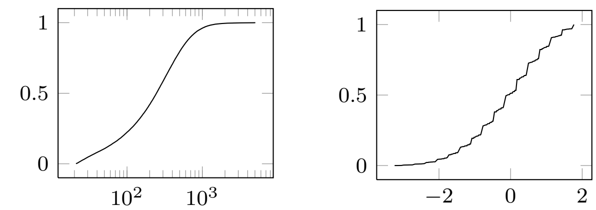


Fig. 1. Screenshot of part of the myPersonality application



(a) Degrees CDF (b) Extraversions CDF

Fig. 2. Cumulative distributions of (*left*) Degrees and (*right*) Extraversions

then provide a broad range of related works and findings on the links between psychological traits and structural features at the network level.

A. Two Datasets

myPersonality is a Facebook application which allows users to take a variety of personality and ability tests (Figure 1). Users also have the possibility to opt in and give their consent to share their personality scores and their social neighborhood on Facebook for scientific purposes. It should be noted that the psychological and structural data come from two different sources – respectively myPersonality and Facebook – and as a consequence are *a priori* independent one from the other.

The application undertakes a series of measures to avoid that users respond in a careless or mischievous way [10], and thus to assure the highest quality of its database, *e.g.* by removing unreliable results through numerous validity tests. It has been shown that the quality of the responses is at least as high as in traditional pen-and-paper studies, with the significant advantage of reaching a much broader and less biased audience. The benefit in using the myPersonality data is that is allows to tap in a unique source which contains both psychological traits of the subjects and link this psychological profile to social information – list of friends and friendships between friends – extracted from their social graph on Facebook.

Personality is measured by the so-called five-factor model of personality [11], which associates to each individual five scores corresponding to five main personality dimensions. Each dimension, labeled as OCEAN, can be summarized as follows:

- **Openness** for spontaneity and adventurousness, denotes an appreciation for emotion, a sensitivity to beauty and intellectual curiosity;
- **Conscientiousness** for ambition and persistence, denotes a tendency to act dutifully and a planned rather than spontenous behavior;
- **Extraversion** for sociability and excitement seeking, denotes an energetic and spontaneous personality and the tendency to seek stimulation in company of others;
- **Agreeableness** for trustingness and altruism, denotes a tendency to be compassionate an cooperative towards others;
- **Neuroticism** for emotional liability and impulsiveness, denotes a personality prone to experiencing negative emotions easily, *e.g.* anger, anxiety and depression.

The **social** data comes from Facebook, and as such we have, for each subject, both the list of their friends and the information about pairs of those who are also friends on Facebook. We adopt an egocentered approach and, rather than considering the impact of all individuals on the whole Facebook social network, we will focus on the way the individual shape the social structure of their ego-network. The ego-network of an individual is the subgraph containing only his friends, that is, Ego does not belong to his ego-network as he would bring no information to the structure given that he is connected to all of his friends.

The ego-network approach has a remarkable property, as there is a direct correspondance between classical network metrics on the original network and the ego-network. Obviously, the degree d – or number of friends – of a subject in the original network is equal to the size of the ego-network. More interestingly, the clustering coefficient of the subject in the original network is equal to the density of his ego-network. We will equivalently mention the *degree of a subject* and the *size of the ego-network*, and similarly the *clustering coefficient of a subject* and the *density of the ego-network*.

Our final dataset consists of a sample of $44{,}096$ Facebook users who have taken the big five personality test. For each of these subjects, we have access to their five personality traits, their age and gender, and their ego-network. We will focus on the users whose number of friends on Facebook is comprised between 50 and 2000 (excluding 35 users with degree greater than 2000 and 3259 users with degree smaller than 50). The final sample contains $49{,}623$ users, whose cumulative degree distribution is represented in Figure 2(a). Given that in the following we will mainly focus on the trait of extraversion, we have represented on Figure 2(b) the cumulative distribution of our sample subjects' extraversion.

B. Related Work

In general, the five-factor model has been shown to predict a broad range of real-world behavior [12], for instance how marriages turn out and people's taste in movies [13]. The five personality factors also relate to people's behavior in a broad variety of social contexts. It is likely that they predispose people's propensity to form more or fewer social ties, and may be related to the extent to which others form relationships with the focal actor. For instance, extroverts are expected to approach others more easily and engage in more social interaction.

