USING FACEBOOK ACTIVITY TO INFER SOCIAL TIES*

Sandra Servia Rodríguez, Rebeca P. Díaz Redondo, Ana Fernández Vilas and José J. Pazos Arias
Department of Telematics Engineering, Escuela de Ingeniería de Telecomunicación, University of Vigo, Vigo, Spain

Keywords: Social Networks, Facebook, Tie Signs, Tie Strength, Social Sphere.

Abstract: Clearly there is a growing omnipresence of social networking sites in particular and social services in general. Given this translation of social relations into the cloud, services are facing the problem of deciding, for every user, what are the really relevant links to provide a social-sensitive response. For this, we postulate that users' interaction is a real sign of social relationships which can supplement the topological connections in social sites. To this end, this paper provides a model for calculating the strength of social ties based on interaction information collected from various social APIs in the cloud. From this general model, we detail the deployment of an interaction network for the Facebook case.

1 INTRODUCTION

Social networks have become increasingly popular, turning into an important mean of communication among people of all ages. Although they do not expect to supply traditional communication, they are an important complement to it, allowing users to keep their contact list, share information and interact with others through cross-posting, messaging, games, social events and applications. Recently, several researches on online social networks have came up as consequence of their importance among Internet users. A plentiful number of them is focused on improving users' social experience by means of enhanced-social applications using, for this purpose, information from their profiles and links in these networks, as in (Wilson et al., 2009)'s and (Chen and Fong, 2010)'s case. The former suggests improving a white-listing system for email using social ties strength that allows emails between friends to bypass standard spam filters or detecting Sybil identities[2] in an online community to protect distributed applications. The latter, in turn, proposes a framework of collaborative filtering on social network, for which study the contribution of trust, similarity between profiles, relation between users and reputation.

These applications, as others recently developed, assume that social ties between users and their friends have not the same strength, i.e. the more interaction they have, the more relevant their tie will be. So, improving the effectiveness of these applications requires distinguishing strong ties from weak ties in social networks. With this aim, we have developed an approach to infer social ties between users from their interactions on Facebook. We have selected Facebook to put into practice our approach of gathering users' interaction activity because it is the largest social network with over 800 million active users (Facebook, 2011). Besides considering different tie signs (which denote interaction) on Facebook, we take into account aspects such as that relationships change over time and that they are more intense when less people are involved in them.

This paper is organised as follows. The following section provides a selection of works related to our proposal. Section 3 describes the context in which our application is included. Before detailing our method to infer ties strength indexes from Facebook in Section 5, in next section (Section 4) Facebook signs that imply interaction between users are indicated. In Section 6, we provide an experimental evaluation that shows how our application works properly. Finally, in Section 7, a discussion in this field is provided.

*Work funded by the Ministerio de Educación y Ciencia (Gobierno de España) research project TIN2010-20797 (partly financed with FEDER funds), and by the Consellería de Educación e Ordenación Universitaria (Xunta de Galicia) incentives file CN 2011/023 (partly financed with FEDER funds).

[2]Sybil attacks happens when a malicious user pretends to have multiple identities -Sybil identities- to get to control a peer-to-peer system.

Servia Rodríguez S., P. Díaz Redondo R., Fernández Vilas A. and J. Pazos Arias J..
USING FACEBOOK ACTIVITY TO INFER SOCIAL TIES.
DOI: 10.5220/0003907803250333
In *Proceedings of the 2nd International Conference on Cloud Computing and Services Science* (CLOSER-2012), pages 325-333
ISBN: 978-989-8565-05-1

2 RELATED WORK

The concept of tie strength was introduced by (Granovetter, 1973), who defines it as a function of duration, emotional intensity, intimacy and exchange of services from which ties are split into 'strong' and 'weak' ties. Although this work, as well as many others about this subject, are included in the field of social science, there are also several studies related to the same topic in computer science. For example, (Mutton, 2004) describes a method of inferring a social network by monitoring an IRC channel in which, to obtain the network, an IRC bot observes the messages exchanged between users in the channel and, from this information, infer the social network in which they are involved. Other example is the case of (Tyler et al., 2005), who propose a method for identifying communities using e-mail data.

In the case of online social networks, they already provide users' social graphs, which are made of links between users and their contacts in these networks. However, initial studies on interaction networks (networks made up of ties among users who often interact through social networks) have brought great insights into how an interaction network is structurally different from the social network. Examples of these works are (Wilson et al., 2009), (Viswanath et al., 2009) or (Backstrom et al., 2011), which study users' activity on Facebook to built on the interaction network, taking into account different interaction signs. In (Wilson et al., 2009) case, for each user, these signs are the social graph (they only consider interactions between Facebook friends), wall-posts and photo comments, whereas (Viswanath et al., 2009) only take into account wall-posts to study the varying patterns of interaction over time affect the overall structure of the interaction network. Finally, (Backstrom et al., 2011) study how Facebook users allocate attention across friends, taking into account, as well as messages, comments and wall-posts, information about how many times one user views another's profile page or photos posted by another user. Both (Wilson et al., 2009) and (Viswanath et al., 2009) use Facebook data obtained using crawlers otherwise, (Backstrom et al., 2011) retrieve data directly from Facebook, since information about users' passive interactions such as browsing updates, photos or profiles from their friends through homepage, is not available.

