How to Improve Group Homogeneity in Online Social Networks

Pasquale De Meo, Emilio Ferrara, Domenico Rosaci, and Giuseppe M. L. Sarné

Abstract— **The formation and evolution of interest groups in Online Social Networks is driven by both the users' preferences and the choices of the groups' administrators. In this context, the notion of *homogeneity* of a social group is crucial: it accounts for determining the mutual similarity among the members of a group and it's often regarded as fundamental to determine the satisfaction of group members. In this paper we propose a group homogeneity measure that takes into account behavioral information of users, and an algorithm to optimize such a measure in a social network scenario by matching users and groups profiles. We provide an advantageous formulation of such framework by means of a fully-distributed multi-agent system. Experiments on simulated social network data clearly highlight the performance improvement brought by our approach.**

Index Terms—**Multi-agent systems, Online Social Networks, Group Recommendation, Group Homogeneity.**

I. Introduction

Online Social Networks (OSNs) such as Facebook, Google+ and Twitter have become very complex realities [6], [7], significantly grown in scale and content [5], [18], [26], with significant social effects [10], [11], [19], [30]. In this context, a relevant role is played by social *groups*, that are sub-networks of users sharing common interests [4], [21], [28], [29], [37].

Recent studies investigated the relationships between users and groups in OSNs [2], [23], [24]. For example, Hui *et al.* [23] considered four popular OSNs and empirically computed the probability that a user joins a group; the problem of choosing which group to join has been studied in [2] for a single user and in [24] for a group of users. So far, to the best of our knowledge, no study considers the evolution of a group as a problem of matching between users and groups profiles.

Although the concept of *social profile* is known in the context of virtual communities [25], that of *group profile* is rather novel. The definition of such concept is useful to face the problem of suggesting a user the groups she could affiliate to, so that to improve her satisfaction.

Commonly, a group might be considered (*i*) as a set of nodes (i.e., users) more densely connected among each other than to the others (i.e., the group formation is viewed as a graph clustering problem [13], [14], [20]); or, (*ii*) as a community of people sharing similar interests [12] (i.e., the group formation accounts for some definition of users similarity).

Satisfaction, on the other hand, is often related to the notion of group *homogeneity*: when the similarity/inter-connectivity among group participants is high, according to both structural and semantic dimensions, a OSN group is regards as homogeneous and this yields better satisfaction among its users [27].

However, if we assume that *homogeneity* should reflect users satisfaction, we argue that other behavioral characteristics of members and groups should be considered as important components [8]. For example, in virtual communities, users are often characterized by multiple interests, and groups enact common rules, define accepted behaviors, exhibit a manifold of communication styles and implement different facilities for sharing media content.

In this paper, we define a novel measure of group homogeneity that exploits users similarity and the other users' features cited above. By means of our new definition we are able to provide an algorithm to match the individual users' profiles with group profiles. The goal of this method is to find the matching between users and groups capable of improving the homogeneity of the social groups. More in detail:

- We introduce the notion of *group profile* in the context of OSNs considering a set of categories of interests, common rules, behaviors, communication styles and facilities for sharing media content. This definition of group profile is coherent with the definition of a *user profile* containing information comparable with those of a group profile.
- Each OSN group is associated with a *group agent* [15]–[17], capable of creating, managing and updating the group profile defined above. Similarly, a *user agent* is associated with each OSN user.
- We present a distributed agent platform to handle group formation [31], [32], [34]–[36]. The agents automatically and dynamically compute a matching between user and group profiles in a distributed fashion. We provide the user agent with a matching algorithm, named *Group Homogeneity Maximization* (GHM), and introduce a homogeneity measure between user and group profiles able to determine the group profiles best matching user ones.
- The GHM algorithm will be executed to improve the intra-group homogeneity as follows: (*i*) the user agent submits some requests for joining with the best groups; (*ii*) each group agent accepts only those requests whose originators have profiles matching with the group profile.
- The experimental evaluation of our matching algorithm, carried out on a set of simulated users and groups, clearly shows the advantages of our proposal.

