Effective Group Formation in Agent Societies

Antonio Liotta*, Fabrizio Messina†, Domenico Rosaci‡, Giuseppe M. L. Sarné §

*Data Science Centre, University of Derby, Lonsdale House, Quaker Way Derby DE1 3HD U.K., a.liotta@derby.ac.uk

†Department of Mathematics and Informatics, University of Catania, Viale Andrea Doria Catania, Italy, messina@dmi.unict.it

*Department DIIES, University of Reggio Calabria, Loc. Feo di Vito, 89122 Reggio Cal., Italy, domenico.rosaci@unirc.it

‡Department DICEAM, University of Reggio Calabria, Loc. Feo di Vito, 89122 Reggio Cal., Italy, sarne@unirc.it

***Abstract*—In this paper we address the problem of measuring the overall effectiveness of group formation in virtual communities. Group formation is often driven by the combination of similarity and trust measures which are usually exploited with the recommendations provided by *all* the community members (*global reputation*). In this work propose a specific index to measure the effectiveness of group formation, and to exploit the *local reputation* in place of the global one. The use of local reputation will allow group administrators to save a significant amount of computationally and/or communicational tasks. We designed an algorithm to form effective groups in virtual communities and tested it on real data.**

***Index Terms*—Group formation, Virtual Communities, Reputation, Trust**

I. Introduction

Group formation in virtual communities have been widely studied in the last years [1]. Novel affiliations are allowed when the members of a group give a positive assessment towards the incoming agent, i.e. when the joining of the new member will not decrease the social capital (i.e., effectiveness) of the group itself [2]. In this context, an important question is how to measure the overall effectiveness of a group.

To this end, let us suppose that *i*) each agent belonging to the community is characterized by a *social value* v (i.e., its quantitative utility for the community) and *ii*) to classify the members of a community in n classes C_i (with $i \in 1 \cdots, n$) on the basis of their social value. Whenever the composition of a group changes – starting from a certain distribution of the social values – it causes a *social variation* (ΔV). For example, when the percentage of agents belonging to the class C_i assumes the value η_i^*, with η_1 representing the previous value, then $\Delta V = |\eta_1 - \eta_1^*|$. The social variation ΔV_g associated with the n components of the group g can be defined as the average $\frac{\sum_{i=1}^{n} |\eta_i - \eta_i^*|}{n}$ where 0 is the value related to the best group, in terms of social value, and vice versa, 1 represents the worst scenario. To this aim, we define the E_k *index* as the percentage of the groups having a social variation less or equal than $k/100$. Moreover, in this paper, in order to test our approach, we will use the E_{10} index to evaluate the effectiveness of a group.

Note that the formation of a group can not be "driven" by the members' social values, that are known only a posteriori and at a global level (based on all the community members' opinions). Moreover, the updated social values could be unknown for one or more members. In other word, the E_k index allows us to evaluate the effectiveness of a group formation process once the group is formed and cannot be used to obtain a set of groups having a high E_k index value or for evaluating a newcomer request.

In this work we also propose to form groups by exploiting *local reputation* [3], [4], in place of the global one. The local reputation considers opinions only coming from the neighboring agent [5], assumed as more reliable than unreferenced opinions. This is particularly useful in a distributed architecture, where each member can manage local information with a limited consumption of resources, instead of processing the global reputation for which the agent is needed to process all the community members. However, similarly to real societies, if the individual experience is insufficient to trust another member and the number of friends (or friends of friends and so on), then further members' opinions are required (although it is needed to decide how to weight their trustworthiness).