Other studies to deduce the interaction network, as in (Gilbert and Karahalios, 2009), are supported by the information kept in users' profiles: age, political ideals or distance between hometowns, for instance. However, (Kahanda and Neville, 2009) study how to infer the nature and strength of relationships among Facebook's members using attribute-based features (gender, relationship-status,...), topological features (connectivity of the users in the friendship graph), transactional features (Wall postings, picture posting and groups) and network-transactional features (Wall posting in another users Wall,...) to obtain users' "top-friends". They have concluded that the most outstanding features to predict tie strength are network-transactional features, followed by transactional ones.

In this paper, we propose an approach to infer social ties between users from their interactions on Facebook. Although studies like (Gilbert and Karahalios, 2009) consider information kept in users's profiles (age, political ideals,...), from our point of view, they are not as reliable as other signs left by users in the network, mainly different modes of interaction which we consider in our application. Apart from other features of interaction, we consider wall-posts, tagged photos or membership of a group as in the case of (Kahanda and Neville, 2009). However, in our propose, as well as using these features, we take into account aspects such as that relationships change over time and that they are more intense when less people are involved in them.

Finally, aforementioned works get Facebook data using crawlers or directly from Facebook servers. However, as our approach is part of a large project to provide personalized services in the cloud, we are interested in using information obtained through social network APIs. In this paper, we propose getting the most out Facebook API and taking into account all users' information available having their suitable permissions, which is relevant from the point of view of interaction. Anyway, the procedure can be easily generalized to any social networking site with a public API.

3 APPLICATION CONTEXT

Our application to infer tie strength from Facebook interactions is part of the project CLOUDIA[3], which has the goal of defining a suitable ecosystem to assist the users in finding existing services in the cloud to satisfy their specific needs, and also to detect and cater for new business opportunities in the form of services demanded but not existing as a unique whole. To this aim, it will use information stored in the different social networks in which the users may participate.

In CLOUDIA, to find suitable services, the assistance has to be personalized to each user, i.e. depend-

[3]http://gssi.det.uvigo.es/index.php?option=com_content &view=article&id=210&Itemid=439&lang=en

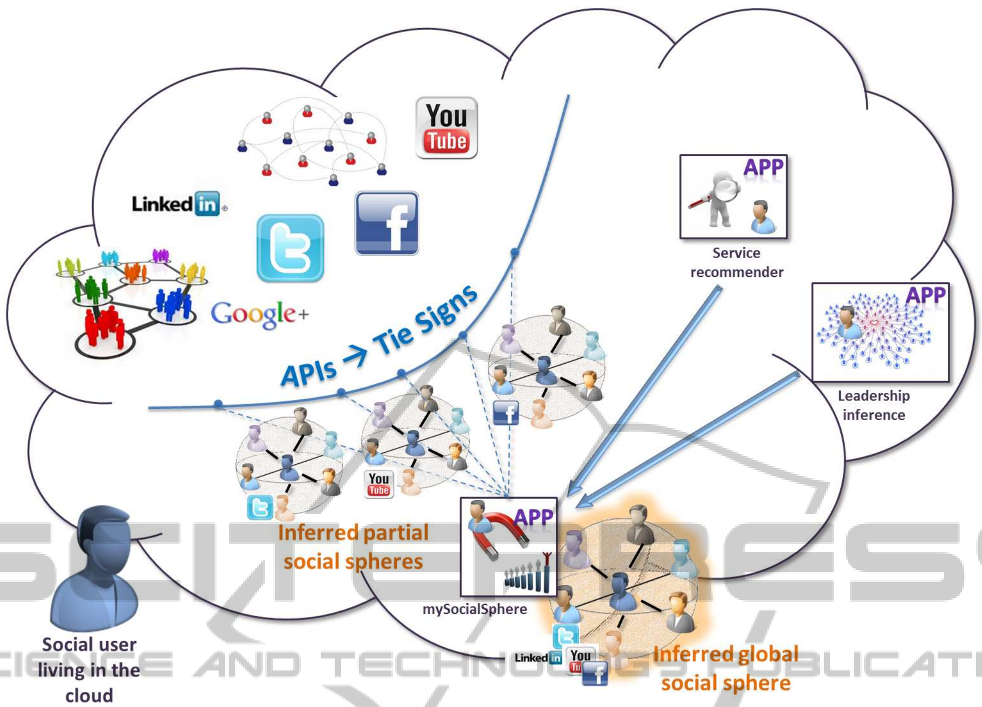


Figure 1: Inferring social ties application.

