Facencounter: bridging the gap between offline and online social networks

Sabrina Gaito, Gian Paolo Rossi and Matteo Zignani
Department of Computer Science
Universitá degli Studi di Milano
Milano, Italy
Email: firstname.lastname@unimi.it

***Abstract*—Human beings are involved in a broad range of social relationships spanning from real life experiences to online media and social networks. This is leading people to act in a multilayered complex network whose relationships among different layers have still to be analyzed and understood in depth. In this paper, we focus on this problem by comparing and overlapping the online sociality (Facebook network) and the offline sociality (encounter network) of a group of students. First we describe the experiment we performed to trace the encounters, occurring with people both inside and outside the group of experimenters, and to gather information about their online friendships. On the basis of the obtained dataset, we obtain the relevant complex networks, whose separated analysis lead us to observe significant structural differences. Moreover, we study the correlations and overlap between the two interrelated networks, showing how users' centralities change in the two networks. Finally, the information transfer across layers of the unified complex network enables us to obtain results about the effects on paths and centrality.**

I. Introduction

If online social networks (OSNs) were to mirror the offline sociality of individuals, they would be able to reflect offline relationships and unveil the social behaviors that impact on online sociality. Unfortunately, there is a growing belief that today's social networks are quickly shifting away from their original goal and, by contrast, sparking fears about the fact that they are drifting towards a highly connected, unstructured and flat social graph. The in-depth understanding of relationships between online and offline sociality, beyond being a key issue of Human Sciences, would produce the practical effect of promoting OSNs to the status of best platform for the effective delivery of mobile computing services (e.g. recommendation systems, advertising, content dissemination, crowd sourcing, social discovery, etc.). In fact, the online deployment of these services would benefit offline social knowledge, for instance, improving trustworthiness of a service, tailoring it according to the target's interests, leveraging context information or predicting the impact on OSN structure of an upcoming event. On the opposite end, it would definitely help in deploying improved mobile services which leverage their online life features.

The above arguments concern a specific research interest whose main goal is to understand the relationship between the two faces of sociality. In this paper, we are interested in answering the following questions:

- Q1: *How do offline and online social networks relate to one another and to what extent do they overlap?*
- Q2: *Is a person's popularity uniform, i.e. more or less the same in all social dimensions, and how do different centrality metrics account for people's popularity in this multilevel network?*
- Q3: *How can we embed into a unique complex network all human social dimensions so as to increase the expressiveness of usual metrics?*

The challenges in answering these questions are both experimental and theoretical. On the one hand, although large datasets describing online social networks have recently been made available together with an extensive literature, datasets concerning offline encounters are few and, due to the limitation of the short range radio technology currently adopted to detect contacts, combine both explicit and opportunistic contacts. As a consequence, the research community has very few opportunities to compare the datasets of offline encounters and online relationships of the same group of individuals. On the other hand, while the modeling of a single layer of sociality has been successfully faced by means of complex network theory, the merging of interrelated complex networks still presents theoretical aspects to be investigated.

In this paper we describe an experiment that enabled us to answer the above questions by exploring the intimate relation between online and offline sociality of a group of students. Data describing the offline sociality of a set of 35 volunteers were purposely collected in a time span of one month and then integrated and compared with relevant data about their online sociality extracted from Facebook. The two layers of the group's sociality were described through the associated complex networks. Our paper offers the following main contributions. First, it shows that *the overlapping degree is low*; in fact, the sets of Facebook and real contacts are quite different. Secondly, by comparing the ranking induced by several centrality metrics, the paper shows that *node centrality is not a universal feature*. In fact, node centrality is not linearly transferred across layers and, as a consequence, the people's popularity is most likely to change in different networks. Finally, the paper introduces a *unified complex network* which allows us to merge offline and online relevant features shedding a light on how human behavior is interwoven

978-0-7695-4911-8/12 $26.00 © 2012 IEEE
DOI 10.1109/SITIS.2012.116

across layers. To the best of our knowledge, this is the first paper addressing the above mentioned issues and providing a set of relevant preliminary answers.

II. Related Work

While there is a very extensive literature on online social networks, research on offline sociality and how it relates to online friendships is still in its infancy. Several works have faced this issue by analyzing mobility traces, also containing information about social ties between the nodes (from WLAN Access Point associations [1], Bluetooth contacts [2] [3] [4] or other technologies [5] [6]). Analysis of such traces has shown that there is some correlation between mobility and social connections. However these studies fail to reveal which nodes would actually experience an encounter during which they could communicate. Some experiments have attempted to collect data on offline and online social relationships. Their main goal is to exploit these data for purposes of designing opportunistic routing algorithms that take into account online sociality. The first one, described in [7], gathers contacts detected by an ad hoc wireless device and the Facebook graph restricted to the participants. A similar approach was performed in [8], where they adopt a different contact detection technology. Both experiments suffer from limitations due to the detecting technology: indeed they detect proximity and not an encounter between willing parties. A small step forward came from the experiments in [9] and in [10]. In [9] the authors developed a Facebook application where a small group of experimenters reported their daily face-to-face meetings with other Facebook friends. In this way, however, only relationships among Facebook friends can be analyzed, so leaving out all friends in real life who are not Facebook contacts. In [10] the authors developed an application which integrates data from online social networks and RFID contacts.