Moreover, in order to perform the best choice about potential newcomers, trust values must be appropriately combined and evaluated within each community. To this purpose, in virtual communities different strategies have been adopted but all of them deeply differ from the processes taking place in human societies. They are mainly based on several different voting mechanisms [6] which gives the advantages deriving from a democratic approach and the disadvantage due to possible manipulations [7] (that we consider as an orthogonal issue with respect to our goals). In this context, we propose the formation of effective groups with respect to the adopted group formation strategy by using *a weighted voting mechanism* (where each vote is represented by a trust value obtained by a suitable combination of reliability and local reputation). In particular, our contribution is mainly represented by the development of an algorithm, called *Effective Group Formation* (*EGF*), aimed at computing individual trust by combining reliability and local reputation information in order to accept or refuse a newcomer affiliation request by using a voting mechanism. The algorithm has been tested on the real data derived by the CIAO [8] community.

The paper is organized as follows: In Section II the related literature is presented and Section III describes the adopted trust measures and the voting procedure. Section IV discusses the EGF algorithm, while Section V presents the experiments we carried out. Finally, in Section VI some conclusions are drawn.

II. RELATED WORK

To form groups within social communities, some researchers proposed to adopt similarity measures as in [9], [10]. Nevertheless, similarity does not guarantee the existence of good interactions among users. Recent studies [11] witness that the larger the mutual members' trust, the larger their interest for mutual interactions [12]. Therefore, to improve the group effectiveness similarity and trust measures can be combined, as in [13], but the computation of similarity measures in huge communities could be too expensive [14], impracticable/unreliable [15]. Differently, forming group techniques only based on trust measures have been proposed [16], but if direct knowledge of a member (i.e., reliability) gives an inadequate knowledge of trust in a community, also opinions of other community members (i.e., reputation) must be used. In particular, the computation of the *global reputation* (i.e., based on the opinions of all the members) could be difficult in presence of unreliable opinions, often due to malicious behaviors [17]. As a consequence the evaluation of the recommender trustworthiness assumes a certain relevance [18], [19]. This leads to realize complex group formation processes dissimilar from those realized in real user societies.

To aggregate trust information we propose using voting to manage individual opinions and interests [20], [21] by reducing conflicts [22] or maximizing the social utility [23]. In the context of huge communities, global voting procedures can be inefficient or unfeasible while a local approach, decomposing the vote in more local votes successively joined, should be desirable [24]. Another aspect is represented by the risks of manipulation due to strategic vote [25], [26]. In particular, software agent societies are more exposed to voting manipulations for the agent aptitude to easily explore different strategies [27].

Our proposal adopts trusts to support voting in virtual communities where local trust and local voting approaches are usually preferred in presence of great population, mobility, lack in infrastructure, communications or limited computational and/or storage capabilities. To this regard, a local trust-voting mechanism is applied in a mobile wireless network context in [28] to establish whether or not a node should be included in a transmission path; the evaluation is based on its trustworthiness as it is perceived by the other nodes. The theory of semi-rings is used in [29] to model trust in Ad-Hoc Networks by a graph where links represent trust relationships on the basis of second-hand information, even though this information is weighted differently from that derived by direct experiences. The authors of [30] discuss a group affiliation procedure where any group joining request is evaluated by means of a democratic group trust-voting mechanism, where each group member exploits an individual local trust-engine.

III. THE TRUST-VOTING PROCESS

Let us denote with $\mathcal{A}$ the agent community and with a directed unlabeled graph $\mathcal{G} = \langle \mathcal{N}, \mathcal{L} \rangle$ the agents relationships in $\mathcal{A}$, where $\mathcal{N}$ is a set of nodes (e.g., the node $n \in \mathcal{N}$ and the agent $a_n \in \mathcal{A}$ represent the same entity) and $\mathcal{L}$ is a set of arcs, where $l_{i,j} \in \mathcal{L}$ represents a trust relationship occurring between the agents $a_i, a_j \in \mathcal{A}$. Below some definitions about trust and local trust will be provided.

a) Trust: Let: *i)* $\hat{t} : \mathcal{A} \times \mathcal{A} \to [0, 1]$ be an *agent trust relationship*, where $0/1$ represents the minimum/maximum trust degree; *ii)* $r_{n,k}$ be the *reliability* measure of the direct trust that a_n has in a_k, derived by his/her direct past experiences with a_k; and *iii)* w_k be the *global reputation* measure of the trust perceived by the whole community about $a_k \in \mathcal{A}$ by averaging all the reliability values $r_{x,k}$, for each $a_x \in \mathcal{A}$. Then, for each agent $a_n \in \mathcal{A}$ the global trust $\hat{t}_{n,k}$ that a_n has about a_k can be computed by weighting reliability and global reputation in the unique measure $\hat{t}_{n,k} = \alpha \cdot r_{n,k} + (1 - \alpha) \cdot w_k$. Note that $\hat{t}$ is an asymmetric measure because it includes r.