ing on user's interests and needs, different services will be recommended. Last years, several collaborative filtering recommender systems have been developed. All of them are based on the premise that users who have historically had similar interests will probably continue having it into the present. An important issue in these systems is finding a set of users, known as neighbors, that have a history of agreeing with the target user (having rated different services similarly, tending to use similar set of services,...). Moreover, several authors, like (O'Donovan and Smith, 2005), have improved neighborhood formation taking into account, besides similarity between profiles, social influence or trust between users. As in real life, when we look for an advice for a service (health, commerce, learning,...) we will often turn to our friends, on the basis that we have similar service preferences overall. In the case of an applications, it is necessary knowing who are user's friends, i.e. users in which target user trusts when looks for a recommendation. Likewise, intensity of social relationships often varies, and recommender should be up on these changes. Since it is not feasible that users report to the application, we may get users' ties from social networks sites in which they have subscribed. However, as (Wilson et al., 2009) shows, social network users usually tend to interact with a few of their friends (Wilson applied it to Facebook, but it can be generalized to the rest social networks). That is, a friendship relation on Facebook does not necessarily indicate a real relationship between them. So, knowing which are the real ties between a user and his friends is essential in social networks.

At this point, our application would come into play, taking care of monitoring users' social networks activity and extracting, from this activity, users' ties strength. From them, the application will be able to build the users' social spheres and, ultimately, find the suitable services to them. Even though in this paper we focus on Facebook social network, our goal is to extend the application to any social network.

In Figure 1 our proposal is shown: a social service, *mySocialSphere*, which lives in the cloud and is in charge of monitoring and processing evidences of relationships to build up the user's social sphere. It builds up user's partial social spheres using user's information obtained from automated queries to different social networks APIs, which are combined to form the inferred global social sphere, which be used in CLOUDIA to find suitable services to users.

4 TIE SIGNS: THE FACEBOOK CASE

Facebook provides its users with the typical interpersonal communication features, although its high-

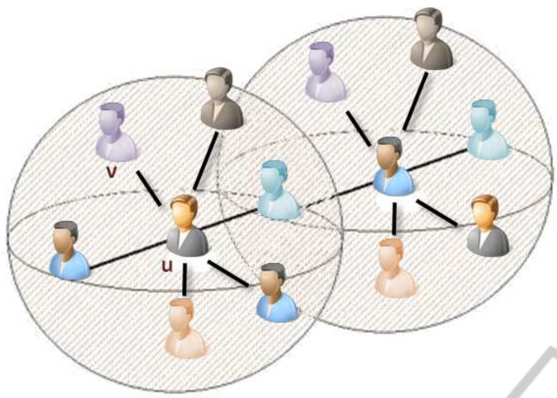


Figure 2: User's social sphere.

light is the *wall*. Subscribers use the wall to post photos, videos, links and messages that may be enriched with any friends' comment. Besides, *mini-feeds* provide detailed logs of each subscriber's actions, so any friend may see at a glance how was his evolution in Facebook over time. As in any social network, security is a key factor and Facebook allows their subscribers to personalize the privacy settings to restrict access to the profile information, mini-feed, wall posts, photos, comments, etc. only to friends, friends-of-friends, lists of friends, no one or all.

On the premise *the more interaction between two users, the more tie strength*, we have developed a Facebook application[4] that extracts user's activity in Facebook and infers the closeness between a target user, u, and one of this friends, v (Figure 2). Since u probably takes advantage from the Facebook facilities to communicate to v (private messages, wall-posts, photos and videos uploads, etc.), we use all this interactions as signs to built a model that calculates the tie strength between u and v, from the u's perspective: $TS_u(v)$. Please, note that this subjective point of view surely cause that the tie strength from the v's perspective, $TS_v(u)$, is different.

After a detailed analysis of Facebook features and how users interact and communicate, we have identified the interaction signs whose mathematical notation is as follows:

Wall-posts. Let $P(x,y) = \{p_1(x,y), p_2(x,y), \dots\}$ be the set of wall-posts user x has writen on y's wall. Thus, $P_s(x) = \bigcup_{\forall y} \{P(x,y)\}$ is the set of x's posts over his friends' wall and $P_r(x) = \bigcup_{\forall y} \{P(y,x)\}$ is the set of wall-posts that x has received.

[4]Using the OAth2.0 protocol, our application requires the target user grants a set of privileges that are explicitly required when joining the application (see Section 6).

Private Messages. The set of private messages that user x has sent to user y is denoted by $PM(x,y) = \{pm_1(x,y), pm_2(x,y), \dots\}$. Consequently, $PM_s(x) = \bigcup_{\forall y} \{PM(x,y)\}$ is the set of private messages x has sent and $PM_r(x) = \bigcup_{\forall y} \{PM(y,x)\}$ is the set of private messages x has received.