The study of the superposition of networks defined on the same set of nodes originates from social sciences where it has been applied to small-scale networks. Only recently multidimensional relationships have been investigated in socio-technological networks, addressing specific problems. For example, in [11] the authors developed a community detection method on multiplex and multiscale networks, while in [12] and [13] the authors introduced new models to represent an interconnected network of networks and a multidimensional sociality and extended classical measures to the multidimensional case. Finally, in [14] the authors studied correlations and overlaps among different kinds of networks by analyzing the social networks of a massive multiplayer online game.

III. Online and offline dataset

A. Client-server application

The data acquisition about encounters and the Facebook friendship graph is performed by means of a simple Client-Server application. The design and development of the required components were assigned to a class of undergraduate students in the Computer Science program at the University of Milano. During the experiment each student used the desktop Client to record and manage his/her daily encounters in the personal storage (student-server) along wiht his/her own list of Facebook friends. All this information can be extracted by means of a dedicated Facebook application accessing Facebook API. Data format was checked during the insertion operation. At the end of the project, all personal records in each student-server were automatically collected into the main Server, where they were merged in order to build the social graph of the experiment.

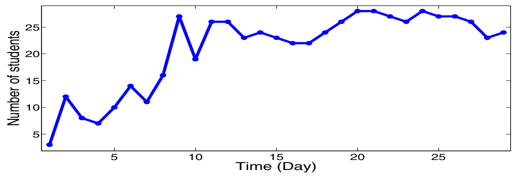


Fig. 1. The number of distinct participants, for each day, who reported at least one meeting.

Each encounter record reported by a student provides the following information:

- *Name and Surname* of the person met. The surname is optional because it might not be known.
- *Facebook name*: The value of the field *name* associated to the object *User* in the Facebook API. This information is optional.
- *Date* in MM-DD-YY format, so as to achieve a variable reporting interval and thereby avoid the problem of daily and persistent reporting [9].

B. Dataset description

To form the experiment team, we gathered more than 70 students from different courses and different years. After the project presentation, 35 out of 73 students volunteered. They were required to develop their own client according to specifications, as well as participate in the experiment. In this type of experiment, initial motivation is essential for obtaining rich and consistent datasets.

During the experiment's lifetime, the 35 students met 1,115 other people, while the corresponding total number of Facebook friends reached 10,291. As clear from these figures, the great majority of Facebook friends never met during the experiment.

The experiment lasted for four weeks straight, from December 13, 2011 to January 10, 2012, including both working and holiday days. By the end of the first week almost 25 students had completed the development of their application, so their reporting phase started before Christmas. In Figure 1, we report the number of students who recorded their daily contacts. We can observe that a stable condition is reached after about just one week. This is an indication of the fact that motivations remained intact throughout the experiment's lifetime, with no drop in the production of contact events.

We should say a little about the group we are investigating. The students who took part in our experiment represent a group with rather homogeneous behavior patterns. Nevertheless, as discusses in Section VI, their encounter structure is highly influenced by whether they are first, second, third, etc. year students. Despite that, all students are socially active, as shown by the fact that the average number of encounters is roughly 40. Meanwhile, each has an average of 311 Facebook friends. This mean figure, higher than the 190 reported by Facebook [1], generated some 10,000 sampled nodes, starting from a seed of 35 users.

C. Technical Issues and Limitations

A few technical issues about managing and cleaning up the experiment dataset deserve examination in greater detail.

The need to compare offline and online social networks advocates a policy to map the set of encounters of each person onto her/his Facebook ID. The Facebook policy can help us by stating: *"We require everyone to provide their real names, so you always know who you're connecting with."* [2], although some users, even in our dataset, ignore this advice. As a consequence, we mainly exploited the Facebook Graph API to get the user Facebook ID. This kind of request is based on public information (such as, the ID and the full name) and does not require any user's authorization.

Nonetheless, we had to deal with many different conditions as to the available data. Of course, when the encounter record contains the Facebook name, the mapping is simply obtained by querying the Facebook Graph. When the fields "'name"' and "'surname"' are used, the query might return namesakes. In this case we operate as follows: if one of the friend lists of the people met is public, we search the encounter name person and extract her/his ID; if both lists are private, we try to find the most likely profile leveraging the public information. When only the person's name is available, we do not perform any mapping (5% of the nodes).

Errors might arise because students happen to be unable to pay attention to details about daily encounters. To enable some statistical adaptation, we estimate the magnitude of these errors by evaluating the one-sidedness of the recorded offline friendship, i.e. when all the records of a relationship are registered by only one person. We calculate that bilateral relationships happen on 90% of links, accounting for the reliabilty of the experimenters.

