# Ego-net Digger: a New Way to Study Ego Networks in Online Social Networks


Massimiliano La Gala
IIT-CNR
via G. Moruzzi,1
Pisa (PI), Italy
m.lagala@iit.cnr.it

Valerio Arnaboldi
IIT-CNR
via G. Moruzzi,1
Pisa (PI), Italy
v.arnaboldi@iit.cnr.it

Andrea Passarella
IIT-CNR
via G. Moruzzi,1
Pisa (PI), Italy
a.passarella@iit.cnr.it

Marco Conti
IIT-CNR
via G. Moruzzi,1
Pisa (PI), Italy
marco.conti@iit.cnr.it



## ABSTRACT

The vast proliferation of Online Social Networks (OSN) is generating many new ways to interact and create social relationships with others. While substantial results have been obtained in anthropology literature describing the properties of human social networks, a clear understanding of the properties of social networks built using OSN is still to be achieved. The presence of a huge amount of records containing users' communication history, provided by OSN, represents a new opportunity to analyse and better understand the human social behaviour. In this paper we present *ego-net digger*, a novel Facebook application for the analysis of ego networks in OSN. Ego-net digger collects users' social data and gives a representation of their personal social networks according to the Dunbar's circles model. To show the potential of our application we analyse a sample data set collected during a data acquisition campaign, finding interesting similarities between the structure of Facebook ego networks and the properties found in the anthropology literature. Specifically, we find that, in our sample, there is a clear evidence of the presence of the same ego network structure - i.e., the Dunbar's circles - as found in human social networks formed offline.



## Categories and Subject Descriptors

H.3.4 [**Information Systems**]: Systems and Software—*Social Networks*

## General Terms

MEASUREMENT





## Keywords

Online Social Networks, Ego networks, Tie Strength


## 1. INTRODUCTION

The worldwide proliferation of online social networks (hereinafter OSN) is rapidly introducing plenty of new means to create and maintain social relationships with others. Although these new ways to communicate are becoming part of our everyday life, we don't have yet a complete view on how they are impacting on human behaviour in the actual society, both in the *physical* (real) and in the *cyber* (virtual) worlds. Human social behaviour is commonly studied using a model for the representation of personal social networks, called ego network - i.e., a social network formed of an individual (ego) and the people with whom ego is in contact (alters).

While the properties of ego networks in the real world have been deeply studied in the anthropology and sociological literature, OSN ego networks are not yet completely understood. Specifically, there is a lack of knowledge regarding the structure and the dimension of ego networks in the virtual world. In addition, the fundamental differences between the properties of OSN ego networks and the well known results about social networks formed in the real world (hereafter referred to as human ego networks) are still under investigation.

In this paper we focus on the characterisation of OSN ego networks in terms of their structure. In particular we want to assess if they show the same structure found in human ego networks. Human ego networks presents a hierarchical structure formed of a series of four concentric circles of alters (called Dunbar's circles) characterised by similar properties and typical proportions - namely, the ratio between the adjacent circles obtained from many different studies is almost constant and close to three. These circles are determined by the frequency of contact between ego and alters and their size increments as this frequency decrease. Further details on the structure of human ego networks are given in Section 2.1.

The study of these social aspects is important to place the first building blocks for a better understanding of the characteristics of OSN, which are already an important part of our everyday life. Moreover, a complete knowledge of the differences between the structures of personal social networks in OSN and in the real life could be exploited to infer and analyse properties of human social relationships using OSN data. For example, leveraging the communication records in OSN, we could be able to obtain precise information regarding weak social ties, which are difficult to be collected in real life, and better study their role within our society.

In this work we describe our contribution aimed at bridging the gap in the understanding of the differences between the characteristics of human social relationships in the physical and the cyber worlds. Specifically, we introduce *ego-net digger*, a novel application for the study of OSN ego networks in Facebook. Ego-net digger has two purposes. The first aim of this application is to give to the users a graphical representation of the structure of their ego-networks, arranging their friends according to the Dunbar's circles model. The second aim of ego-net digger is to provide a tool to download social data from Facebook using public API. To have a complete view of users' ego networks, we also collect explicit evaluations of their social relationships in both the real world and in Facebook, asking them to rate their friendships using a dedicated user interface. With the collected data we are able to perform accurate analysis on the properties of Facebook ego networks, and to depict their structure as a graphical output for the user. Moreover, using ego-net digger, we aim to obtain enough social interaction data to build and validate a model for the estimation of tie strength using Facebook data. We think that this model should be sufficiently accurate, since the information we are collecting from Facebook is not limited to the mere indication of the existence of a contact between pairs of people, but also includes other variables, useful to estimate tie strength. For example, we collect the number of photos in which two people appear together or their degree of kinship. These observable quantities could reveal the intimacy between people involved in a social relationship in Facebook. In addition, other variables like the number of events in common or the number of likes on the same objects could indicate the social similarity between individuals.