Comments. Let $C(x,y) = \{c_1(x,y), c_2(x,y), \dots\}$ be the set of comments done by x about y's entries (photos, wall-posts, etc.). Then, $C_s(x) = \bigcup_{\forall y} \{C(x,y)\}$ is the set of comments done by x about any Facebook user's entries and $C_r(x) = \bigcup_{\forall y} \{C(y,x)\}$ is the set of comments that the x's entries have received.

Likes. Let $L(x,y) = \{l_1(x,y), l_2(x,y), \dots\}$ be the set of likes done by x over y's entries. Then, $L_s(x) = \bigcup_{\forall y} \{L(x,y)\}$ is the set of likes done by x and $L_r(x) = \bigcup_{\forall y} \{L(y,x)\}$ is the set of likes that x's entries have received.

Photos and Videos. The set of x's photos where user y has been tagged is defined as $PH(x,y) = \{ph_1(x,y), ph_2(x,y), \dots\}$. Analogously, the set of x's videos where user y has been tagged is defined as $VD(x,y) = \{vd_1(x,y), vd_2(x,y), \dots\}$. Then, $PH(x) = \bigcup_{\forall y} \{PH(x,y)\}$ ($VD(x) = \bigcup_{\forall y} \{VD(x,y)\}$) is the set of x's photos (videos) where any of this friends is tagged.

Belonging to the Same Groups. $G_p(x) = \{g_{p_1}(x), g_{p_2}(x), \dots\}$ denotes the set of the public groups to which x belongs and $G_s(x) = \{g_{s_1}(x), g_{s_2}(x), \dots\}$ denotes the set of the private or secret groups to which x belongs.

Event Attendance. $EV_p(x) = \{ev_{p_1}(x), ev_{p_2}(x), \dots\}$ denotes the set of public events to which x has shown intention to participate and $EV_p(x) = \{ev_{p_1}(x), ev_{p_2}(x), \dots\}$ denotes the set of secret or private events to which x has intention to go.

5 TIE STRENGTH INFERENCE

This paper only focuses on infering tie strength indexes between Facebook friends, so we only analyze how to assess the closeness u perceives about his relationship with v: $TS_u(v) \in [0,1]$. In order to obtain the index value, we propose the following logarithmic function:

$$f(x) = \begin{cases} 0 & \text{if } 0 \leq x \leq \frac{\bar{x}^2}{x_{max}} \\ \frac{\ln(\frac{x_{max}}{\bar{x}^2} x)}{\ln(\frac{x_{max}^2}{\bar{x}^2})} & \text{if } \frac{\bar{x}^2}{x_{max}} < x \end{cases} \quad (1)$$

being $\bar{x}$ and x_{max} the mean and maximum value, respectively. So, $f(x)$ is close to 1 if $x > \bar{x}$, close to 0 if $x < \bar{x}$ and, finally, close to 0.5 if $x \sim \bar{x}$; exactly the behavior we are looking for. With this function as base, Section 5.1 shows our approach to calculate the tie strength index. However, as life itself, tie strength should be a dynamic index reflecting that old interactions are progressively less important and, so, have less influence in the index calculation. Additionally, some signs' influence vanishes as the number of participants increase, so we have also add the concept of *relevance*. Section 5.2 shows as time and relevance are taken into account in the index calculation.

5.1 Tie Strength Calculation

We propose obtaining the strength index of the tie between u and v, from the u's perspective, as a weighted addition of three kind of interactions: (1) *on-line*, $TS_u|_o(v)$; (2) *face-to-face*, $TS_u|_p(v)$; and (3) *interest-based*, $TS_u|_i(v)$:

$$\begin{aligned} TS_u(v) =& \beta \cdot TS_u|_o(v) + \gamma \cdot TS_u|_f(v) \\ &+ (1-\beta-\gamma) \cdot TS_u|_i(v) \end{aligned} \quad (2)$$

On-line Interactions ($TS_u|_o(v)$). Under this name we include those signs that happen exclusively in the Facebook world and do not require a previous face-to-face contact: wall-posts, comments, likes and private messages. We define two subsets: addressed-signs and open-signs. The former draw together the interactions that v explicitly sends to u –a private message, for instance; whereas the latter are those interactions without an explicit receiver –any like, for example. Therefore, $TS_u|_o(v)$ is obtained as follows:

$$TS_u|_o(v) = \alpha \cdot f(x_d(v,u)) + (1-\alpha) \cdot f(x_o(v,u)) \quad (3)$$

where

$$\begin{aligned} x_d(v,u) =& |P(v,u)| + |P(u,v)| + |PM(v,u)| \\ &+ |PM(u,v)| \\ x_o(v,u) =& |C(v,u)| + |C(u,v)| + |L(v,u)| + |L(u,v)| \end{aligned}$$

are the number of addressed-signs and the number of open-sings, respectively and $f(x)$ is the logarithmic function in Equation 1. Since α reflects the importance of addressed-signs, that we consider is significantly more relevant than open-signs, it should be more than 0.5.