Previous research has explored the relationship between personality and structural properties of ego-networks. In that respect, it has consistently been shown that extraversion is associated with the size of the ego-network and with greater social status [14], [15], [16], [17]. Other dimensions have also been argued to have an effect on social network topology [18], but findings tend to be inconsistent in the literature, and no significant correlation has been found in a recent study on a large sample of users of myPersonality [17].

In studies interested in other aspects of the ego-network, it has been observed that extroverts are not emotionally closer to individuals in their network [19], despite an increased size. It has also been shown that self-monitors, the chameleons of the social, are more likely to have a high centrality [20]. Brokerage also appears to be related to personality. People whose networks bridge structural holes are more likely to have an entrepreneurial personality [21]. In another study, it was shown that extraversion is positively associated with closed triads of strong ties and neuroticism with closed triads of weak ties [22]. The existence of different types of structural configurations has also been proposed, each associated to the social and psychological characteristics of an individual: people embedded in dense networks, people having several subsets of alters, etc. [23].

II. From Friends to Communities

Contrary to previous work, which studied the relationship between psychological traits and graph metrics attached to nodes, such as degree, centrality or clustering, our aim is to capture more precisely the topological features of the subject entourage and confront these to his personality. To that effect, we rely on the notion of social communities, which help at understanding the organization of their social neighborhood.

For a long time, communities were thought of as disjoint set of nodes, which is not compatible with the fact that individuals typically belong to multiple communities, e.g. to a "family" community and a "friends" community. In the following, we thus consider the problem of detecting overlapping communities. Instead of "partitioning" the nodes into disjoint communities, we thus aim at covering each of them by at least one community. To do so, we cover each ego-network by using the notion of social cohesion, a graph metric which quantitatively captures to which extent a set of nodes forms a social community [8]. This covering also provides us with additional metrics which capture the structural organization of the ego-network. In this section, we will first recall the definition of the metric and describe the community detection algorithm derived from it.

A. Cohesive Communities

Instead of relying on the number of edges either inside or crossing the boundary of a community, as it has traditionally been done with limited success, the metric draws from the classical notion of triadic closure in sociology. We will say that a *triangle* is a set of three pairwise connected nodes in a network and we write $\textcircled{\triangle}(S)$ the number of triangles contained in a set of nodes S and $\underline{\triangle}(S)$ the number of outbound triangles of S, that is the number of triangles in the network having exactly one edge in S. Cohesion is defined (Eq. 1) as the product of two terms which respectively quantify the triangular density – or transitivity – of the associated subgraph, and the relative isolation of the set from the rest of the network.

$$\mathcal{C}(S) = \underbrace{\frac{\textcircled{\triangle}(S)}{\binom{|S|}{3}}}_{\text{transitivity}} \times \underbrace{\frac{\textcircled{\triangle}(S)}{\textcircled{\triangle}(S) + \underline{\triangle}(S)}}_{\text{isolation}} \qquad (1)$$

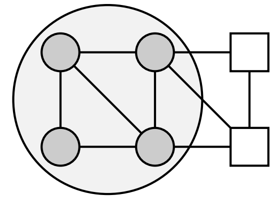


Fig. 3. In this example, the set of circle nodes contains 4 nodes, features 2 inbound triangles and only 1 outbound triangles, leading to a cohesion $\mathcal{C} = \frac{1}{3}$.

See for example the diagram on Figure 3, where the set of circles has a cohesion $\frac{1}{3}$. Cohesion is reminiscent of the classical idea that a community is a set of nodes featuring a high inner density and relatively isolated remains, even though this property is measured in terms of triangles instead of edges. The use of this metric to measure the quality of communities was validated through a large-scale experiment on Facebook in which we had observed that it was highly correlated to the users' perception of the quality of social communities. As a consequence, the cohesion of a set of nodes is a good indicator of its communiti*ness*.

B. Finding Communities with C^3

Although the problem of finding communities with maximal cohesion is an $\mathcal{NP}$-hard problem, we have proposed C^3, a heuristic algorithm to cover a network with highly cohesive communities [24]. The fundamental idea behind C^3 is twofold, first we build communities locally and independently from one another by choosing appropriate seeds in the graph and then expand around those in order to find the most cohesive community containing each seed. In a second phase, we take the set of communities and construct the final covering of the network by merging selected communities in order to bound the maximal amount of overlap between communities.