P. De Meo is with the Dept. of Ancient and Modern Civilizations, University of Messina, 98166 Messina, Italy, e-mail: pdemeo@unime.it

E. Ferrara is with the Center for Complex Networks and Systems Research, School of Informatics and Computing, Indiana University, Bloomington (IN), USA, e-mail: ferrarae@indiana.edu

D. Rosaci is with the Dept. DIEES, University of Reggio Calabria, Loc. Feo di Vito, 89122 Reggio Calabria, Italy, e-mail: domenico.rosaci@unirc.it

G.M.L. Sarné is with the Dept. DICEAM, University of Reggio Calabria, Loc. Feo di Vito, 89122 Reggio Calabria, Italy, e-mail: sarne@unirc.it

II. The Reference Scenario

In our scenario, we consider an OSN, the set of its *users*, and the set of its *groups*, denoted by $\mathcal{S}, \mathcal{U}$ and $\mathcal{G}$, respectively. In $\mathcal{S}$, each group of users $g \in \mathcal{G}$ represents a subset of $\mathcal{U}$ (i.e, $g \subseteq \mathcal{U}\ \forall g \in \mathcal{G}$). A multi-agent system is associated with $\mathcal{S}$, such that: (*i*) each user u is supported by her personal agent a_u in the activities of participation to groups; and, (*ii*) each group g is supported by an administrator agent a_g managing all the received requests to join with the group.

A. The agents knowledge

To represent the knowledge that each agent a_u (resp., a_g) has about the orientations of its user u (resp., group g), a *profile* p_u (resp., p_g) is associated with it. This profile stores *preference* and *behavioral* information referred to the user u (resp., the users of g) in four section (called *interests*, *access preference*, *behaviors* and *friends*) storing data on topics of interest, mode to access groups, ways of performing activities and friends, respectively. The profile of a user u (resp., a group g) is represented by a 4-tuple $\langle I_u, A_u, B_u, F_u \rangle$ (resp., $\langle I_g, A_g, B_g, F_g \rangle$), where each component describes the properties of u (resp., g).

Let C be the set of all categories considered in the OSN, where each element $c \in C$ is an identifier representing a given category (e.g. *music*, *sport*, etc.). Each OSN user u (resp., group g) deals with some categories belonging to C where I_u (resp., I_g) denotes a mapping that, for each category $c \in C$, returns a real value $I_u(c)$ (resp., $I_g(c)$), ranging in $[0..1]$. This represents the level of interest of the user u (resp., the users of the group g) with respect to discussions and multimedia content dealing with c. The values of this mapping are computed based on the actual behavior of u (resp., of the users of g) — see Section II-B for the details.

The access mode property represents the policy regulating the access to a group (described by an identifier, e.g. *open*, *closed*, *secret*, etc.) preferred by u (set by the administrator of the group g) and denoted by A_u (resp., A_g).

The property B_u represents the types of behavior adopted (resp., required) by u in her OSN activities, for instance "publishing posts shorter than 500 characters". Let $b \in \mathcal{B}$ a behavior adoptable by user u (admitted in the group g) and described by a boolean variable set to $true$ if b is adopted (resp., tolerated) or $false$ otherwise and let $\mathcal{B}$ be the set of possible behaviors associated with the OSN (e.g., $B = \{b_1, b_2, \cdots, b_n\}$). Therefore, let B_u (resp., B_g) be a mapping that, for each $b \in \mathcal{B}$, returns a boolean value $B_u(b)$ (resp., $B_g(b)$), where $B_u(b_i) = true$ means that such behavior is adopted by u (resp., tolerated in g).

The property F_u (resp., F_g) represents the set of all users that are friends of u (resp., that at least have a friend among the members belonging to the group g).