Face-to-Face Interactions ($TS_u|_f(v)$). This contribution reflects any interactions showing a previous physical contact between u and v. It is obtained as follows:

$$TS_u|_f(v) = f(x(u,v)) \quad (4)$$

where

$$x(u,v) = |PH(u,v)| + |VD(u,v)|$$

denotes the number of u's photos where v is tagged and $f(x)$ is the logarithmic function in Equation 1.

Interest-based Interactions ($TS_u|_i(v)$). This contribution assesses the common interests that u and v have explicitely shown. In the Facebook universe this may be done by subscribing to a group as well as accepting an event invitation. Thus, it is obtained as follows:

$$TS_u|_i(v) = \alpha \cdot f(y_d(v,u)) + (1-\alpha) \cdot f(y_o(v,u)) \quad (5)$$

where

$$\begin{aligned} y_d(v,u) =& |G_s(u) \cap G_s(v)| + |EV_s(u) \cap EV_s(v)| \\ y_o(v,u) =& |G_p(u) \cap G_p(v)| + |EV_p(u) \cap EV_p(v)| \end{aligned}$$

are the number of addressed sings (private and secret groups and events, and the number of open-signs (public groups and events), respectively and $f(x)$ is the logarithmic function in Equation 1; α, since has the same meaning than in Equation 3, should have the same value and be always over 0.5.

5.2 Impact of Time and Relevance

None all Facebook signs, even belonging to the same kind, should have the same relevance in the index calculation. For instance, being tagged together in a five-people photo it is clearly more relevant than being tagged together in a twenty-people photo; at least, it may be assumed that in the first case the situation entails more closeness. So, some signs' relevance vanishes as the number of participants increase. For time we adopt the same pattern: relevance vanishes as time goes by. Thus, we propose to modify the previous equations by using the following decreasing function:

$$d(x) = e^{-\mu \cdot x} \quad (6)$$

where μ represent the strength of the slope, i.e. the velocity to *vanish* signs' importance: μ_r for relevance and μ_t for time.

Relevance Impact. This aspect only affects face-to-face and interest-based contributions in Equation 2 (photos, videos, events and groups). Face-to-face contribution is obtained by:

$$TS_u|_f(v) = f(x(v,u))$$

where

$$x(v,u) = \sum_{\forall j \in PH(u,v)} d(|tags_j|) + \sum_{\forall j \in VD(u,v)} d(|tags_j|)$$

being $|tags_j|$ de number of tags in the j-picture (or video) and $d(|tags_j|)$ the result of applying Equation 6. To obtain interest-based index, Equation 5, we use the following contributions:

$$\begin{aligned} y_d(v,u) = & \sum_{\forall j \in (G_s(u) \cap G_s(v))} d(|users_j|) \\ & + \sum_{\forall j \in (EV_s(u) \cap EV_s(v))} d(|users_j|) \\ y_o(v,u) = & \sum_{\forall j \in (G_p(u) \cap G_p(v))} d(|users_j|) \\ & + \sum_{\forall j \in (EV_p(u) \cap EV_p(v))} d(|users_j|) \end{aligned}$$

being $|users_j|$ the number of users that are expected to attend j-event or are subscribed in j-group, and $d(|users_j|)$ the result of applying Equation 6.

Gradual Forgetting. Time, however, affects all Facebook signs: the older an interaction is, the lower its weight should be. Thus, applying the decreasing function, the contributions to Equation 3 to calculate $TS_u|_o(v)$ are as follows, being $d(t_j)$ the result of applying Equation 6 to the time of the latest updated of j-Facebook sign:

$$\begin{aligned} x_d(v,u) = & \sum_{\forall j \in P(u,v)} d(t_j) + \sum_{\forall j \in P(v,u)} d(t_j) \\ & + \sum_{\forall j \in PM(u,v)} d(t_j) + \sum_{\forall j \in PM(v,u)} d(t_j) \\ x_o(v,u) = & \sum_{\forall j \in C(u,v)} d(t_j) + \sum_{\forall j \in C(v,u)} d(t_j) \\ & + \sum_{\forall j \in L(u,v)} d(t_j) + \sum_{\forall j \in L(v,u)} d(t_j) \end{aligned}$$

In the case of calculating $TS_u|_f(v)$, the new contribution to Equation 4 is:

$$\begin{aligned} x(v,u) = & \sum_{\forall j \in PH(u,v)} d(|tags_j|) \cdot d(t_j) \\ & + \sum_{\forall j \in VD(u,v)} d(|tags_j|) \cdot d(t_j) \end{aligned}$$