The first part of the algorithm, can be described as follow: we select a node in the network and an uncovered edge connecting to that node, then we expand around this edge by repeatedly adding new candidate nodes. Those added nodes have to satisfy one of two criteria, either their addition directly increases the cohesion of the set, or it increases the transitivity of the set. While quite similar in spirit to a simple greedy optimization of the cohesion, this second condition is added in order to test branches which drill into the heart of communities by locally increasing the transitivity of the set. Once all the set of possible candidates has been exhausted, the algorithm returns the most cohesive group it has encountered and repeats the same process as long as there are nodes which are not part of a community

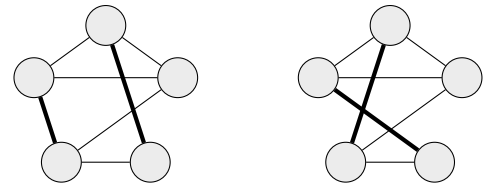
One should note that the first part of the algorithm maximizes locally the cohesion and thus does not take into account another aspect of social communities which is the extent to which they overlap. In some cases, it might desirable to impose no constraint on this overlap – the most extreme case being when we allow communities to be embedded into one another. In this case, however, we wish to control the amount of overlap between the communities, without completely prohibiting it either. To do so, let us take advantage of an analogy between finding cohesive groups in a network and finding sets of strongly mutually overlapping communities. Consider two overlapping communities, S_1 and S_2. We introduce an overlapping weight defined as $O = \frac{|S_1 \cup S_2|}{\min(|S_1|,|S_2)}$. With this at hand, we can construct a "community network", or meta-network, in which each node is a community, and each edge represent the overlap between communities, if the overlapping weight is above a certain treshold, $O \geq 0.5$ in our case. Then, finding sets of mutually overlapping communities boils down to the know problem of findings communities in the meta-network – or, to give them a name, meta-communities. We can therefore recursively use C^3 to compute the meta-communities, which gives us sets of strongly mutually overlapping communities. We can then merge then communities contained inside each meta-community in order to limit the maximum overlap.

C. Communities of the Ego-Networks

Let us now apply C^3 to the ego-networks we have described in Section I-A in order to obtain a covering of each ego-network with overlapping communities for each Ego. Note that, given the definition of the cohesion, isolated nodes and more generally nodes which do not belong to a triangle are not considered to be part of any communities As a consequence, contrary to other community detection algorithms C^3 does not force people into communities and therefore the covering might not be complete.

As we wish to study the links between Ego's psycho-social characteristics and the community structure of his ego-network, we will restrict the set of users to those whose surrounding topology consist at least of 50% of friends present in one or more community (24,285 subjects). By using C^3 on our final data set, we have obtained $974{,}677$ communities.

In order to compare our findings to a baseline, we introduce, for each subject, a random null model ego-network. Since we wish to study the impact of the psychological traits on the community structure of the ego-network, we need to control for other topological factors such as size and density of the ego-network. In order to do so, we will construct the null model ego-network G_R by randomly rewiring edges of the original ego-network G in the following manner: choose two distinct edges randomly, such that the ends of those two edges are four distinct nodes, and swap the ends of those two edges,



(a) before rewiring (b) after rewiring

Fig. 4. Illustration of the rewiring process: (a) two edges are chosen and marked with thick lines, (b) swap the ends of the two chosen edges.

as illustrated on Figure 4. We obtain the null model ego-network G_R after repeating this procedure m times, where m is the number of edges in G. During this random rewiring, each edge has been rewired on average twice, guaranteeing a the randomness of the graph G_R.

Note that at each step of the rewiring, no node or edge are added nor deleted, which guarantees that the null model ego-network has same size and density that the original ego-network. Furthermore, because at each step we swap the ends of edges, the degree of each node remains constant, and therefore the distribution of degrees of both G and G_R are identical. This means that the null model would be the ego-network of an individual which would have same degree and clustering coefficient that the original subject.