B. The agents tasks

The agent a_u (resp., a_g) automatically updates the profile p_u (resp., p_g) of its user u (resp., group g) after that u (resp., a user affiliated to g) performs an action involving an information stored in its profile. In particular, every time u deals with a category c, the associated value $I_u(c)$ is updated as the weighted mean between its previous value and the new contribution to $I_u(c) = \alpha \cdot I_u(c) + (1 - \alpha) \cdot \delta$. In detail, α and δ are real values arbitrarily set by u in [0..1], where δ is the increment to give to the u's interest in c due to her action, while α weights the two components of $I_u(c)$. Similarly, every time the $I_u(c)$ value of any user $u \in g$ changes, the $I_g(c)$ value of a group g is updated by the agent a_g as the mean of all the $I_u(c)$ values $\forall c \in g$. For each action performed by the user u (e.g. publishing a post, etc.) its agent a_u sets the appropriate boolean values of the variables in B_u. Analogously, the agent a_g updates the variables contained in B_g every time the administrator of g changes the associated rules. Besides, when u (resp., the administrator of g) modifies her preferences about the access mode, the associated agent updates A_u (resp., A_g). Also, when u (resp., a user of g) modifies her friends list, the associated agent updates F_u (resp., F_g). Note that a_g computes F_g as the union of the sets F_u of all the users of g.

Periodically, the agent a_u (resp., a_g) executes the user (resp., group) agent task described above, to contribute to the *group matching* activity of the OSN.

To perform the above tasks, the agents can reciprocally interact, send and receive messages thanks to a *Directory Facilitator* agent (DF), associated with the OSN, that provides a indexing service. The DF stores the names of each user and group belonging to the OSN and those of their agents. Note that the DF is the only centralized component in the proposed scenario, while the the GHM matching algorithm is completely distributed on the whole agent network.

C. Definition of homogeneity

In order to represent the potential attitude of the user u to stay in the same group with the user v (resp., to stay in the group g), we define the *homogeneity* between two users u and v (resp., a user u and a group g) as a measure representing how much u and v (resp., u and g) are similar (or, different) with respect to the properties I, A, B and F.

The homogeneity $h_{u,v}$ between the users' profiles of u and v is defined as a weighted mean of the contributions c_I, c_A, c_B and c_F, associated with the properties I, A, B and F, measuring how much the values of each property in p_u and p_v are similar. To this purpose:

- c_I is the average of the differences (in the absolute value) of the interests values of u and v for all the categories present in the social network, that is $c_I = \sum_{c \in C} |I_u(c) - I_v(c)|/|C|$.
- c_A is set to 0 or 1 if A_u is equal or not equal to A_v.
- c_B is the average of all the differences between the boolean variables stored in B_u and B_v, where this difference is set to 0 or 1 if the two corresponding variables are equal or different.
- c_F is computed as the percentage of common friends of u and v, with respect to the total number of friends of u or v as $c_F = |F_u \bigcap F_v|/|F_u \bigcup F_v|$. Note that, to make them comparable, the contributions are normalized in $[0..1]$.

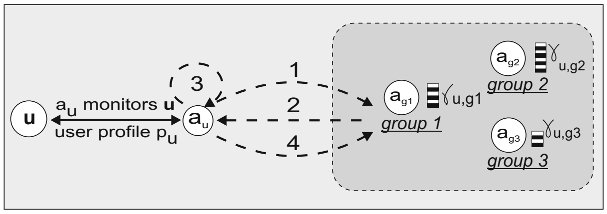

Fig. 1. User agent task schema.

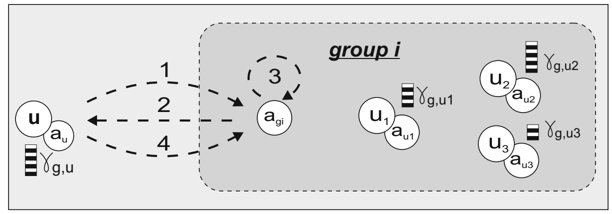

Fig. 2. The group agent task schema.

The homogeneity $h_{u,v}$ is then computed as

$$h_{u,v} = \frac{w_I \cdot c_I + w_A \cdot c_A + w_B \cdot c_B + w_F \cdot c_F}{w_I + w_A + w_B + w_F} \quad (1)$$

Similarly, homogeneity $h_{u,g}$ between a user u and a group g is simply computed as $h_{u,v}$ substituting user v with group g.