Finally, the new contributions to Equation 5 are as follow:

$$\begin{aligned} y_d(v,u) = & \sum_{\forall j \in (G_s(u) \cap G_s(v))} d(|users_j|) \cdot d(t_j) \\ & + \sum_{\forall j \in (EV_s(u) \cap EV_s(v))} d(|users_j|) \cdot d(t_j) \\ y_o(v,u) = & \sum_{\forall j \in (G_p(u) \cap G_p(v))} d(|users_j|) \cdot d(t_j) \\ & + \sum_{\forall j \in (EV_p(u) \cap EV_p(v))} d(|users_j|) \cdot d(t_j) \end{aligned}$$

6 EXPERIMENTAL EVALUATION

Our evaluation is focused on three stereotyped Facebook users: (1) users having many friends that usually interacts with only a few (our instance is user A having 130 friends, average Facebook user (Facebook, 2011)); (2) users having only a few close friends and interacting with all of them (our instance is user B having 11 friends); and (3) users having a few friends with which hardly interact (our instance is user C having 62 friends). With the objective of assessing the goodness of the previous formulation, we have developed a Facebook application that uses its API to access the available information the subscribers have upload in their Facebook profile. OAth 2.0 is the protocol Facebook uses for authentication (users and applications) and authorization (applications). OAuth provides a method for clients to access server resources on behalf of a resource owner (such as a different client or an end-user). It also provides a process for end-users to authorize third-party access to their server resources without sharing their credentials (typically, a username and password pair), using user-agent redirections. Facebook implementation of the OAuth 2.0 involves three different steps: user authentication (users are prompted to enter their credentials), application authorization (users are asked to authorize the application to access some of their information through permissions) and application authentication (using the application secret, available from the Developer Application).

Since our objective is retrieving information about subscribers' activity in Facebook, our application requires from the following permissions (which are explicitly asked to subscribers whenever they run our application): (i) *basic information permission* (to access name, gender, profile picture, list of friends, networks and any other information the subscriber have shared with everyone), (ii) *offline permission* (to access the previous information any-time), (iii) permissions to access the subscriber's mailbox (*read_mailbox*), wall-posts (*read_stream*), photos (*user_photo*), videos (*user_videos*), events (*user_events*) and groups (*user_groups*). Using this information, we are able to obtain the tie strength index and all the data needed to this experimental evaluation.

6.1 Index Calculation

After several analysis, we have decided that the importance of the directed addressed-signs (α) is 5 times greater than the opened ones, as well as 60% is the weight for online interactions, 25% for physical interactions and 15% for interest-based interactions. Concretely, we have fixed $\alpha = 0.8$, $\beta = 0.6$, $\gamma = 0.25$, $\mu_r = 0.035$ y $\mu_t = 0.01$. We established these values because most of the users do not have many photos or, if they have many, they are not tagged. Also, they attended events and were members of groups with a lot of attendees (members), which indicated that this type of interaction would not be very relevant for them. Also, we observed that online interaction was their main type of interaction. Moreover, we consider that the importance of the event of interaction loses half of its value when there are about 20 users tagged in a photo (video) or members of a group (attendees an event) or when the event of interaction happened approximately two months before the moment in which indexes are calculated.

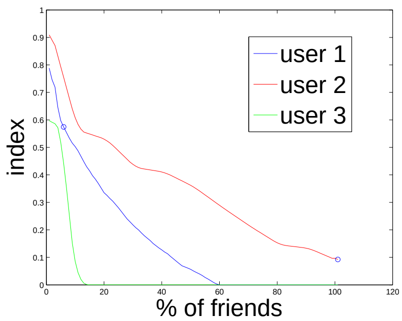


Figure 3: $TS_A(x)$, $TS_B(x)$ and $TS_C(x)$.

Under these conditions, Figure 3 shows the results for each stereotyped user, *A*, *B* and *C*. These results show that the value of the index is greater than 0.5 in, at most, the 23% of the considerate cases. It is in keeping with Wilson's article (Wilson et al., 2009), which indicates that for most of the Facebook users, the large majority of interactions occur only across a small subset of their social links.

Besides, the more spread the allocation attention across friends (or not friends) is, the more difference exists among their indexes. For example, users *B* and *C*, who spread their interactions among their friends more uniformly than *A*, have a lower slope. For example, *X* is a very active user and has the same interaction signs with *A* and *B*, however $TS_A(C) = 0.58$ and $TS_B(C) = 0.009$, as expected. Thus, the tie strength index depends on how the allocation attention across friends is: high for *A* (*A* does not pay attention to many of his or her contacts), whereas for *C* is much lower, since *C* spreads his or her attention more uniformly among them.