In this work we also verify that the structure of the ego networks extracted by ego-net digger and presented to the user in the form of groups of friends having different levels of intimacy with her (following the Dunbar's circles principle) is similar to that found in human ego networks. To do so we analyse a small sample of social data collected by ego-net digger, as will be explained in greater details in Section 4. The preliminary results, even if they need to be validated on a much more larger data set, represent an interesting indication of the similarity of the structure of ego networks in OSN and in real life. Specifically, we find that the dimensions and the scaling factors of the Dunbar's circles in Facebook, obtained using a partitioning clustering algorithm applied to the frequency of contact between users, are compatible with those found in the physical world.

The rest of this paper is organised as follows. In Section 2 we give an overview of the existing work in literature regarding the properties of ego networks in real life and some of the preliminary results found in OSN. Then, in Section 3 we describe ego-net digger, our novel application for OSN data collection and analysis. In Section 4 we describe a preliminary analysis, based on a small data set sampled using ego-net digger, to support the correctness of the graphical output provided by our application based on Dunbar's circles concept and to show the similarities we have found between ego networks structure in Facebook and in real life. Hence, in Section 5 we draw the conclusions of our work, giving some highlights on our future work.

# 2. RELATED WORK

In this section we give an overview of the background work in literature on social ego networks analysis from two different standpoints. Firstly, we introduce the relevant findings in the anthropology literature regarding human ego networks. Then we summarise the recent results aimed at describing the properties of OSN to analyse social aspects of humans in virtual environment.

## 2.1 Properties of Human Ego Networks

Human personal social networks are modelled using ego networks. The key properties of this model is the tie strength, represented as the distance between the alter involved in a relationship and the ego. The tie strength is not directly observable, but can be estimated using several dimensions of a social relationship. In [6], the author gives an informal definition of tie strength, defining it as a (probably linear) combination of the amount of time, the emotional intensity, the intimacy and the reciprocal services which characterise the tie. In [9] authors demonstrated that *emotional closeness* or *intensity* of the relationship are the best indicators of tie strength. Other authors define the emotional closeness as a synonym of tie strength and use the frequency of contact to estimate it, since the two concepts appear to be tightly correlated [3, 7].

Studying human ego networks, anthropologists found that an individual can maintain only a limited number of *active* social links - i.e., relationships for which ego spends a non negligible amount of cognitive resources [2]. Maintaining a social relationship with an alter costs in terms of time spent in communicating and interacting with her, and cognitive resources in terms of remembering facts about the relationship on the alter. There is substantial evidence that the cognitive capacity devoted to this task are limited, and therefore the number of active relationships is also limited.

The portion of ego network containing active relationships is called *active network*. In human ego networks the average active network size is around 150, the so called "Dunbar's number". This number was found by the anthropologist Robin Dunbar, who has been the first to hypothesise the presence of this limit through a correlation analysis between primates' brain neocortex dimension and the size of their social groups [2]. After him, his hypothesis has been validated by different studies on humans [12, 7].

Human ego networks show a typical hierarchical structure formed of four concentric subgroups of alters arranged in a concentric sequence of circles (called "Dunbar's circles") defined by the frequency of contact between ego and alters. Alters placed in the different circles have an increasing level of frequency of contact from the outermost circle to the innermost one. The findings in the anthropology literature indicate that Dunbar's circles have typical proportions, with a scaling factor between adjacent circles close to three. In addition, each circle has a typical frequency of contact and the alters contained in it share similar properties. For example the innermost circle, called *support clique*, contains people contacted at least weekly, from whom ego seeks advice in case of severe emotional distress or financial disaster [10]. The support clique is limited, on average, to five members. The second circle, called *sympathy group* contains 15 people, contacted at least monthly. The last circle (active network) defines the boundary between active and inactive relationships and is limited to 150 people, contacted at least yearly. The third circle (called *affinity group*) is not yet completely characterised and no detailed properties have been found regarding its typical contact frequency and dimension, which is estimated to be around 50.