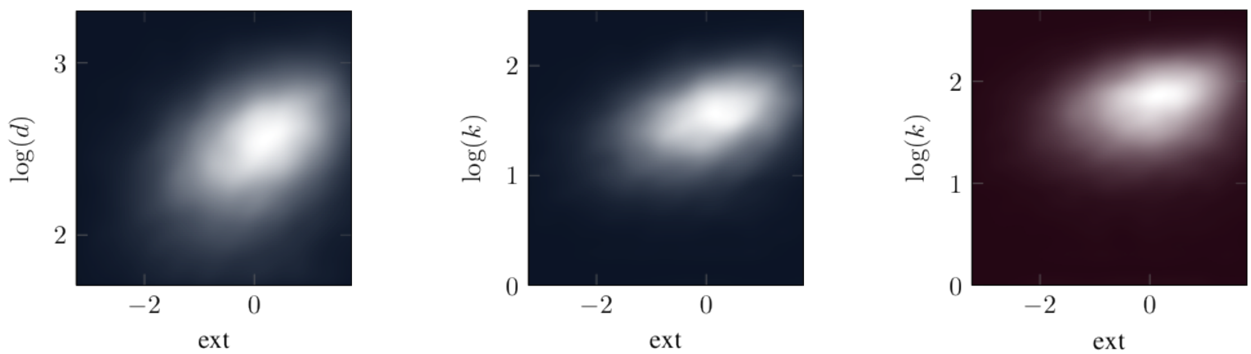
We have applied C^3 to each null model ego-network and obtained $1{,}709{,}883$ communities. In the following section, where relevant, we will apply the same computations both to the original ego-network and the null model in order to highlight the effects due to the deeper community structure.

III. Psychology of Structural Features

Given the personality data from Section I, and the social and null model ego-networks which we have enriched by computing the social communities in Section II, we will now exhibit the links between psychological traits and community structure, using the null model as a baseline to compare to. Given that we observed no significant correlations between either Openness, Conscientiousness, Agreeableness or Neuroticism and structural features, the results presented here are mostly related to the Extraversion factor, although we will also observe the impact of age on the topology of social neighborhood.

A. Degree & Number of Communities

As stated in Section I-B, it has been observed repeatedly that there is a linear Pearson correlation between the number of social connexions maintained by an individual and his extraversion. Note that we will consider the logarithm $\log(d)$ of the number of contacts (or degree) rather than the actual number of friends due to the high variability of its value. On Figure 5(a), we have represented a kernel density estimation of $\log(d)$ as a function of the extraversion. As expected, we observe a moderate correlation $r = 0.301$ (p-value $< 10^{-100}$) which indicates that the more extroverted users tend to add more friends on Facebook.



(a) Degree vs. Extraversion

(b) Number of Communities vs. Extraversion

(c) Number of Communities vs. Extraversion (null model)

Fig. 5. Kernel density estimations of (a) degree as a function of extraversion, and number of communities as a function of extraversion on the (b) original data and (c) null model.

Given this correlation between number of friends and extraversion, it is only natural to look at the link between the number of communities in a subject's ego-network and extraversion. We also observe a moderate correlation $r = 0.293$ (p-value $< 10^{-100}$) between extraversion and number of communities (Fig. 5(b)). This result holds in part on the null model where we observe a somehow smaller correlation $r = 0.196$ (p-value 1.2×10^{-207}) between extraversion and number of communities (Fig. 5(b)).

A possible explanation to this observation is that there is a strong correlation $r = 0.894$ (p-value $< 10^{-100}$) between the number of communities $\log(k)$ and his number of friends $\log(d)$. Interestingly, on the null model, we observe a smaller correlation $r = 0.788$ (p-value $< 10^{-100}$) between these two quantities.

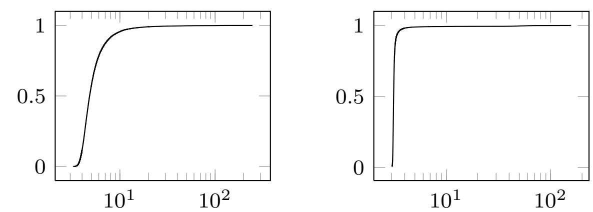
It is important to point out that all the correlations are less important in the null model than in the original data. This leads us to conclude that there is a direct contribution of the actual community structure in the original ego-networks to correlations. That is, part of the correlation between extraversion and communities which cannot only be explained by the degree alone.

Summary: As expected, we have observed that more extroverted subjects tend to have more friends on Facebook. More interestingly, we have also observed that the friends of more extroverted subjects are split across more communities. Our data suggest data suggest that extroverted people not only maintain more social relationships but also interact with a larger number of social groups.