Finally it is interesting to observe that the approach we adopted overcomes intrinsic limits of the methodology that only captures the encounters among Facebook friends. In fact, we also record the encounters between strangers and between familiar strangers.

IV. Definitions

In this section we provide some definitions to formally describe the two complex networks and the respective interleaving. As a matter of fact, the different layers used in the experiment, i.e. online and offline sociality, introduce a variety of nodes and, as a consequence, many types of edges. As for nodes, we have three sets: V_s, the students involved in the experiment; V_f the students and their Facebook friends; and V_c, the students and the people they meet.

[1] https://www.facebook.com/notes/facebook-data-team/anatomy-of-facebook/10150388519243859

[2] Facebookpage:https://www.facebook.com/help/?faq=112146705538576

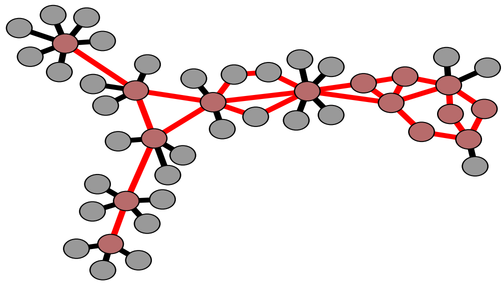


Fig. 2. Example of an inner graph. The red subgraph represents the graph of the nodes with degree greater than 1. The "'leaves"' of the graph are grey nodes.

Based on these node sets, we define the different objects we analyze and compare:

- We define the undirect $G_f = (V_f, E_f)$ as the *Facebook graph*, where E_f represents the link set retrieved from the student friend lists, i.e. $(u, v) \in V_s \times V_f$ belongs to E_f if u and v are Facebook friends .
- We define the *contact graph* $G_c = (V_c, E_c)$, where E_c represents the link set retrieved from the contact record of the students. Specifically, $(u, v) \in V_s \times V_c$ belongs to E_c if u and v experience at least one encounter during the experiment.
- We extend the contact graph G_c to the *weighted contact graph* $W_c = (V_c, E_c, w_c)$ by adding a weight function $w_c : E_c \to \mathbb{R}$. The function $w_c((u, v))$ assigns to each edge (u, v) the number of contacts between u and v.
- Let finally be $W_{fc} = (V_c \cup V_f, E_{fc} = (E_c \cup E_f), \phi, w_{fc})$ the *merged graph*. The link labeling function $\phi : E_{fc} \to \{0, 1, 2\}$ is defined as

$$\phi((u,v)) = \begin{cases} 0 & (u,v) \in E_f - E_c \\ 1 & (u,v) \in E_c - E_f \\ 2 & (u,v) \in E_f \cap E_c \end{cases} \quad (1)$$

that is, ϕ indicates if two nodes have a relationship only on Facebook, only in real life, or both. While w_{fc} is defined as

$$w_{fc}((u,v)) = \begin{cases} 1 & \phi((u,v)) = 0 \\ w_c((u,v)) & otherwise \end{cases} \quad (2)$$

It is to note that the w_{fc} definition depends on the dataset we analyze, in particular, as we do not have information on the Facebook link weights, we assign the contact weights only when possible.

Besides, we introduce the notion of *inner graph* $I_{V_1}(G)$ of a graph G as the subgraph induced by the set of nodes V_1 with degree greater than 1, deleting peripheral nodes, i.e. the leaves of the graph. This is shown in the Fig.2.

V. Overlapping degree of the networks

As a first step in comparing the encounter and the online social network structures, we analyze and confront student neighborhoods on Facebook and the contact graphs (see Fig.3 and 4). The graph G_f, shown in Fig.3(a), is made up of 10,326 nodes and 10,864 edges. The weighted contact graph W_c is made up of 1,150 nodes and 1,201 edges. It is visualized in Fig.3(b), where the thickness of an edge is proportional to its weight.

We first measure how many Facebook friends a person has met during the course of the experiment. Results indicate that on average only 4% of the Facebook friends were met and, apart from some nodes, percentages oscillate between 0% and 10%.

So far we have considered the direction from Facebook to offline life. Now we take into account the opposite direction. We examine the people involved in the encounters, look at factors such as how many have no Facebook account, how many are on Facebook but are not friends with one another and how many are on the social network. As to the first point, we find that the average number of people met who were not on Facebook is 18. In particular we observe that for a third of the students 50% of the meetings involve people not on Facebook. As for the second variable, we discover 75 people met with a Facebook account but not Facebook friend with the encountered. For the last quantity we find that on average 45% of contacts involved Facebook friends.

An important measure used in many friend recommendation algorithms is the number of common neighbors (overlap) between two nodes. This type of property represents a similarity measure for nodes. In fact the higher the overlap the more the nodes share the same interests and the same features. Results are in line with previous behaviors, in particular in the Facebook graph we find 411 common neighbors, but only 15% of them (54) were met during the experiment. This fact has a big impact on the common neighbor relevance, as it sheds light on its value as a similarity measure. In particular this observation makes us wonder about its effectiveness in cases where common neighbors measure is employed in real life recommendations.