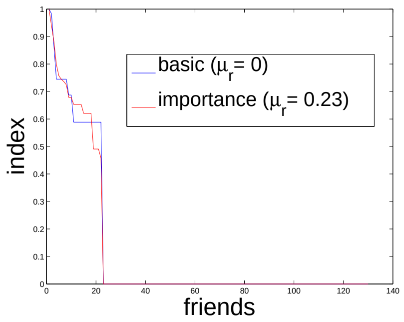
6.2 Relevance and Gradual Forgetting

Now, we consider an user whose Facebook social graph is formed by 130 friends. We study how the index varies using different values for the params that control relevance and time (μ_r and μ_t). We chose this user because is similar than the average user, who has 130 friends (Facebook, 2011).

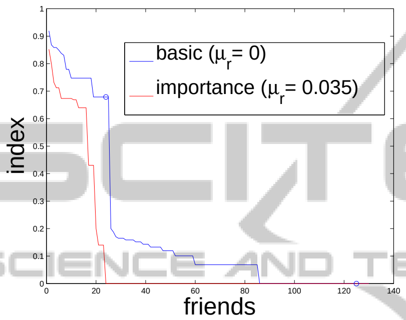
We consider interaction sign's relevance and each kind of interaction separately. Figures 4(a) and 4(b) show how the index varies among the user's friends, taking in account only face-to-face and interest-based interactions, respectively.

In graphic 4(a) we used $\mu_r = 0.23$, which means that the importance of the interaction sign loses half of its value when there are 3 users tagged in a photo (video). In this case, the index is similar when we take in account or do not take in account the relevance of the interaction. The reason is that the user's photos has hardly more than 3 people tagged in them. To interest-based interaction (figure 4(b)) we fixed $\mu_r = 0.035$, which means that the importance of the interaction sign loses half of its value when there are about 20 attendees an event (or members of a group). In this case we observe differences if we consider or do not consider the relevance. When we do not consider it, for the 35% of the users their index is zero, while if we consider it, this percentage rises to the 80% of the users. The main reason is because most of the groups (events) have about 5000 members (attendees), which means that the fact that the user belongs to the group is irrelevant. For example, the user *B* has an basic index of 0.68, while this value drops to 0 when the relevance is considered.

Finally, we consider the importance of the time in each contribution to the index (online, face-to-face and interest-based). In figures 5(a), 5(b) and 5(c), index variation over the user's friends is shown. We fixed $\mu_t = 0.01$, $\mu_t = 0.002$ and $\mu_t = 0$, which means that the event of interaction loses half of its value when it happened two months or a year before the moment in which the index is calculated. The value

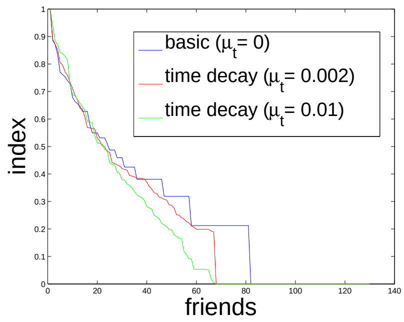


(a) Tie strength index (only face-to-face interactions).

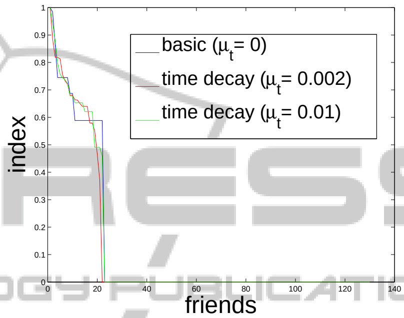


(b) Tie strength index (only interest-based interactions).

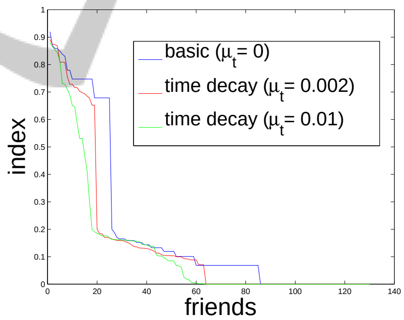
Figure 4: Relevance influence in tie strength index.



(a) Tie strength index (only online interactions).



(b) Tie strength index (only face-to-face interactions).



(c) Tie strength index (only interest-based interactions).

Figure 5: Time influence in tie strength index.

$\mu_t = 0$ happens when time factor is not considered. Wed use the same value for the rest of the parameters that in previous section.

Results of this study are shown in the figures 5(a), 5(b) and 5(c). We obtain that the number of friends with index value is equal to zero is greater when we consider the time factor in the interaction. Also, the lower μ_t is, the more friends have index 0. It is keeping with Wilson's article (Wilson et al., 2009), which indicates that the lower the size of the temporal window in which the interactions happened is, the lower the number of friends with the user interacts is. Also It makes sense that a friend who only is tagged in some photos uploaded 2 or 3 years ago has an index 0 when we consider the importance of the time in the interaction. On the other hand, a friend that is tagged in few photos may have a greater index than another friend tagged in many older photos when we consider the importance of the time in interactions.