B. Community Sizes

We have seen in the previous section that there is a correlation between number of communities and extraversion. In this context, it is legitimate to look at the relationship between the size of those communities and extraversion. For each user, we will define the average size $\bar{s}$ of his communities (c) as:

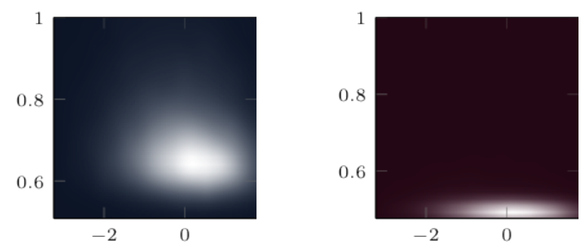
$$\bar{s} = \frac{\sum_c \mathcal{C}(c)|c|}{\sum_c \mathcal{C}(c)}$$



(a) Av. Size CDF

(b) Av. Size CDF (null model)

Fig. 6. Cumulative distributions of average community size in (a) original data and (b) null model.



(a) log Av. Size vs. Extraversion

(b) log Av. Size vs. Extraversion (null model)

Fig. 7. Kernel density estimations of average community size as a function of extraversion in (a) original data and (b) null model.

We weight the sizes by cohesion in order to give more importance to good communities. The cumulative distribution of average sizes represented on Figure 6(a) shows that most communities have a rather small size (50% have $\bar{s} \leq 4$ and 95% have $\bar{s} \leq 10$). In the case of the null model, the average size of communities is evan smaller, which contrast with the original data, as 95% of the null model ego-networks have an average community size $\simeq 3$ (Figure 6(b)).

There is a small negative correlation $r = -0.11$ (p-value 1.57×10^{-66}) between $\log(\bar{s})$ and extraversion, which shows that more extroverted subjects have smaller communities, and conversely (Figure 7(a)). Strikingly, this has to be contrasted with the absence of correlation $r = -0.0574$ (p-value 4.16×10^{-19}) between average size and extraversion in the case of the null model (Figure 7(b)). From this we conclude that correlations observed in the original data are mainly due to the community structure.

Summary: We have observed that introverted subjects tend to be in larger groups and that extroverts tend to be in smaller groups. This observation suggests that introverts are more naturally inclined to blend into larger groups, rather smaller groups, to avoid being the center of attention. On the contrary, extroverts might prefer being part of a large number of smaller groups in order to have more chance of attracting attention.

C. Variability of Cohesion

Until now, we have analyzed the community structure in terms of number of communities of the subjects and in terms of number of people in those communities, that is the average size of those communities. We now focus on the the average cohesion – weighted by size – of a user's communities, which is defined as:

$$\bar{\mathcal{C}} = \frac{\sum_c \mathcal{C}(c)|c|}{\sum_c |c|}$$

The average cohesion $\bar{\mathcal{C}}$ captures, for each user, to which extent he is part of socially cohesive environments. Interestingly, we have found that there is no significant correlation between this quantity and any of the big five personality traits. However, we have observed a small correlation ($r = 0.123$, p-value 1.96×10^{-82}) between the standard deviation of the cohesion $\sigma_{\mathcal{C}}$ and extraversion.

$$\sigma_{\mathcal{C}} = \sqrt{\frac{\sum_c |c|(\mathcal{C}(c) - \bar{\mathcal{C}})^2}{\sum_c |c|}}$$

In the null model, this correlation is not present ($r = 0.0676$ p-value 7.05×10^{-26}) which suggests a relation between extraversion and the heterogeneity of social communities.

This correlation between cohesion variability and extraversion is related to the fact that that subjects who belong on average to larger communities tend to evolve in social groups of similar cohesion there is a strong negative correlation $r = -0.806$ (p-value $< 10^{-100}$) between $\sigma_{\mathcal{C}}$ and $\log(\bar{\mathcal{C}})$. One mechanism may explain such a correlation: the number of large communities is much lower, and those are typically of low cohesion, given that a number of inbound triangles proportional to the cube of the size of the community would be needed in order to maintain a high cohesion.