VI. Structural analysis

The high number of nodes derives from the multiple star structures associated to each node. They are due to the design of the experiment, which concerns the description of the network of a group. The stars are composed by nodes in the ego-network of someone participating in the experiment who is not known by any other participant. As for some metrics, only the network of student nodes and their overlaps are interesting. We visualize the inner graphs in Fig.4(a) and 4(b). All students are present in the inner graphs since they all have degree greater than one. Obviously the number of nodes is considerably lower, 446 and 65 respectively, while the links number 1,153 and 116. It is interesting to see, in Fig.4(a) and 4(b), the number of persons who share a relationship with more than one participant in the experiment. While on Facebook they number 411, in real life they are only 30 of them.

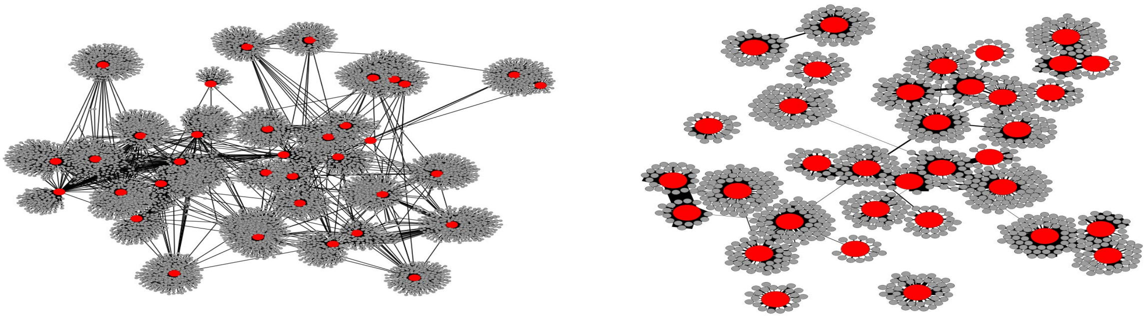
A. Connected components

In Fig.3(b) we explore the structural properties of the Facebook graph G_f of the classroom. As we can see, there is only a giant component. In other words, each node pair is connected. While this observation may seem trivial it is actually not, for the experimental environment is a quite heterogeneous one, consists of students of different years, and the network has a very low density equal to 0.012. As we indicate in the following, this property is due to the presence of a few nodes that act as a bridge between different groups in the class. Analogously, in Fig.3(a) we explore the structural properties of the contact graph G_c of the classroom. We can promptly note that G_c is not connected; there are, in fact, 6 components. This produces a less connected scenario in comparison to the Facebook one. The giant component is composed by 914 nodes and characterized by a low density (0.014). We must underscore that the remaining components contain eight students, forming groups marginal to the class.

B. Degree centrality

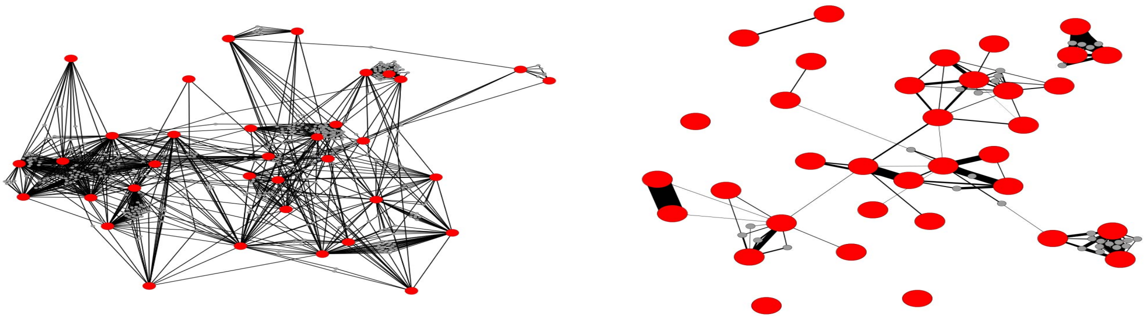
The simplest centrality measure in a graph is the degree. We take into account two kinds of degree depending on the network we analyze. The first type, which we call *total degree*, is the usual definition applied to graph $G_.$, while the second, known as *inner degree* is computed on the corresponding $I_{V_1}(G_.)$. We compute the above quantity both on G_f and on the unweighted contact graph G_c and on its weighted extension W_c. Obviously, in W_c we apply the strength of nodes. The last metric allows us to measure the popularity of a person not only by the number of friends s/he has but also on the basis of how often s/he meets with them.