## 2.2 Studies on OSN Ego Networks

Since people have hundreds (or even thousands) of personal contacts inside OSN (e.g., 50% of the Facebook users have more than 99 friends [11]), having a clear view of their social relationships in virtual environment is becoming quite a complex task. Some recent applications give a graphical representation of the social links of the users relying on the frequency of contact to express the importance of relationships[1]. One of the most interesting feature of these applications is the ability to divide OSN friends into different communities, thanks to the analysis of the records of past communication between them. Although the presence of this interesting function, these applications are not able to provide any indications regarding the real tie strength between users. For example two colleagues using an OSN to communicate during their work could produce a lot of communication data over time, but their social tie strength could be lower than expected from the frequency of contact.

Some recent work found evidences of the presence of the Dunbar's number in Facebook and Twitter [1, 5], as a first proof of the similarity of the ego networks inside physical and cyber environments. Although these interesting findings, the relation between tie strength and the properties of ego networks is not yet completely understood. The design of measures of tie strength is still an open issue. In fact, the introduction of this concept has been carried out on an "intuitive" basis. Some research work has been done to give a more precise definition of tie strength using survey data [9]. Recently, some authors have made a first attempt to build a model for tie strength prediction using data coming from Facebook [4]. However, the proposed model uses a high number of variables (i.e., it has a high complexity since requires a lot of time to collect a large set of Facebook data) and has not been validated using an enough large data set. Using Ego-net digger we aim to come up with a large data set of Facebook social interactions with the following objectives in mind: (i) give a precise definition of tie strength using a combination of Facebook variables, validating the hypothesis in [6]; (ii) study the structure and other important properties of OSN ego networks and the relation between human behaviour in the virtual and the real worlds.

# 3. EGO-NET DIGGER APPLICATION

Ego-Net Digger is a web-based Facebook application able to retrieve and analyse social interaction data between the user and her friends, giving as output the ego network of the user partitioned into its Dunbar's circles. Ego-net digger is the result of the extensive work we performed on the basis of the know-how acquired during the creation of the prototype application described in [1]. However, ego net-digger is a much more advanced application, built to overcome some limitations of the cited prototype, summarised by the following points:

- The download of social data was only applicable to one user at a time
- The prototype did not support background data download, thus requiring the user to remain connected for the entire download process, otherwise the process would have ended before its completion
- The number of Facebook posts downloaded by the prototype (per each ego) was limited to 400/500, far less than the complete history accessible from Facebook communication records
- The manual evaluation of Facebook tie strength was long and tedious, not giving an easy method to rank friends and to visually compare them
- People were not incentivised to use the application because of the lack of a reward for their time spent

Ego-net digger is designed to overcome these limitations, allowing a much more refined analysis of the structure of ego networks. Specifically, ego-net digger is able to retrieve the entire communication history between the user and her alters in Facebook, includings posts, comments, likes, tags, events in common, private messages, notes, photos, status updates, video, family relationship and other information related to the profile of the user and her friends. This data clearly contains more than just the indication of contact between people, and can be leveraged to obtain a model for the estimation of tie strength much more refined than those based only on the frequency of contact between users.

We also provide a new method to collect explicit evaluations of tie strength from the users. This method is based on a simple but effective graphical interface that allows users to evaluate her friends dragging their pictures on a graphical ruler. All the pictures related to previously evaluated friends remain stick on the ruler, so that the user is able to easily compare her friends, producing a much more accurate evaluation compared to [1].

Ego-net digger introduces also the ability to obtain explicit evaluations of tie strength in the real life, in addition to those related to the Facebook world. These evaluations could be extremely useful for further characterisation of the differences between the physical and the real worlds.

In addition to the need to overcome the limitations of the prototype application [1], we developed ego-net digger with the following goals:

- give a graphical representation of the Dunbar's circles of users' ego networks through the analysis of Facebook data. This represents an immediate feedback to the user, motivating her to use the application and to share it with others
- support a wide range of different data analysis techniques, acting as a testbed for OSN analysis algorithms
- minimise the effort needed to update the application to cope with the OSN platform updates

[1]For example http://inmaps.linkedinlabs.com for LinkedIn and http://danielmclaren.com/node/77 for Facebook