The measure $\hat{t}_{n,k}$ can be used to derive the measure $\hat{t}_{n,g}$ (where $g \subset \mathcal{A}$ is an agent group) to determine the "trustworthiness" of g as perceived by a_n and computed by averaging all the values $\hat{t}_{n,k}$ for all the agents $a_k \in g$. Similarly, $\hat{t}_{g,n}$ represents a synthetic measure of the trust that the whole group g has in a_n and computed by averaging all the trust values $\hat{t}_{k,n}$ for all the agents $a_k \in g$. Formally, $\hat{t}_{g,n} = \sum_{k \in g} \hat{t}_{k,n} / |g|$, $\forall k \in g$, where $|g|$ is the size of g.

b) Compactness: The *compactness* combines the *similarity degree* between two agents, or an agent and a group, and their *trust levels* [13]. The similarity $s_{n,k}$ is usually computed by comparing some "features" of the a_n and a_k agents' profile (e.g., interests, item categories and so on), while the similarity $s_{n,g}$ between an agent a_n and a group g by weighting the similarities existing between a_n and all the agents of g. More in detail, the compactness measures $c_{n,k}$ (between agents a_n and a_k), $c_{n,g}$ (between an agent a_n and a group g) and $c_{g,n}$ (between a group g and an agent a_n) are obtained as:

$$c_{n,k} = \gamma \cdot s_{n,k} + (1 - \gamma) \cdot \hat{t}_{n,k}$$
$$c_{n,g} = \gamma \cdot s_{n,g} + (1 - \gamma) \cdot \hat{t}_{n,g}$$
$$c_{g,n} = \gamma \cdot s_{g,n} + (1 - \gamma) \cdot \hat{t}_{g,n}$$

The measure $c_{n,g}$ can be used by any agent to evaluate the goodness of joining with g, while the measure $c_{g,n}$ is useful to a the agent administrator of g for evaluating if accepting a_n into the group.

c) Local trust: Let $t : \mathcal{A} \times \mathcal{A} \to [0, 1]$ be the *local trust*, where 0/1 represents the minimum/maximum trust level. The local trust for an agent $a_n \in \mathcal{A}$ arises by all the agents a_k linked to a_n by a path $(a_n, \ldots, a_k)$ and by all the oriented paths connecting a_n with all the agents of such sub-set (i.e., sub-graph). For each pair of agents a_n, a_k the local trust $t_{n,k}$ of a_n about a_k is given by the combination of the reliability before defined and a *local reputation* (i.e., $w_{n,k}$) computed by summing the contributions of how much the agents, belonging to the ego-network of a_n, trusts a_k.

Let $D(n,k)$ be the set of agents belonging to the ego-network of a_n directly connected with a_k. Let $s(n,k)$ be the sum of the contributions, in term of indirect trust, given by the agents $a_h \in D(n,k)$ and let $l_{(n,k)}$ be the shortest path between a_n and a_k. Then the (normalized) local reputation $w_{n,k}$ is defined as:

$$w_{n,k} = \frac{\sum_{h \in D(n,k), h \neq n,k} \frac{1}{2^{(l_{(n,h)}-1)}} \cdot t_{h,k}}{\sum_{h \in D(n,k), k \neq n,k} \frac{1}{2^{(l_{(n,h)}-1)}}} \tag{1}$$