1) Facebook: Observing the Fig.5(a), relevant to the Facebook graph, we obtain different behaviors involving the same nodes. In the figure the size and the color of the nodes are respectively proportional to the inner and total degree. For example, node 18 has the higher total degree (787) while its inner (44) is lower w.r.t. the other nodes. To quantify the agreement between Facebook importance and classroom importance we perform a rank correlation analysis. Rank correlation analysis allows us to test if the ranking induced by the different degrees is similar or not. As a rank correlation method, we compute the Spearman's rank correlation coefficient ρ on the ranking induced respectively by total and inner degree on V_s. We obtain $\rho = 0.4$, which indicates that the two degree measures induce different rankings. So, some nodes, relevant for example in G_f, lose their importance in the relative $I_{V_1}(G_f)$. An explanation of these changes is rooted in the numbers of common neighbors in the induced subgraph $I_{V_1}(G_f)$. In fact, nodes with a high total degree and a small inner degree have few neighbors and share few connections with other nodes in the subgraph. Generally, the above results suggest that Facebook popularity is not uniform among groups a person belongs to and so people with a



(a) (b)

Fig. 3. Fig.3(a) and 3(b) respectively represent the Facebook graph and the weighted contact graph, where link size is proportional to its weight. In all the graphs, red nodes represent the students.



(a) (b)

Fig. 4. Fig.4(a) and 4(b) are the corresponding inner graphs $I_{V_1}(G_f)$ and $I_{V_1}(W_c)$ of 3(a) and 3(b). In all the graphs, red nodes represent the students.

high overall importance may not be popular in a specific community. We find that students have an average degree equal to 312 and a 0.8-quantile equal to 447. As for the induced subgraph $I_{V_1}(G_f)$, the average student degree is 35, while the 0.8-quantile corresponds to 52.

2) Encounter: We analyze the degree distribution of G_c and W_c to highlight the number of people met and the number of contacts per person. In particular we focus only on the degree of the students, since for the other nodes we have incomplete information. The degree results are presented in Fig.5(c). In the figure the node color is proportional to its degree computed on W_c, while the size is proportional to the one computed on G_c. On average the number of people met by each participant is 37 and the average number of encounters is 125. As suggested by the figure and by the Spearman coefficient $\rho = 0.6$, a clear relation between the degree and its weighted version does not exist; actually, there are many nodes having a high degree yet a color that indicated a medium-low weighted one. This explains why maintaining many close friendships proves difficult.

We also compare the different degrees nodes have in Facebook and encounter networks. By analyzing the Spearman coefficient matrix, we find quite heterogeneous results. For example, the Facebook total degree quite positively correlates with the inner degree in the contact graph, while, at the same time, it has no correlation with the total degree in the contact weighted graph. Generally, we have shown that the centrality measure given by the degree does not maintain the rank, so that popularity in Facebook does not correspond to the same popularity in the encounter networks.

C. Eigenvector centrality

We calculate the eigenvector centrality defined by $x' = \lambda_1^{-1} \sum_j a_{i,j} x_j$ where A is the adjacency matrix of the graph and λ_1 is the largest eigenvalue of A. The eigenvector centrality relates the node importance to the importance of its neighbors; in particular it may be large either because a vertex has many neighbors or because it has important ones.

1) Facebook: In Fig.5(b) we can see this effect at the bottom right of the graph; in fact, node 2 gains its centrality from its numerous neighbors and conversely spreads its value among them. In this respect, comparing Fig.5(a) and 5(b), we can see that the degree centrality is different from the eigenvector centrality: specifically node 17 has a high degree yet is connected to nodes low in importance.

2) Encounter: As to this measure, we calculate the eigenvector centrality considering $I_{V_1}(G_c)$ and $I_{V_1}(W_c)$ on each component of the relative graph. In particular, in the weighted case we apply the general centrality proposed in [15] which still corresponds to the leading eigenvector of the adjacency matrix, with matrix elements being equal to the edge weights.