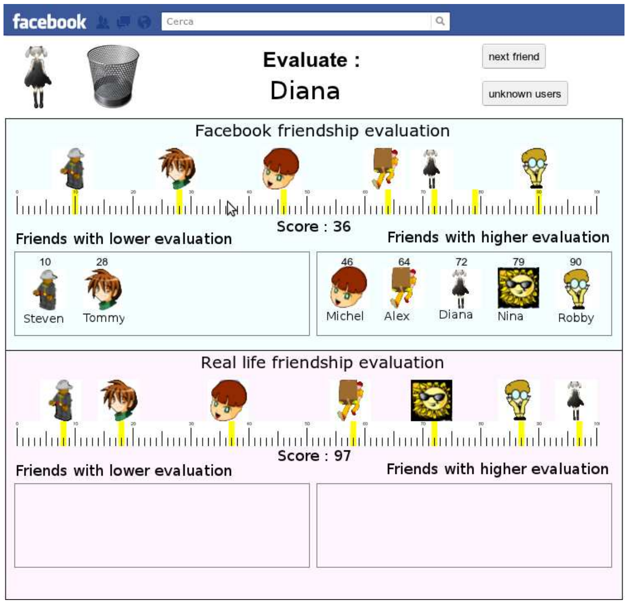


**Figure 1: Screenshot of ego-net digger tie strength evaluation module**

- support an high number of users, without the detriment of user experience
- make the manual evaluation of tie strength as fast and simple as possible to require a limited amount of time even for users with an high number of friends

Ego-Net Digger is composed of three main modules: (i) a data fetcher component; (ii) a web based user interface for tie strength evaluation and (iii) a graphical viewer that shows the Dunbar's circle representation of users' ego networks.

**Data fetcher module** is the server side component of the application. It retrieves all the interaction data related to the user and her friends, according to her privacy settings. This data can be then elaborated using one or more data analysis algorithms. Ego-net digger is designed to be easily extended and modified to be used with other OSN different than Facebook (e.g., Twitter).

The data fetcher module performs the data retrieval and elaboration phases in background, to allow the user to disconnect from the application without blocking the data retrieval procedure. The presence of this background data download procedure also ensures a better distribution of the application workload, since the concurrent downloads can be limited or delayed to avoid network congestion.

The actual implementation of ego-net digger data fetcher module uses Facebook Query Language (FQL) to access the social data of the users. We have spent a lot of effort to make the data fetcher module easily updatable and extendible to cope with Facebook API changes, building a tool for the automatic generation of FQL queries and table relying on the database structure, described on Facebook documentation[2].

**Web interface for tie strength evaluation** is a module used to collect manual evaluations of tie strength from the user connected to the application. We intend to use this data only during a preliminary phase of our work, to build and tune models for automatic emotional closeness estimation. With these models we would like to study the properties of ego networks and their relation with tie strength without requiring user interaction.

As experienced in [1], for the users is not easy to assign a numeric score to their friends without a graphical comparison between them. Moreover, the duration of the ranking process highly affects the accuracy of the evaluation. Therefore is important to give an easy to use user interface for the evaluations. Using ego-net digger, a user can rank a friend clicking on a graphical "ruler", graduated from 0 to 100. After doing so, the Facebook picture of the evaluated friend appears on the ruler at the position indicating the given score. In this way the user can easily compare her friends and change previously assigned scores in case she wants to refine the evaluation. To facilitate the visualisation of previously evaluated friends in case they are too many to be displayed on the ruler, ego-net digger shows the pictures of five people with higher/lower tie strength with respect to the current position of the mouse cursor (on the ruler) in two separated panels placed beneath the ruler. Figure 1 depicts a screenshot of the web interface for tie strength evaluation.

For each friend, the user is asked to express two different evaluations: the first one concerning the tie strength she feels with her friends in the "real life" and the second one for the tie strength in Facebook. We collect these different evaluations to analyse the differences between the users' active networks in real and virtual environments. Further details on this evaluation are given in Section 4.

From previous work [1] we know that the amount of inactive relationships which receive a tie strength evaluation equal to zero is rather substantial. To speed up the evaluation process without biasing the results, we introduce a button by which the user can declare a friend as a mere acquaintance for both the physical and the cyber worlds.

Since we want to be able to check the goodness of the evaluations, collected by ego-net digger, we introduce some additional information to track the behaviour of the user. Specifically, we collect data regarding the duration of the entire evaluation process of a user and the timestamp related to each single score given. In this way we are able to study the distribution of the speed of the evaluation, performed by the user. This analysis allows us to identify if and how the evaluation changes over time and to detect not enough accurate evaluations in order to remove outliers from the collected data.