III. The GHM Algorithm

The GHM algorithm is a global activity distributed and periodically executed by each user agent a_u (resp., group agent a_g), where we call *epoch* every time the task is executed and T the (constant) period between two consecutive epochs.

A. The user agent task

Let X be the set of the n groups u is affiliated to, where $n \leq n_{MAX}$ and n_{MAX} is the maximum number of groups a user can join with. We suppose that a_u stores into a cache the profile p_g of each group $g \in X$, contacted in the past, with the date $date_g$ of its acquisition. Let m be the number of group agents that at each epoch is contacted by a_u. In such a context, a_u behaves as follows (see Figure 1):

- From the DF repository, a_u randomly selects a set Y of m groups so that $X \bigcap Y = \{0\}$ and let $Z = X \bigcup Y$ the set consisting of all the groups present in X or in Y.
- For each group $g \in Y \bigcap X$ such that $date_g > \psi$ (i.e., a fixed threshold), u sends a message to the agent a_g to ask the profile p_g associated with g (*cf.* Action 1, Fig. 2).
- For each received p_g (*cf.* Action 2, Fig. 1), u computes the homogeneity measure $h_{u,g}$ between her profile and that of the group g (*cf.* Action 3, Fig. 1).
- The groups belonging to Z and having the highest homogeneity values such that $h_{u,g} > \tau$, where τ is a real value ranging in $[0..1]$, are inserted by a_u in the set of good candidates, named $GOOD$, to join with (up to a maximum of n_{MAX} groups). For each group $g \in GOOD$ if $g \notin X$, a_u sends a join request and the profile p_u of u to a_g (*cf.* Action 4, Fig. 1). Otherwise, if $g \in X$ but $g \notin GOOD$, then a_u deletes u from g.

B. The group agent task

Let K be the set of the k users affiliated to the group g, where $k \leq k_{MAX}$, being k_{MAX} the maximum number of members allowed by the administrator of g. Suppose that a_g stores into its cache the profiles of the users $u \in K$ obtained in the past along with the date $date_u$ of their acquisition. When a_g receives a join request by a user agent u (along with u's profile p_u), it behaves as follows (see Fig. 2):

- For each user $u \in K$ such that $date_u > \eta$ (i.e., a fixed threshold), it sends a message to the agent a_u to require the profile p_u associated with u (*cf* Action 1, of Fig. 2).
- When a_g receives the required users' profiles (*cf.* Action 2, Fig. 2), it computes the homogeneity measure $h_{g,u}$ between the profile of each user $u \in K \bigcup \{r\}$ and the profile of the group g (*cf.* Action 3, Fig. 2).
- The user u having the highest homogeneity values such that $h_{g,u} > \pi$, where π is a real value ranging in $[0..1]$, is inserted by a_g in the set of good candidates, named $GOOD$, to join with (up to a maximum of k_{MAX} users). If $u \in GOOD$, a_g accepts its request to join with g (*cf.* Action 4, Fig. 2). Moreover, if $u \in K$ but $u \notin GOOD$, a_g deletes u from g.

IV. Evaluation

We evaluate the effectiveness of the GHM algorithm in increasing the homogeneity of the groups of an OSN by using a simulator, called GHM-Sim, capable of modeling all the required users and groups activities. The experiments involve a simulated OSN having 30.000 users and 100 groups, ad hoc generated by GHM-Sim, each one provided with a profile, having the structure described in Section II. More in detail, the profile p_u of a user u is generated as follows:

- The values of $I_u(c)$ are randomly chosen from a uniform distribution in the interval $[0..1]$;
- A_u is assigned the value *open* (resp., *closed* and *secret*) with a probability of 0.7 (resp., 0.2, 0.1) to implement the variability of OSNs group access restrictions;
- B_u contains the values, randomly generated, of six boolean variables representing in average the user's attitude to: (*i*) publish more than 1 post per day; (*ii*) publish posts longer than 200 characters; (*iii*) comment at least two posts of other users per day; (*iv*) respond to comments associated with her posts; (*v*) leave at least 2 "Like" rates per day; (*vi*) respond to the messages.
- The set of friends F_u are randomly generated choosing in the set of the users.