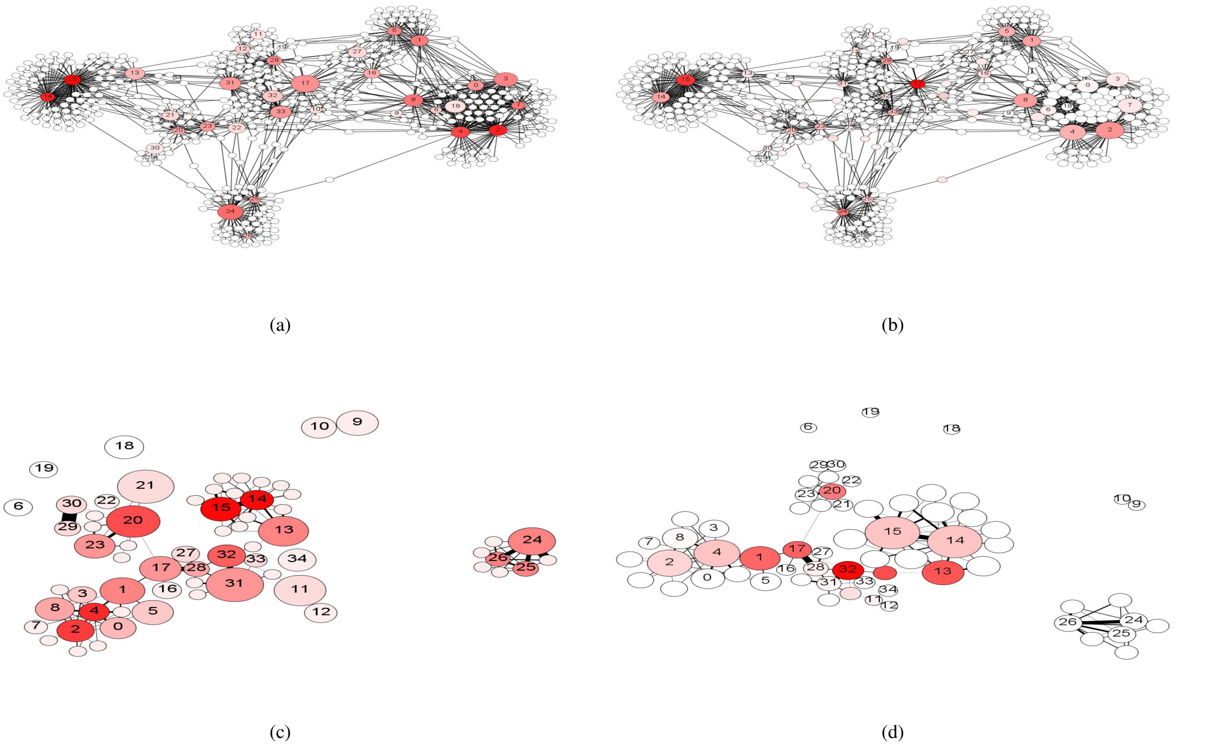


Fig. 5. Fig.5(a) Facebook graph: size and color (from white to red) nodes are respectively proportional to the total and the inner degrees. Fig.5(b) Eigenvector and betweenness centralities: size and color nodes are proportional to their eigenvector centrality and betweenness centrality. Fig.5(c) Unweighted contact graph: size and color (from white to red) nodes are respectively proportional to the total and the inner degrees. Fig.5(d) Eigenvector and betweenness centralities: size and color nodes are proportional to their eigenvector centrality and betweenness centrality.

The meaning of this measure is quite similar to the one in a citation network. In fact, if we use the frequency encounters as link weights, eigenvector centrality would then give people high ranks in either of two cases: when they are met by many others and if they meet frequently with a few others. Weights play a fundamental role in comparing the ranking induced by the two measures; in fact, analyzing only the two most numerous components, we find opposite results. In one case we observe a strong monotone increasing relation between the weighted and the unweighted centrality ($\rho = 0.8$), while in the other we observe a substantial lack of correlation between the variables ($\rho = -0.3$). These results depend on the distribution of the weights: in one case the highest weights are among central nodes, in the other case the opposite is true.

If we consider both Facebook and encounter networks, we find results which shows a substantial lack of correlation among the eigenvector centralities of the student nodes computed on the different graphs. In fact for each pair of centralities involving the Facebook and the contact graphs, we obtain correlation values near to zero. Also in this case, these findings claim the observation that eigenvector centrality is not linearly transferred across layers.

D. Betweenness centrality

A different concept of centrality is betweenness centrality. It captures the extent to which a node is on paths between other nodes. We may formally define the betweenness of the node i as

$$b_i = \sum_{s,t \in V} n^i_{st} / g_{st}$$

where n^i_{st} is the number of shortest path from s to t passing through i and g_{st} is the total number of shortest paths from s to t. The betweenness measures the amount of information passing through each vertex, if it follows the shortest path. Therefore, nodes with high betweenness may have a high influence due to a sort of control over the information passing among nodes.

1) Facebook: Betweenness values are depicted in Fig.5(b) where the node dimension is proportional to them. As expected, the betweenness values are different from the other centralities. In particular, node 17 gains the maximum betweenness in that it acts as a bridge among the different areas of the graph [3].

[3]Closeness centrality shows values similar to the betweenness centrality

2) Encounter: In the Fig.6(a) we report the values of the betweenness centrality computed on the simple and weighted induced subgraph, where the weights are equal to the inverse of the value returned by the w_c up to a scaling factor which flavors paths passing through strong links. Comparing the relative values, we can observe how the introduction of the weights changes some node centralities. In particular, weights enhance the probability that information passes through some paths. For example, if we consider an unweighted graph and two minimum paths between two nodes, the probability that a message follows one of them is an even split. In the weighted case, the path might be unique concentrating all the probability in it. We can observe this phenomenon in node 14 values where in the unweighted case the betweenness is distributed between 15 and 14, while strength force paths to pass through 14. As reasonable to expect, weight introductions not only changes the betweenness values but also the ranking the student nodes. In fact, Spearman coefficients measured on each pair of betweenness types show uncorrelation between the different centralities. Therefore, even on the social dimension (offline sociality) a node can assume different relevance depending on the features of the network we consider.