Summary: We observe a positive correlation between the standard deviation of cohesion and extraversion. We understand this in terms of social adaptability, which is key in the definition of extraversion: extroverts are members of different communities with highly varying cohesion, and are thus members of social communities which can be tight groups of close friends as well as more sparse communities of more distant acquaintances. On the other hand, as we have seen before, introverts tend to hide in larger groups and those groups tend to have an average cohesion, i.e. introverts tend to lack highly cohesive communities which would increase the variability of cohesion.

D. Partition

Another aspect of communities is the amount of overlap. As described earlier, one of the strengths of C^3 is that it computes communities without imposing the constraint that a subject belongs to one and only one community. As a matter of fact, C^3 does not impose the constraint that a subject belongs to at least one community either: a node might be present in 0, 1 or more communities. For the sake of clarity, nodes of an ego-network which are in at least one community will be referred to as *covered*.

A simple way to capture the notion that some of the covered nodes are in more than one community is to look at the partition ratio $p = \frac{|\bigcup c|}{\sum |c|}$.

This partition ratio quantifies the extent to which the communities of an ego-network are disjoint from one another. It is equal to 1 when all covered nodes are exactly in one community, that is when covered nodes are partitioned into communities, and it decreases as more nodes are present in several communities, *i.e.* when the overlap increases.

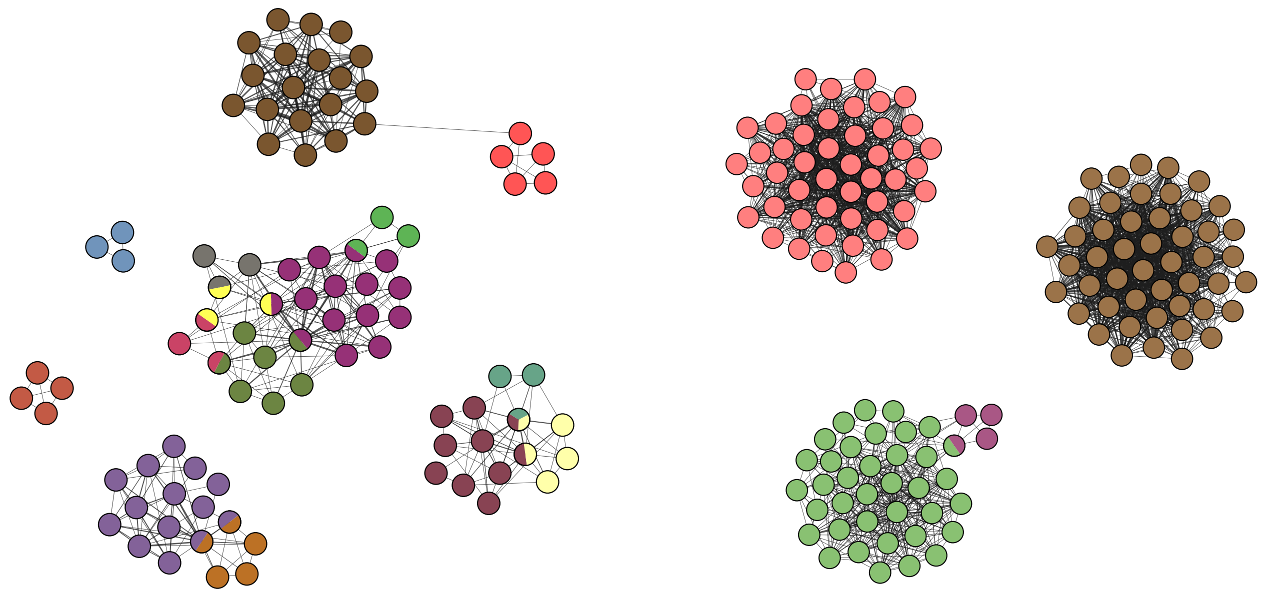
We observe a small negative correlation $r = -0.132$ (p-value 1.56×10^{-95}) between δ and extraversion in the original data whereas the correlation $r = -0.0857$ (p-value 1.03×10^{-40}) is negligible on the null model.

Summary: We observe a negative correlation between extraversion and the partition ratio, which implies that there is a link between the compartmentalization of the ego-network and the subjects extraversion. More extroverted subjects tend to be in groups which are intricately linked to each other whereas less extroverted subjects tend to be in more distinct and separate social groups. This observation is compatible with the hypothesis that extroverts act as bridges between communities and introduce individuals form one community to those in another one.