**Dunbar's circles viewer** is a web-based interface that displays to the user her Facebook active ego network divided into four different groups, according to the Dunbar's circles concept (see Section 2 for further details). This module uses the social interaction data retrieved by the data fetcher module to calculate the frequency of contact between the user and her online friends. Using the frequency of interactions is a first-order model to estimate the strength of social ties. We will refine this through the model we are developing, based on all the interaction variables between users (besides the frequency of contacts) Hence, the module applies a clustering techniques on the frequency of contact to split user's friends into 4 different groups (Figure 2 depicts a screenshot of an example output of ego-net-digger Dunbar's

[2]https://developers.facebook.com/docs/reference/fql/

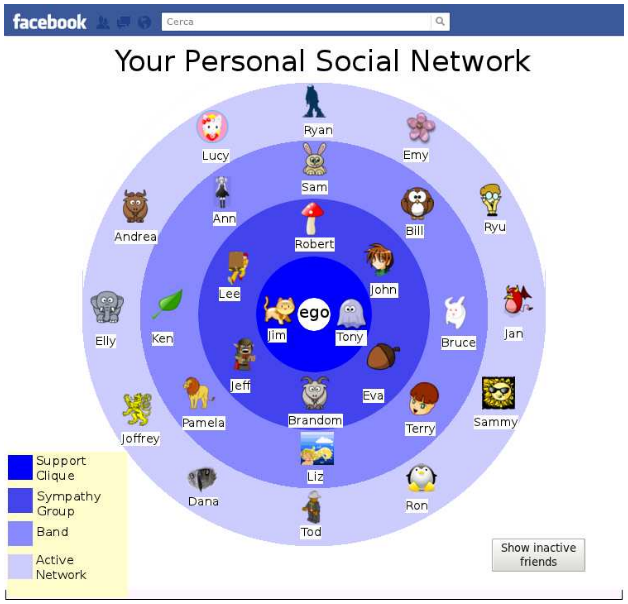


**Figure 2: Screenshot of ego-net digger Dunbar's circles viewer**

circles viewer). Further details regarding the clustering algorithm used by ego-net digger are provided in Section 4.

# 4. PRELIMINARY ANALYSIS ON FACEBOOK EGO NETWORKS

To validate the effectiveness of Dunbar's circles visualisation provided by ego-net digger and to show the potential of our application we make a preliminary data analysis on a small sample data set. This data set is composed of social data regarding 27 people from our research centre. For each participant, we downloaded all the social data related to the communication between them and all their friends (making sure to take care of Facebook privacy policies, and anonymising the data). We also asked the users to explicitly evaluate the strength of their relationships with their friends, both in the real and virtual worlds. With this data we perform a first analysis focused on the study of ego networks structure following the concept of Dunbar's circles. Then we make a second analysis aimed at comparing the tie strength evaluations in real and virtual worlds.

## 4.1 Analysis of Ego Network Structure

The structure of Dunbar's circles in human ego networks is based on the frequency of contact between ego and alters. Hence, to compare our results with the findings in the anthropology literature, we firstly extract the frequency of contact from the social interaction traces collected using ego-net digger. Although the frequency of contact alone is not enough to fully describe the social tie strength [9], we use it as a first order estimate of the intensity of social relationship between ego and her alters, maintaining thus an analogy with the results in the anthropology literature. This allows us to compare our results we the findings in human ego networks. We leave for future work the analysis of the estimation of tie strength using other indicators of social intimacy.

All the distributions of the Facebook variables in the sample data set concerning the communication made by ego to her alters show a long tail shape, as can be seen in Figure 3, depicting the complementary cumulative distribution function (CCDF) of the number of likes sent by ego to alters, calculated for each ego network, then averaged for all the networks. This result indicates the presence of a small set of alters with whom ego has a high frequency of contact, and a larger set of friends with low frequency of contact. This is in accordance with the findings in [1, 7] and with the idea behind Dunbar's circles [10].

We distinguish between outgoing and incoming communication between ego and alters modelling each social relationship as two different directional links. Therefore we assign a weight to each link equal to the frequency of outgoing contact as regards the links directed from ego to alters and equal to the frequency of incoming communication for the opposite direction. The frequency of outgoing communication represents how frequently ego contacts each of her friends. In this analysis we consider it as the total number of direct interactions sent by ego to the alter divided by the duration of the considered social relationship. Since Facebook API does not provide a method to retrieve the time at which two people became friends, we estimate the duration of a social link using the number of days elapsed since the first communication occurred between the involved users. This is a standard way to estimate the duration of social relationships in OSN [4]. On the other hand, the frequency of incoming communication expresses how frequently each alter contacts ego and is calculated as the number of direct interactions received by ego from the alter divided by the duration of the social link, calculated as done previously.