Users are initially randomly assigned to at least 2 and at most 15 of the available groups. The properties I_g, A_g, B_g and F_g of the profile p_g of each group g are randomly generated. The values of the parameters introduced in Section III are shown in Table I. We also limit to: (*i*) 250 the users who can join a given group; (*ii*) 15 the groups that a user can be joined with; (*iii*) 5 the maximum number of requests that a user can send in each epoch to new groups.

TABLE I
THE VALUES OF THE PARAMETERS USED IN THE GHM-SIM SIMULATOR.

τ	π	$KMAX$	$NMAX$	$NREQ$
0.4	0.4	250	15	5

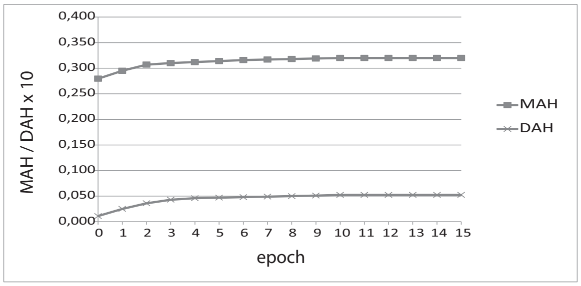

Fig. 3. Variation of MAH and DAH (x10) vs epochs obtained with the GHM-comp and GHM-diff algorithms, for a SN with 30.000 users and 100 groups.

To measure the internal *homogeneity* of a group g we use the *average homogeneity* AH_g, derived by [33], computed as $\sum_{x,y \in g, x \neq y} h_{x,y}/|g|$, while to measure the global homogeneity of the OSN groups we compute the *mean average homogeneity* MAH and the *standard deviation average homogeneity* DAH of all the AH_g, defined as

$$MAH = \frac{\sum_{g \in G} AH_g}{|G|} \quad (2)$$

$$DAH = \sqrt{\frac{\sum_{g \in G}(AH_g - MAH)^2}{|G|}} \quad (3)$$

In the simulations, the initial values for the above measures were $MAH = 0.266$ and $DAH = 0.0011$, denoting a very low homogeneity, due to the random generation. Applying the GHM algorithm we have simulated 15 epochs of execution per user. We can observe that the GHM algorithm quickly converges after few iterations (see Figure 3). The experimental results show that the GHM algorithm increases the homogeneity in OSN groups of about 14 percent on average, with respect to a random assignment of users to groups, achieving a stable configuration (e.g., $MAH = 0.320$ and $DAH = 0.0052$) after about 10 epochs. It is reasonable to suppose that the GHM algorithm, when applied to real OSNs, should lead to concrete benefits in terms of homogeneity.

V. RELATED WORK

In this section we describe some recent research results achieved in the fields covered by this paper, illustrating the main novelties brought in by our approach.

In the latest years, an increasing number of authors focused on the problem of recommending items to the member of a group [1], [22]. This implies the need to construct a *group profile*, often by simply aggregating the individual orientations of its members. This task is usually called *group modelling*.

More formally, let $\mathcal{U}, \mathcal{I}$ and $G \subseteq \mathcal{U}$ be the user population, a collection of items and a group of users, respectively. Suppose that a *rating function* $r : \mathcal{U} \times \mathcal{I} \rightarrow R$ is available, where R (*rating space*) is a discrete set. The function r receives a user $u_i \in \mathcal{U}$ and an item $i_k \in \mathcal{I}$ as input and returns an element $r_{ik} \in R$ as output. Building the profile of G is equivalent to compute a function $f_G : \mathcal{I} \rightarrow R$ receiving an item i_k as input and returns how much the members of G are satisfied by i_k.