E. Cohesion, Degree and Age

In this last section, we explore the effect of age on network topology. Although not a psychological trait, age is part of the identity of the subjects and as such has an impact on the structure of their ego-network. We observe, as it was the case in [17], that there is a negative correlation $r = -0.194$ (p-value 8.51×10^{-193}) between age and degree, which tends to indicate that the elder have a smaller social neighborhood on Facebook whereas the younger have more friends.

Our community based approach reveals a moderate correlation $r = 0.271$ (p-value $< 10^{-100}$) between age and the average cohesion $\bar{\mathcal{C}}$ which is not present in the null model, where the correlation is $r = 0.0843$ (p-value 2.37×10^{-37}). This observation suggests that older individuals belong to denser communities on Facebook, whereas the younger are part of sparser ones.



(a) User A, 26 years old: high extraversion (ext $= 1.33$), 101 friends of which 91 are split across 15 communities of size varying between 3 and 19, and average cohesion $\bar{\mathcal{C}} = 0.46$.

(b) User B, 19 years old: low extraversion (ext $= -1.21$), 145 friends of which 136 are split across 4 communities of size 4, 37, 48 and 48, and average cohesion $\bar{\mathcal{C}} = 0.31$.

Fig. 8. Examples of two ego-networks of subjects with different psychological traits and structural features.

Finally, let us mention a small correlation $r = 0.171$ (p-value 4×10^{-150}) between age and the Conscientiousness factor, which indicates that the older subjects exhibit less spontaneous behavior than younger ones. Intriguingly, though, we do not observe any correlation between degree and conscientiousness, nor between average cohesion and conscientiousness.

Summary: We have observed an impact of age on the structural properties of ego-networks, as the older a subject gets the less friends he has on Facebook. This observation can be explained by the fact that younger subjects are more active on Facebook and tend to add more friends than older ones. Our analysis also reveals that older subjects tend to be part of more cohesive groups than the younger ones, which might come from the fact that younger subjects are less careful before adding someone as a friend on Facebook.

F. Illustration

In this section, we focus on visualizations of ego-networks in order to illustrate the quantitative findings of the previous sections. Figure 8(a) shows a drawing of the ego-network of a user, which we will call A, with high extraversion (ext $= 1.33$). Figure 8(b) shows the ego-network of user B who is more introvert (ext $= -1.21$). Let us recall that an ego-network is the subgraph containing the neighborhood of Ego, and that Ego and its links are not represented. Each circle in the visualization represents a friend of the subject and their community is represented by their color. If a friend belongs to several communities, the circle is divided into equal regions and each slice is colored with the color of a community. This procedure makes the overlapping community structure of the ego-network immediately visible.

Notice the differences between both networks. First, user A has a degree ($d_A = 101$) which is slightly less than B ($d_B = 145$) even if both numbers have the same order of magnitude. More interestingly, the organization of those friends into communities is strikingly different. B has four communities, three of which contain more than 35 friends. Moreover, two of those groups are totally isolated from the others. On the other hand, A exhibits 15 communities which are much more interconnected and of smaller size.

This figure illustrates the different ways in which personality impacts the structure of the ego-network. More specifically, it shows that a user with low extraversion, such as B, is part of a few large cohesive groups which are compartementalized, whereas a user with high extraversion, such as A, evolves in more different social groups, of smaller sizes, varying cohesion and overlapping one another.

IV. Conclusion

In this work, we have exhibited several ways in which the personality of individuals may affect the structure of their social network. To do so, we have studied a data-set from the web application myPersonality which provides psychological traits of its users, together with Facebook data about their social neighborhood – list of friends and friendships between friends.

We have applied a community detection algorithm, C^3, to the ego-network of each subject and compared their community structure to their personality profile. We have first verified that the number of friends is positively correlated to the trait of extraversion, a well established result. More importantly, we have shown that extraversion not only affects the size but also the topology of the ego-network. First, it is positively correlated with the number of communities in the ego-network of a subject, and negatively correlated with the size of those communities. We have also observed a positive correlation between extraversion and how the cohesion of the communities of a subject varies, and a negative correlation between extraversion and the compartmentalization of the ego-network into communities.