Intuitively, we expect the frequency of outgoing communication to be a good indicator of the amount of cognitive resources spent by ego to maintain each social relationships with others. On the contrary, the incoming communication - as we will show later - only partially explains this concept, since many alters are not reciprocated by ego. These

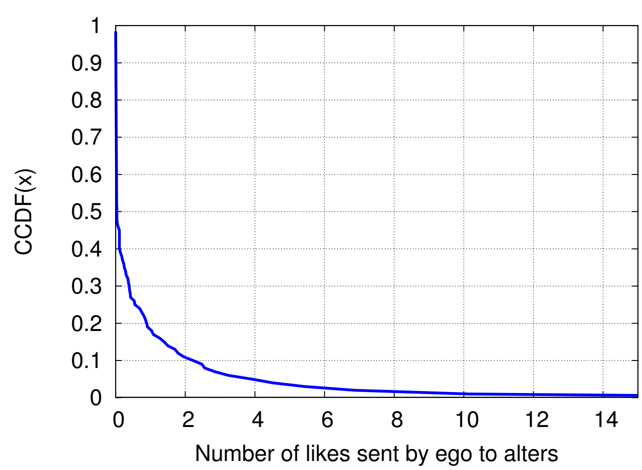


**Figure 3: CCDF of the number of likes sent by ego to alters**

alters represent a set of individuals for whom ego does not spend efforts to maintain the relationship active. Hence, using the outgoing frequency of contact we should be able to characterise the structure of our ego networks, maintaining an analogy with the results in the anthropology literature. This approach has been taken also in [5], where the authors consider as active friends all the users to whom ego sends direct replies to tweets.

We first consider the ego networks identified by using the frequency of outgoing communications as the strength of social ties. After removing messages that Facebook implicitly strongly encourages to send (e.g., wishes for the birthday), we consider as *active* the alters contacted at least yearly by ego. We find that the average active network size of the ego networks in our sample is equal to 77.8, over a total size of Facebook network (also considering inactive relationships) of 341.64[3]

If we consider the frequency of incoming communication and we define as active all the relationships related to alters who contact ego at least yearly, we find that the average active network size is equal to 142.04. Moreover, this active network does not completely contain the active network based on the outgoing communication. Indeed, only 66.52% of the active network based on outgoing communication is contained in the active network of the incoming communication. 63.55% of the social links in the active network built on the incoming communication turn out to be inactive in the other model. The discrepancy between the active network size found using the two different frequencies of contact indicates that these two quantities are not always symmetric. Hence, we decided to base our analysis on the active network built on the outgoing communication, to be sure to analyse the limits of ego in terms of her cognitive resources, according to Dunbar's studies. In addition, we decided to use the incoming frequency of contact to strengthen all the relationships that show reciprocity in the communication (i.e., the frequencies of contact in the two directions are both greater than zero). In this way we want to capture the resources ego spends to read the messages received from alters she actively contacts. To include this information in our analysis we defined the following index:

$$AdjFreq_{jk} = f_{jk} + \frac{f_{jk} * f_{kj}}{f_{jk} + f_{kj}} \tag{1}$$

Where $f_{jk}$ is the frequency of outgoing communication from ego $j$ to alter $k$ and $f_{kj}$ is the frequency of incoming communication from alter $k$ to ego $j$. The increment given by the additional term is maximised when the incoming and the outgoing interactions are balanced. The CCDF of the frequency of outgoing communication, of the frequency of incoming communication and of the $AdjFreq$ index are depicted in figure 4. The shape of these CCDF is similar to that presented in figure 3

To check whether ego networks can be divided into different circles - according to the Dunbar's circles model - we apply a partitioning clustering algorithm on the two different data spaces. The former is compesed of one dimensional points representing the values of the frequency of outgoing communication of each ego network, while the latter contains the points representing the $AdjFreq$ index. We verify that the Facebook ego networks of our sample show a typical number of circles similar to that found in humans (i.e., equal to 4). Specifically, we iteratively apply the $k$-means algorithm with an increasing value of $k$ on each ego network, until the variance explained by the partitioning scheme remains under a threshold of 3% (this is a standard way of finding optimal clusters using the k-means algorithm [8]). Then we take the resulting $k$ as the optimal number of clusters for that ego. We find that the average optimal number of clusters in the data space of the frequency of outgoing communication is equal to 3.76 with a 95% confidence interval of (3.46,4.06) and a median equal to 4. As far as the $AdjFreq$ index, we obtain an optimal number of clusters equal to 3.88 with a 95% confidence interval of (3.58,4.18) and median 4. Hence, we re-apply the $k-means$ algorithm on all the ego networks fixing $k$ to be equal to 4, using both the frequency of outgoing communication and the $AdjFreq$ index. We study the dimensions, the scaling factors and the typical contact frequency of each Dunbar's circle represented by the union of each cluster found by $k$-means with all the other previous clusters with higher frequency of contact, comparing them with the results found in human ego networks.