7 DISCUSSION

This paper describes an approach to infer social ties from Facebook through its public API, which is included in a large project to infer social ties from any social network in the cloud. The solution goes one step further to integrate the various links maintained by users in all their social networks sites, the ones in which they are registered. We are interested in separating the sheep from the goats, i.e. separating the relevant ties from these other ties which are almost figurative. For instance, despite maintaining Facebook a contact network, users are allowed to interact with people outside their contact network (send a private message, tag a photo with, etc.). In the same way Twitter users can retweet any post, even if they are not following the original poster. So, users may interact frequently with users with who have not direct link in any social network. Reversely, also it may occur that a user has never interact with another user with who shares a direct link in any social network. Consequently, taking into account the users' activity and not only the figurative relation is a more effective approach.

Our proposal is oriented to the construction of the user's social sphere in the cloud taking into account two different-nature contributions. Firstly, the interaction network which can be computed from the formulation in this paper by extending the online, face-to-face and interest-based interactions to other social network sites. Secondly, the topological networks, i.e. the real links which the user maintains (and implicitly accepts) in a plethora of social services. We make out the topological network as a surface where social tie strength is deployed. So, to obtain the social influence between two users we consider (1) the tie strength inferred from the interaction between them and (2) the accumulate tie strengths of paths through, at most, one intermediate user (in Facebook, for example, it would be between friends and friends of friends).

Although, in this paper, the tie strength is only based on Facebook interactions between friends, the proposed formulation may also used to obtain the tie strength between any two Facebook users, not necessarily friends[5]. Despite of the fact that at first sight it is expected for two Facebook friends to have a stronger tie than two non-friends, statistics show Facebook users only regularly relate with a small subset of their 130 friends, on average (Facebook, 2011). Thus, it is perfectly possible for them to have more interaction with a non-friend than with one of their friends, which must not be ignored to obtain the users' social sphere.

Besides, and as aforementioned, users' social sphere should be obtained not only with the Facebook data, but also taking into account their interactions in other social networks. Along this line, we are currently working on extending this approach to other social sites having a public API and adapting the interactions in each social network accordingly to our classification (online, face-to-face and interest-based interactions). For instance, in Twitter, retweets, replies or private messages would be included into the online category, as well as private messages or photo comments in Flickr, private messages in LinkedIn or Wall-posts, comments or +1 in Google+. However, tags in photos or videos in Google+, Picasa or Flickr would belong to the face-to-face category; whereas interactions among users in the same group in Google+, Flickr, LinkedIn, etc. would be categorized as interest-based. Finally, interactions among users occur in any Web 2.0 application, even if it does not have a declarative network as, for example, in the case of blogs or wikis. Consequently, we bear in mind extending our proposal to cover all the range of Web 2.0 application.

[5]Please, note that some of the signs, like wall-post are only available for friends, so the absence of these contributions entails a reduction in the tie strength.

REFERENCES

Backstrom, L., Bakshy, E., Kleinberg, J., Lenton, T., and Rosenn, I. (2011). Center of attention: How facebook users allocate attention across friends. In *Proceedings of the 5th International AAAI Conference on Weblogs and Social Media*.

Chen, W. and Fong, S. (2010). Social network collaborative filtering framework and online trust factors: a case study on facebook. In *The 5th International Conference on Digital Information Management (ICDIM 2010)*, pages 266–273, Thunder Bay, Canada.

Facebook (2011). Facebook statistics.

Gilbert, E. and Karahalios, K. (2009). Predicting tie strength with social media. In *Proceedings of the 27th international conference on Human factors in computing systems*, pages 211–220. ACM.

Granovetter, M. (1973). The strength of weak ties. *American journal of sociology*, pages 1360–1380.

Kahanda, I. and Neville, J. (2009). Using transactional information to predict link strength in online social networks. In *International AAAI Conference on Weblogs and Social Media (ICWSM)*.

Mutton, P. (2004). Inferring and visualizing social networks on internet relay chat. In *Information Visualisation, 2004. IV 2004. Proceedings. Eighth International Conference on*, pages 35–43. IEEE.

O'Donovan, J. and Smith, B. (2005). Trust in recommender systems. In *IUI '05: Proceedings of the 10th International Conference on Intelligent User Interfaces*, pages 167–174, New York, NY, USA. ACM Press.

Tyler, J., Wilkinson, D., and Huberman, B. (2005). E-mail as spectroscopy: Automated discovery of community structure within organizations. *The Information Society*, 21(2):143–153.

Viswanath, B., Mislove, A., Cha, M., and Gummadi, K. (2009). On the evolution of user interaction in facebook. In *Proceedings of the 2nd Workshop on Online social networks*, pages 37–42, Barcelona.

Wilson, C., Boe, B., Sala, A., Puttaswamy, K. P., and Zhao, B. Y. (2009). User interactions in social networks and their implications. In *Proceedings of the 4th ACM European conference on Computer systems*, Nuremberg, Germany.