These observations lead us to hypothesize different types of behavior for introverts and extroverts, as introverts tend to hide into larger groups in order to avoid attracting attention and are part of segregated groups, whereas extroverts display a greater adaptability to a wide variety of social contexts, are involved in more malleable social groups and act as bridges between social groups.

REFERENCES

[1] S. Wasserman and K. Faust, "Social Network Analysis," Cambridge University Press, Nov. 1994.

[2] P. Bourdieu and L. J. D. Wacquant, *An invitation to reflexive sociology*. University Of Chicago Press, Jul. 1992.

[3] R. S. Burt, *Brokerage and Closure*, ser. An Introduction to Social Capital. Oxford University Press, Aug. 2005.

[4] G. A. Hillery Jr, "Definitions of community: Areas of agreement," *Rural sociology*, vol. 20, no. 2, pp. 111–123, 1955.

[5] D. E. Poplin, "Communities: A survey of theories and methods of research," Macmillan, 1979.

[6] S. Fortunato, "Community detection in graphs," *Physics Reports*, vol. 486, no. 3-5, pp. 75–174, Jan. 2010.

[7] M. McPherson and L. Smith-Lovin, "Birds of a feather: Homophily in social networks," *Annual review of sociology*, vol. 27, pp. 415–444, 2001.

[8] A. Friggeri, G. Chelius, and E. Fleury, "Triangles to Capture Social Cohesion," in *2011 IEEE Third International Conference on Social Computing (SocialCom 2011)*, Cambridge, 2011, pp. 258–265.

[9] M. Granovetter, "The Strength of Weak Ties," *American journal of sociology*, vol. 78, no. 6, pp. 1360–1380, May 1973.

[10] T. Buchanan and J. L. Smith, "Using the Internet for psychological research: Personality testing on the World Wide Web," *British Journal of Psychology*, vol. 90, no. 1, pp. 125–144, 1999.

[11] P. T. Costa and R. R. McCrae, "NEO Personality Inventory Revised NEO-PI-R Test Manual," SAGE Publications, May 2005.

[12] D. Nettle, "Personality: what makes you the way you are," Oxford University Press, 2007.

[13] O. Chausson, "Assessing the impact of gender and personality on film preferences," *Cambridge University*.

[14] C. Anderson, O. P. John, D. Keltner, and A. M. Kring, "Who attains social status? Effects of personality and physical attractiveness in social groups." *Journal of Personality and Social Psychology*, vol. 81, no. 1, p. 116, 2001.

[15] D. S. Berry, J. K. Willingham, and C. A. Thayer, "Affect and personality as predictors of conflict and closeness in young adults' friendships," *Journal of Research in Personality*, vol. 34, no. 1, pp. 84–107, 2000.

[16] R. Swickert, C. J. Rosentreter, J. B. Hittner, and J. E. Mushrush, "Extraversion, social support processes, and stress," *Personality and Individual Differences*, vol. 32, no. 5, pp. 877–891, 2002.

[17] D. Quercia, R. Lambiotte, M. Kosinski, D. Stillwell, and J. Crowcroft, "The personality of popular Facebook users," *Proc. ACM Conf. Comput. Support. Cooperat. Work*, 2012.

[18] L. A. Jensen Campbell and W. G. Graziano, "Agreeableness as a moderator of interpersonal conflict," *Journal of personality*, vol. 69, no. 2, pp. 323–362, 2001.

[19] T. V. Pollet, S. G. B. Roberts, and R. I. M. Dunbar, "Extraverts Have Larger Social Network Layers," *Journal of Individual Differences*, 2011.

[20] M. Kilduff and D. Krackhardt, "Bringing the individual back in: A structural analysis of the internal market for reputation in organizations," *Academy of Management Journal*, pp. 87–108, 1994.

[21] R. S. Burt, J. E. Jannotta, and J. T. Mahoney, "Personality correlates of structural holes," *Social Networks*, vol. 20, pp. 63–87, 1998.

[22] Y. Kalish and G. Robins, "Psychological predispositions and network structure: The relationship between individual predispositions, structural holes and network closure," *Social Networks*, 2006.

[23] B. Wellman and University of Toronto. Centre for Urban and Community Studies, "Studying personal communities in East York," Sage, Beverly Hills, 1982.

[24] A. Friggeri and E. Fleury, "Finding cohesive communities with C3," Tech. Rep. RR-7947, Apr. 2012.