Table 1 reports the results of the $k$-means analysis with $k$ fixed to 4, indicating with "size" the dimension of the circles and "sc. f." the scaling factors between adjacent circles. The results are divided in the table into three different parts: (i) results concerning the $k$-means analysis applied to the frequency of outgoing communication; (ii) results of $k$-means applied to the $AdjFreq$ index and (iii) results in the anthropology literature. The mean value of the scaling factors is equal to 3.14 for the frequency of outgoing communication and 3.12 as regards the $AdjFreq$ index. These results are really close to the mean scaling factor found in human ego networks. The size of the circles in Facebook (for both the frequency of outgoing communication and the $AdjFreq$ in-

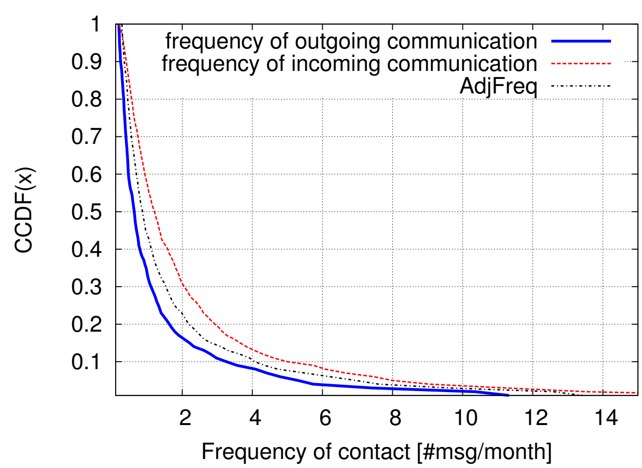


**Figure 4: CCDF of the frequency of outgoing/incoming communication and the index AdjFreq**

[3]The active network size of our sample differs from the results in [1], but in that work the authors considered as active relationship all the social links related to alters who received a manual evaluation greater than zero from the participants of the experiment. This clearly has an impact on the numerical results, although the analysis in [1] also suggests a hierarchical ego network structure.

**Table 1: Results of $k$-means with $k = 4$. 95% confidence intervals are reported in square brackets.**

| | support clique | sympathy group | affinity group | active network |
|---|---|---|---|---|
| Frequency of Outgoing Communication | | | | |
| size | 2.52 [.57] | 7.84 [1.88] | 23.04 [7.41] | 77.8 [33.49] |
| sc. f. | - | 3.11 | 2.94 | 3.38 |
| min freq | 10.49 | 3.94 | 1.15 | .19 |
| AdjFreq Index | | | | |
| size | 2.56 [.63] | 8.24 [2.18] | 24.6 [8.28] | 77.8 [33.24] |
| sc. f. | - | 3.22 | 2.99 | 3.16 |
| min AdjFreq | 12.3 | 5.03 | 1.56 | .28 |
| Results in Human ego networks | | | | |
| size | 4.6 | 14.3 | 42.6 | 132.5 |
| sc. f. | - | 3.10 | 2.98 | 3.11 |
| Projections | | | | |
| outFreq | (4.29) | (13.35) | (39.24) | (132.5) |
| AdjFreq | (4.34) | (14.03) | (41.9) | (132.5) |

dex) is lower than that found in human ego networks. This could be ascribed to the fact that online ego networks represent only a partial subset of the ego networks of a person in real life and on-line ego networks are currently still in the infancy and in a growing phase (the size of an ego network also depends on how long the ego has been active in OSN). Despite this, the structure of Facebook ego networks presents the same hierarchical pattern of the structure of human ego networks. The last two rows of Table 1 report the projection of the results in Facebook calculated in order to make the size of the larger circle fit with its counterpart in real life, for both the frequency of outgoing communication and the $AdjFreq$ index. The results confirm the structure similarity between the Dunbar's circles of the ego networks obtained by ego-net digger and those found in human real ego networks. Moreover, the $AdjFreq$ index, built as a combination of the frequency of outgoing and incoming communication, produces an ego network structure closer to that found in the anthropology literature.

The typical frequency of contact for each circle is reported in Table 1 as "min freq", expressed by the minimum number of posts sent from ego to alters per month within the considered circle. The resulting typical frequency of contact allows us to define the Dunbar's circles in our sample data, considering the frequency of outgoing communication, the group of people contacted at least ∼ *three times a week* (support clique), ∼ *weekly* (sympathy group), ∼ *monthly* (affinity group) and ∼ *twice a year* (active network). The structure we find in Facebook ego networks is thus compatible with that found in human ego networks.