To compute $f_G(\,)$ two popular strategies are: (*i*) *Average* [1], where $f_G(i_k)$ is equal to the average of the ratings the member of G have given to i_k. If none of the users in G has rated in i_k, then $f_G(i_k)$ is set equal to $\perp$ (this symbol specifying a not rated item)); (*ii*) *Least Misery* [3], where the rating that group G would assign to i_k is defined as $f_G(i_k) = \min r(u_i, i_k)$ if $\exists u_i \in \mathcal{U} : r(u_i, i_k) \neq \perp$ and $\perp$ otherwise.

In the *Average* strategy the score of an item i_k depends on how many users in G liked it and, if $f_G(i_k)$ is large, i_k could be recommended also to whom in G dislikes it. Otherwise, with *Least Misery* the opinion of who liked the less i_k has the biggest weight in computing $f_G(i_k)$ to minimize the chance that i_k is recommended to someone in G who dislikes it.

For example, if all of the group members *but one* like i_k and the *Least Misery* strategy is applied, i_k will automatically get a low score although almost all users in G are interested in it. Differently, in the *Average* strategy few low ratings on i_k are largely compensated by the ratings of other users.

Besides, most approaches assume that user's preferences are independent of users joining (or not) with a group: if a user alone likes (or dislikes) an item, she will continue liking (or disliking) it if she decides to join a group.

In the literature there are few papers dealing with the matching of a user and a group profile. Most of this work has been designed to recommend to an OSN user groups to join with (such a problem is also called *affiliation recommendation* in [39]). This differs from the group recommendation problem where the objects to recommend are items whereas the affiliation recommendation problem deals with groups.

Spertus *et al.* [38] presented a proposal that describes an empirical comparison of six distinct measures for computing the similarity of a user and a community to exploit for communities recommendation. Chen *et al.* [9] provide an algorithm called CCF (Combinational Collaborative Filtering) which is able to suggest users new friendship relationships as well as the communities they could join with. CCF considers a community from two different but related perspectives (e.g., users and interests) to alleviate the data sparsity arising when only information about users (resp., on words) is used.

Vasuki *et al.* [39] studied the co-evolution of the user's social network of relationships with the affiliation network modelling the affiliation of users to groups. The authors show how such information can be a good predictor to recommend to a user the groups she should join in the future.

Summarizing the benefits provided by our approach are: (*i*) to models user interests, behaviors, friendship relationships and the policies for accessing groups; (*ii*) to manage both group and user profiles by means of a multi-agent architecture where agents provide all the required affiliation activities; (*iii*) to provide a distributed greedy algorithm to match users and groups that computes, at each stage, how good a group is for a given user and selects, uniformly at random, some of these groups; (*iv*) to manage large networks with a large number of groups in a flexible and computationally feasible manner.

VI. Conclusions

The problem of dynamically increasing the intra-group homogeneity is emerging as a key issue in the OSN research field. The introduction of high-structured user profiles, the large dimensions of current OSNs and the increasing number of groups require to face efficiency and scalability issues. In this paper, we presented the *Group Homogeneity Maximization* algorithm that allows a set of software agents, associated with the OSN user profiles, to dynamically and autonomously manage the evolution of the groups, detecting for each user the best groups to join with based on the measures of homogeneity. The agents associated with the group administrators accept only those users having a profile compatible with that of the group. Our experiments on simulated social network data clearly show that the execution of the matching algorithm increases the internal homogeneity of the groups composing the social network, bringing about 15% of improvement with respect to the baseline.

In order to obtain more accurate results, in our ongoing research we are considering to combine the homogeneity measure with a new measure taking into account the trustworthiness of the users. Indeed, in virtual communities, interacting users reciprocally measure the trustworthiness of their counterparts to decide if these are reliable interlocutors or not. To this aim, we are planning a specific experimental session on real OSN data to evaluate our approaches.

References

[1] S. Amer-Yahia, S. Roy, A. Chawlat, G. Das, and C. Yu. Group recommendation: Semantics and efficiency. *Proc. of VLDB Endowment*, 2(1):754–765, 2009.