the contribution in (1) given by a_h to $w_{n,k}$ is raised by $1/2^{(l_{(n,h)}-1)}$ so that less importance is given to the trust relationships (h,k) which are "far" from a_n. Finally, the local trust $t_{n,k}$ combines reliability and local reputation:

$$t_{n,k} = \alpha \cdot r_{n,k} + (1-\alpha) \cdot \beta \cdot w_{n,k} \tag{2}$$

where the real parameters $\alpha, \beta \in [0,1]$. The former parameter weights reliability and local reputation to give relevance to one or other. The parameter β considers the dependability of $w_{n,k}$ by the number of agents in $D(n,k)$ that contributed to compute $w_{n,k}$. Specifically, β yields 1.0 if $\|D(u,x)\| \geq \overline{N}$ or $\|D(u,x)\| \cdot \overline{N}^{-1}$ if $\|D(u,x)\| < \overline{N}$, where $\overline{N}$ specifies how many agents of an ego-network are necessary to compute a reliable value of the local reputation. Indeed, for a small number of nodes in $\|D(n,k)\|$ then a_n will not have sufficient information about a_k from his ego-network and the local reputation measure will not be suitably relevant.

Note that the capability to provide reliable opinions is unrelated to other aspects and, therefore, it needs a specific trust/reputation measure [17]. For this reason, in computing our local reputation we prefer to tune its relevance in computing t by means of the parameters α and β above defined.

d) The Local Trust-Voting Mechanism: Voting is the main approach used in deliberative assemblies to assume a decision [31]. The voting mechanism adopted here exploits the local trust above defined to decide whether an agent belonging to A can join with a group. In particular, each member gives a vote based on its local trust measures with respect to the agent presented the joining request to the group. For instance, the agent a_n may express a vote $v_{n,k}$ to accept a_n or not the requester agent a_k in the group g whether the local trust measure $t_{n,k}$ is greater or equal to a threshold $T_g \in [0,1]$. In the former case, $v_{n,k} = 1$, otherwise it is 0. More formally:

$$v_{n,k} = \begin{cases} 0 & \textit{if } \tau_{n,k} < T_g \\ 1 & \textit{if } \tau_{n,k} \geq T_g \end{cases} \tag{3}$$

We assume that the result of the voting process of a group g for a potential new member a_k and a particular voting criterion v (like that in (3)) is the output of a function $V(v, g, a_k)$. For instance, the requester will be accepted into the group only if the majority of its members has voted for its acceptance.

e) Discussion: In Figure 1 we represent a simple example of community of 8 agents in order to explain the computation of the local trust and the voting procedure. Let a_b and a_d (depicted in yellow) be the nodes asking to join with the group. The voting mechanism proposed above requires that all the group members compute their local trust in a_b and a_d. With respect to the node a_a (depicted in orange), its ego-network consists of all the green nodes and of the yellow node a_b. Note that $r_{a,b} = 1$ and $r_{a,d} = 0$ because any edge exists between a_a and a_d. Moreover, in Figure 1 the nodes giving their contribution to compute the local reputation measures $w_{a,b}$ and $w_{a,d}$ are respectively $\langle a_g, a_c, a_h \rangle$ and $\langle a_c, a_e \rangle$. In computing $w_{a,b}$, the contributions of the nodes a_g and a_c, directly connected with a_a, are weighted by 1, while the contribution of a_h is weighted by 0.5, since it is 2 the shortest path with a_a. Therefore, $w_{a,b} = (1 \cdot 0.75 + 1 \cdot 0.75 + 0.5 \cdot 0.5)/(1+1+0.5) = 0.6$. Similarly, in computing $w_{a,d}$ both the contributions of a_c and a_e are weighted by 1 and in this way $w_{a,d} = (0.5 \cdot 1 + 0.5 \cdot 0.25)/(0.5+0.5) = 0.625$. Furthermore, if we adopt $\alpha = 0.5$ and $\beta = 1$ for a_b and a_d, then the two measures of the local trust are $\tau_{a,b} = 0.5 \cdot 1 + (1-0.5) \cdot 1 \cdot 0.6 = 0.8$ and $\tau_{a,d} = 0.5 \cdot 0 + (1-0.5) \cdot 1 \cdot 0.625 = 0.31$, respectively. Finally, if $T_g = 0.5$, the voting result will be that a_b is admitted (i.e., $v_{a,b} = 1$) and a_d is not admitted (i.e., $v_{a,d} = 0$) into the group.