In comparing the two social layers by the Spearman coefficient matrix, we find results in accordance with the one presented in the above paragraph. In fact betweenness centrality values on the student set are almost uncorrelated. This further corroborates the fact that betweenness centrality does not transfer monotonically too.

E. Small world properties

To see if our network presents a small-world phenomenon, we analyze the average clustering coefficient C and the average path length L. L is the number of hops in the shortest path averaged over the pairs of nodes, while C is the average of C_v. C_v is defined as the fraction of edges that exists between all the possible links connecting the neighbors of v. We have a small-world situation if L is similar to L_{rand} (characteristic shortest path of a random graph with n nodes and average degree k equal to the real one) and $C >> C_{rand}$. We perform the following computation only on the induced subgraphs because the star structures in the corresponding graphs would artificially decrease the average clustering coefficient.

1) Facebook: Comparing the above computed quantities, we can see that our network is a small-world one as $L = 3.55$ and $L_{rand} \approx ln(n)/ln(k) = 3.65$ while $C = 0.73$, which is much greater than $C_{rand} \approx k/n = 0.012$. In general the Facebook network contains nodes that are high clustered and a few shortcuts that reduce the distance between the nodes. In the Fig.5 the role of shortcut is played by the central area around the node 17, which links link the different high clustered groups. In fact, as many paths pass through the center of the graph the most likely distance is 3.

2) Encounter: As for small-world properties, the induced subgraph is quite cryptic to classify. In fact the average clustering coefficient $C = 0.764$ is greater than the expected one in a random version, i.e. $C_{rand} = 0.053$, while the average path length $L = 4.03$ is greater than $L_{rand} = 3.33$ and the diameter is equal to 7. So, as shown in Fig.5(d) , the structure presents highly clustered regions (explaining the high C) connected through few links (explaining the path features). In particular, we can observe a sort of backbone, comprised naturally by the nodes with high betweenness centrality.

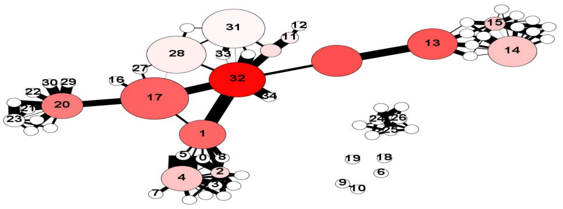
VII. Merging the complex networks

In this Section we analyze the merge graph W_{fc}. Our main goal is to blend the two social layers in a unique network and check if student nodes, in this merged scenario, maintain their centralities or the merging modify the ranking among nodes. In the following analysis we adopted the same methodology applied in Section VI. In particular we compare total and inner degrees, eigenvector centrality and betweenness centrality on the graphs G_f,G_c,W_C,W_{fc},G_{fc} (unweighted version of the merge graph) and their induced subgraphs. In general G_f influences the most of the measured centralities on the merge graph because of its denser and more compact structure. That happens despite the weight function flavors links that correspond to encounters.

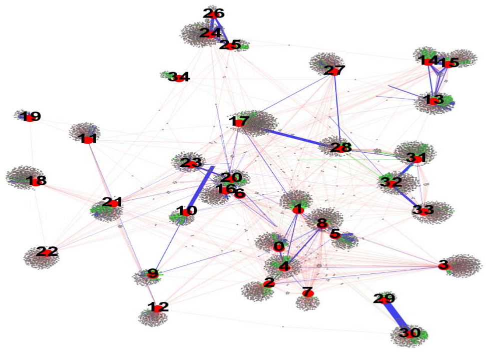
As for inner and total degrees, we find similar results so we only present the one on the second variable. First we observe a strong correlation between the total degree on G_{fc} and G_f as the number of Facebook friends is much higher than the encounter one. Otherwise the total degree on W_{fc} correlates with the total degree on W_c and G_c as the bias introduced by w_{fc}.

The eigenvector centrality measured on W_{fc} and G_{fc} has a particular meaning as it mixes the contribution of the degree and the connectivity of the two sociality layers. Furthermore in the weighted case it depends on the attitude a nodes has to connect with other important nodes through the strong link given by the contacts. By analyzing results we obtain that the eigenvector centrality on $I_{V_1}(W_{fc})$ positively correlates with that on $I_{V_1}(G_f)$. An interesting result concerns the centrality on W_c, in fact it correlates with G_f and unexpectedly negatively correlates with W_c. This shows that the Facebook connectivity differently allocates the centrality portion given by the strong links. We observe the same effect in the betweenness centrality where in G_{fc} and W_{fc}, it correlates with the betweenness measured on G_f.