[2] E. Baatarjav, S. Phithakkitnukoon, and R. Dantu. Group recommendation system for Facebook. In *On the Move to Meaningful Internet Systems: OTM 2008 Work.*, pages 211–219. Springer, 2008.

[3] L. Baltrunas, T. Makcinskas, and F. Ricci. Group recommendations with rank aggregation and collaborative filtering. In *Proc. of ACM Conf. on Recommender Systems 2010*, pages 119–126. ACM Press, 2010.

[4] F. Buccafurri, D. Rosaci, G. M. L. Sarné, and L. Palopoli. Modeling cooperation in multi-agent communities. *Cognitive Systems Research*, 5(3):171–190, 2004.

[5] E. Ferrara, and G. Fiumara. Topological features of online social networks. *Comm. in Appl. and Ind. Math.*, 2(2):1–20. 2011.

[6] S. Catanese, P. De Meo, E. Ferrara, G. Fiumara, and A. Provetti. Crawling Facebook for social network analysis purposes. In *Proc. of Int. Conf. on Web Intelligence, Mining and Semantics*. ACM, 2011.

[7] S. Catanese, P. De Meo, E. Ferrara, G. Fiumara, and A. Provetti. Extraction and analysis of Facebook friendship relations. In Abraham, A., ed.: Computational Social Networks. Springer, pages 291-324, 2012

[8] D. Centola. The spread of behavior in an online social network experiment. *Science*, 329:1194–1197, 2010.

[9] W. Chen, D. Zhang, and E. Chang. Combinational collaborative filtering for personalized community recommendation. In *Proc. of ACM Conf. on Knowledge Discovery and Data mining*, pages 115–123. ACM, 2008.

[10] M. D. Conover, C. Davis, E. Ferrara, K. McKelvey, F. Menczer, and A. Flammini. The geospatial characteristics of a social movement communication network. *PloS one*, 8(3):e55957, 2013.

[11] M. D. Conover, E. Ferrara, F. Menczer, and A. Flammini. The digital evolution of Occupy Wall Street. *PloS one*, 8(5):e64679, 2013.

[12] S. Currarini, M. O. Jackson, and P. Pin. An economic model of friendship: Homophily, minorities, and segregation. *Econometrica*, 77(4):1003–1045, 2009.

[13] P. De Meo, E. Ferrara, G. Fiumara, and A. Provetti. Enhancing community detection using a network weighting strategy. *Information Sciences*, 222:648–668, 2013.

[14] P. De Meo, E. Ferrara, G. Fiumara, and A. Provetti. Mixing local and global information for community detection in large networks. *Journal of Computer and System Sciences*, 2013.

[15] P. De Meo, A. Nocera, G. Quattrone, D. Rosaci, and D. Ursino. Finding reliable users and social networks in a social internetworking system. In *Proc. of 2009 Int. Database Engineering & Applications Symp.*, pages 173–181. ACM, 2009.

[16] P. De Meo, A. Nocera, D. Rosaci, and D. Ursino. Recommendation of reliable users, social networks and high-quality resources in a social internetworking system. *AI Communications*, 24(1):31–50, 2011.

[17] P. De Meo, G. Quattrone, D. Rosaci, and D. Ursino. Dependable recommendations in social internetworking. In *Web Intelligence and Intelligent Agent Technologies, 2009*, pages 519–522. IET, 2009.

[18] E. Ferrara. A large-scale community structure analysis in Facebook. *EPJ Data Science*, 1(1):1–30, 2012.

[19] E. Ferrara, O. Varol, F. Menczer, and A. Flammini. Traveling trends: Social butterflies or frequent fliers? In *Proc. of 2013 ACM Conf. on Online Social Networks (COSN'13)*, 2013.

[20] S. Fortunato. Community detection in graphs. *Physics Reports*, 486(3):75–174, 2010.

[21] S. Gauch, M. Speretta, A. Chandramouli, and A. Micarelli. User Profiles for Personalized Information Access. In *The Adaptive Web*, volume 4321 of *LNCS*, pages 54–89. Springer, 2007.