## 4.2 Comparison Between Virtual and Real Ego Networks

In addition to the analysis of the frequency of contact between users, we perform a preliminary study on the differences between the evaluations of the tie strength in Facebook and in the real life collected by the tie strength evaluation module of ego-net digger. Specifically, we assess how the tie strength differ in the two worlds and we study how much the active ego networks in OSN and in real life overlap. To this aim we calculate the normalised root mean square deviation between real life and Facebook scores for each ego network, then averaging the result for all the egos. We find that the tie strength values between Facebook and the real life differ, on average, only for 31%. Moreover, the correlation between real life and Facebook tie strength evaluations (calculated for each ego and averaged for all the egos) is equal to 0.46. This result suggest the possibility to infer the tie strength in real life using Facebook data.

According to the tie strength evaluations, 6.65% of the online user contacts are users' friends only on Facebook. On the contrary, the number of friends contained within users' active network (in real life) and not present inside their Facebook active network, are 19.73%.

# 5. CONCLUSIONS AND FUTURE WORK

In this paper we present ego-net digger, a novel Facebook application for the study of ego networks in Facebook. Ego-net digger, in addition to be a valuable tool for OSN data collection and analysis, gives as output a graphical representation of the Dunbar's circles of users' ego networks. We analyse a sample data set obtained using ego-net digger, showing that there is a similarity between the structure of ego networks in OSN and in real life. Specifically, we find that the typical number of Dunbar's circles in Facebook ego networks is compatible with that found in human ego networks, and the size and the scaling factors of these circles also show an interesting analogy with their real life counterparts. These similarities validates the correctness of the usage of DunbarâĂŹs circles model to display ego networks in Facebook. We also analyse the differences of tie strength evaluations collected by ego-net digger regarding social relationships in the real world and in Facebook, finding interesting results about the overlap between the cyber and the physical active networks. We will continue to use ego-net digger to obtain a larger data set from Facebook. In addition we will provide ego-net digger with other modules for the retrieval of social interaction data from other OSN different from Facebook (e.g., Twitter). With the data we will collect we aim to build models to estimate the tie strength between Facebook users and to study other relevant properties of OSN. We will also continue studying the relation between ego networks in OSN and in real life, analysing how OSN variables could affect the two worlds.

# 6. ACKNOWLEDGMENTS

This work was partially funded by the European Commission under the SCAMPI (FP7-FIRE 258414), RECOGNITION (FP7 FET-AWARENESS 257756), and EINS (FP7-FIRE 288021) projects.

# 7. REFERENCES

[1] V. Arnaboldi, A. Passarella, M. Tesconi, and D. Gazzè. Towards a Characterization of Egocentric Networks in Online Social Networks. In *OTM Workshops*, volume 7046 of *Lecture Notes in Computer Science*, pages 524–533. Springer, 2011.

[2] R. I. M. Dunbar. The social brain hypothesis. *Evolutionary Anthropology*, 6(5):178–190, 1998.

[3] R. I. M. Dunbar and S. G. B. Roberts. Communication in Social Networks: Effects of

Kinship, Network Size and Emotional Closeness. *Personal Relationships*, 2010.

[4] E. Gilbert and K. Karahalios. Predicting tie strength with social media. In *International conference on Human factors in computing systems*, New York, New York, USA, 2009. ACM Press.

[5] B. Goncalves, N. Perra, and A. Vespignani. Validation of Dunbar's number in Twitter conversations. *Networks*, 2011(28):8, 2011.

[6] M. S. Granovetter. The Strength of Weak Ties. *The American Journal of Sociology*, 78(6):1360–1380, 1973.

[7] R. A. Hill and R. I. M. Dunbar. Social network size in humans. *Human Nature*, 14(1):53–72, Mar. 2003.

[8] D. J. Ketchen and C. L. Shook. The Application of Cluster Analysis in Strategic Management Research: an Analysis and Critique. *Strategic Management Journal*, 17(6):441–458, 1996.

[9] P. V. Marsden and K. E. Campbell. Measuring Tie Strength. *Social Forces*, 63(2):482–501, 1984.

[10] S. G. Roberts. Constraints on Social Networks. In *Social Brain, Distributed Mind (Proceedings of the British Academy)*, pages 115–134. 2010.

[11] J. Ugander, B. Karrer, L. Backstrom, and C. Marlow. The Anatomy of the Facebook Social Graph. *CoRR*, abs/1111.4, 2011.

[12] W.-X. Zhou, D. Sornette, R. a. Hill, and R. I. M. Dunbar. Discrete hierarchical organization of social group sizes. In *Biological sciences*, volume 272, pages 439–44, 2005.