As the strong influence of the Facebook graph we investigate how weights act on the different centralities of student nodes only on the Facebook relationships. We compare F_c and F_w, i.e. the induced subgraph of W_{fc} containing links with ϕ values equal to 1 or 3. By comparing the Spearman coefficient of each centrality, we find a strong correlation between the total weighted and unweighted degree. This fact can be explain by the low degree of overlapping between contacts and Facebook friends seen in Section V. An opposite result has been observed on eigenvector centrality. For this measure we observe a low negative correlation ($\rho = -0.33$). So the weight insertion drastically changes the importance a person acquires in the network. Last we consider the betweenness centrality



(a)



(b)

Fig. 6. Fig.6(a) Betweenness centralities measured on $I_{V_1}(G_c)$ and $I_{V_1}(W_c)$. Size and color nodes are proportional to their weighted and unweighted betweenness centrality. Fig.6(b) Merge graph W_{fc}: link color indicates the value of the link labeling function ϕ: 1 (blue), 2 (green) or 3 (red). The link size is proportional to its weight assigned by the function w_{fc}.

on G_f and W_f. In this case we obtain a ρ value equal to 0.6. This finding implies that the ranking does not drastically change, although, also in this case, some ranks consistently increase or decrease.

In general we observe that the merging of the two social networks induces a different ranking on the set of the student nodes and that the Facebook structure and the weights inferred from play a fundamental role in making the centralities always different.

VIII. Conclusion

In this paper we describe an experiment that explores the intimate relation between online and offline sociality of a group of 35 students. The two layers of the group's sociality were described through the associated complex networks. Our work shows that the overlapping degree is low as the sets of Facebook and real contacts are quite different. Secondly the paper shows that node centrality is not a universal feature so that the people's popularity is most likely to change in different networks. Finally, the paper introduces a unified complex network which allows us to merge offline and online relevant features shedding a light on how human behavior is interwoven across layers.

IX. Acknowledgement

This work was partially funded by the Italian Ministry for Instruction, University and Research under the PRIN PEOPLENET (2009BZM837) Project.

References

[1] T. Henderson, D. Kotz, and I. Abyzov, "The changing usage of a mature campus-wide wireless network," in *Proceedings of the 10th annual international conference on Mobile computing and networking*, ser. MobiCom '04, 2004.

[2] A. Mtibaa, A. Chaintreau, J. LeBrun, E. Oliver, A.-K. Pietilainen, and C. Diot, "Are you moved by your social network application?" in *Proceedings of the first workshop on Online social networks*, ser. WOSN '08, 2008.

[3] P. Hui, A. Chaintreau, J. Scott, R. Gass, J. Crowcroft, and C. Diot, "Pocket switched networks and human mobility in conference environments," in *Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking*, ser. WDTN '05, 2005.

[4] N. Eagle, A. Pentland, and D. Lazer, "Inferring friendship network structure by using mobile phone data," *Proceedings of The National Academy of Sciences*, vol. 106, 2009.

[5] S. Gaito, E. Pagani, and G. P. Rossi, "Strangers help friends to communicate in opportunistic networks," *Computer Networks*, vol. 55, no. 2, pp. 374–385, 2011.

[6] M. Zignani, S. Gaito, and G. Rossi, "Extracting human mobility and social behavior from location-aware traces," *Wireless Communications and Mobile Computing*, 2012, article in Press.

[7] G. Bigwood, D. Rehunathan, M. Bateman, T. Henderson, and S. Bhatti, "Exploiting self-reported social networks for routing in ubiquitous computing environments," in *Networking and Communications, 2008. WIMOB '08. IEEE International Conference on Wireless and Mobile Computing,*, 2008.

[8] A.-K. Pietiläinen, E. Oliver, J. LeBrun, G. Varghese, and C. Diot, "Mobiclique: middleware for mobile social networking," in *Proceedings of the 2nd ACM workshop on Online social networks*, ser. WOSN '09, 2009.

[9] T. Hossmann, T. Spyropoulos, and F. Legendre, "Putting contacts into context: Mobility modeling beyond inter-contact times," in *Twelfth ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc 11)*. Paris, France: ACM, May 2011.

[10] M. Szomszor, C. Cattuto, W. Van den Broeck, A. Barrat, and H. Alani, "Semantics, sensors, and the social web: The live social semantics experiments," in *The Semantic Web: Research and Applications*, ser. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2010, vol. 6089.

[11] P. J. Mucha, T. Richardson, K. Macon, M. A. Porter, and J.-P. Onnela, "Community structure in time-dependent, multiscale, and multiplex networks," *Science*, vol. 328, no. 5980, pp. 876–878, 2010.

[12] M. Berlingerio, M. Coscia, F. Giannotti, A. Monreale, and D. Pedreschi, "Foundations of multidimensional network analysis," in *Advances in Social Networks Analysis and Mining (ASONAM), 2011 International Conference on*, 2011.

[13] M. Magnani and L. Rossi, "The ml-model for multi-layer social networks," in *Advances in Social Networks Analysis and Mining (ASONAM), 2011 International Conference on*, 2011.

[14] M. Szell, R. Lambiotte, and S. Thurner, "Multirelational organization of large-scale social networks in an online world," *Proceedings of the National Academy of Sciences*, vol. 107, no. 31, 2010.

[15] N. Perra and S. Fortunato, "Spectral centrality measures in complex networks," *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, vol. 78, no. 3, pp. 036 107+, 2008.