[22] J. Gorla, N. Lathia, S. Robertson, and J. Wang. Probabilistic group recommendation via information matching. In *Proc. of Int. World Wide Web Conf. (WWW '13)*, pages 495–504. ACM Press, 2013.

[23] P. Hui and S. Buchegger. Groupthink and peer pressure: Social influence in online social network groups. In *Proc. of Int. Conf. on Advances on Social Network Analysis and Mining*, pages 53–59. IEEE, 2009.

[24] J. Kim, H. Kim, H. Oh, and Y. Ryu. A group recommendation system for online communities. *International Journal of Information Management*, 30(3):212–219, 2010.

[25] C. A. Lampe, N. Ellison, and C. Steinfield. A familiar face (book): profile elements as signals in an online social network. In *Proc. of SIGCHI Conference*, pages 435–444. ACM, 2007.

[26] J. Lehmann, B. Gonçalves, J. Ramasco, and C. Cattuto. Dynamical classes of collective attention in Twitter. In *Proc. of the 21st International Conference on World Wide Web*, pages 251–260, 2012.

[27] K. Lewis, M. Gonzalez, and J. Kaufman. Social selection and peer influence in an online social network. *Proc. of National Academy of Sciences*, 109(1):68–72, 2012.

[28] F. Messina, G. Pappalardo, D. Rosaci, C. Santoro, and G. M. L. Sarné, "A Distributed Agent-Based Approach for Supporting Group Formation in P2P e-Learning," in *Advances in Artificial Intelligence*, ser. LNCS. Springer, 2013, vol. 8249, pp. 312–323.

[29] F. Messina, G. Pappalardo, D. Rosaci, C. Santoro, and G. M. L. Sarné, "Hyson: A distributed agent-based protocol for group formation in online social networks," in *Multiagent System Technologies*, ser. LNCS. Springer, 2013, vol. 8076, pp. 320–333.

[30] P. T. Metaxas and E. Mustafaraj. Social media and the elections. *Science*, 338(6106):472–473, 2012.

[31] L. Palopoli, D. Rosaci, and G. M. L. Sarné. Introducing Specialization in e-Commerce Recommender Systems. *Concurrent Engineering: Research and Applications*, 21(3):187–196, 2013.

[32] L. Palopoli, D. Rosaci, and G. M. L. Sarné. A multi-tiered recommender system architecture for supporting e-commerce. In *Intelligent Distributed Computing VI*, pages 71–81. Springer, 2013.

[33] R. Pearson, T. Zylkin, J. Schwaber, and G. Gonye. Quantitative evaluation of clustering results using computational negative controls. In *Proc. of SIAM Int. Conf. on Data Mining*, pages 188–199, 2004.

[34] D. Rosaci and G. M. L. Sarné. Efficient personalization of e-learning activities using a multi-device decentralized recommender system. *Computational Intelligence*, 26(2):121–141, 2010.

[35] D. Rosaci and G. M. L. Sarné. A multi-agent recommender system for supporting device adaptivity in e-commerce. *Journal of Intelligent Information Systems*, 38(2):393–418, 2012.

[36] D. Rosaci and G. M. L. Sarné. Recommending multimedia Web services in a multi-device environment. *Inf. Sys.*, 38(2):198–212, 2013.

[37] D. Rosaci and G.M.L. Sarné. Matching users with groups in social networks. In F. Zavoral, J.J. Jung, and C. Badica, editors, *Intelligent Distributed Computing VII*, volume 511 of *Studies in Computational Intelligence*, pages 45–54. Springer, 2014.

[38] E. Spertus, M. Sahami, and O. Buyukkokten. Evaluating similarity measures: a large-scale study in the orkut social network. In *Proc. of ACM Int. Conf. on Knowledge Discovery and Data Mining*, pages 678–684. ACM, 2005.

[39] V. Vasuki, N. Natarajan, Z. Lu, B. Savas, and I. Dhillon. Scalable affiliation recommendation using auxiliary networks. *ACM Transactions on Intelligent Systems and Technology*, 3(1):3, 2011.