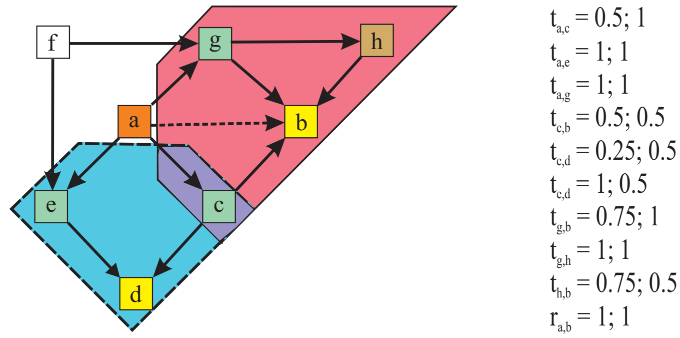

Fig. 1. An example of agent community (the labels report the trust/reliability values and their weights).

IV. The distributed Group Formation procedure

In this Section we present the distributed algorithm EGF to form groups in an agent community by using local trust information and a voting procedure. This algorithm consists of two parts executed by the software agents that operate in behalf of their users. The former part is executed on the agent requester side, while the second part of the algorithm will be executed by the agent managing the group.

A. The EGF algorithm running on the requester agent side.

In Figure 2-A is listed the EGF pseudocode running on the a_n side, where X_n is the set of the groups to which the agent a_n belongs to and $\overline{N}_{MAX}$ is the maximum number of groups that an agent can analyse by fixing that $\overline{N}_{MAX} \geq |X_n|$. Besides, the generic agent a_n stores the profile of each group g_j contacted in the past and the time elapsed d_j from the last EGF execution for that group. Moreover, let δ_n be a timestamp threshold and $\psi_n \in [0,1]$ be a threshold fixed by a_n. The agent

a_n tries to improve its benefits when joining with a group and, to this aim, the values of c are recalculated when older than a threshold δ_i (lines 1-4). Then, candidate groups are sorted in a decreasing order with respect to the compactness c (see the previous section) and the N_{Max} groups are selected in the loop of lines 7-16. If some groups in the set L_{ok} are not in X_n, then a_n could improve the overall compactness by joining with those groups within the maximum number of groups that a_n can join with.

EGF Procedure, executed by the agent a_n

Input:
$X_n, N_{MAX}, \delta_n, \psi_n$;
$Y = \{g \in G\}$ a set of groups randomly selected :
$|Y| \leq N_{MAX}$, $X_n \bigcap Y = \{0\}$, $Z = (X_n \bigcup Y)$

1: $m \leftarrow 0$;
2: **for** $g_j \in Z : d_g > \delta_i$ **do**
3: Send a message to $\hat{a}_{g_j}$ to retrieve the profile P_j.
4: Compute c_{i,g_j}
5: **end for**
6: Let be $L_{ok} = \{g \in Z : c_{n,g_j} \geq \psi_n\}$, with $|L_{ok}| \leq N_{MAX}$
7: $k \rightarrow 0$
8: **for** $g_j \in L_{ok} \wedge g_j \notin X_i$ **do**
9: send a join request to $\hat{a}_{g_j}$
10: **if** $\hat{a}_{g_j}$ accepts the request **then**
11: $m \leftarrow m + 1$
12: **end if**
13: **end for**
14: **for** $g_j \in \{X_i - L_{ok}\} \wedge m > 0$ **do**
15: Sends a leave message to g_j
16: $m \leftarrow m - 1$
17: **end for**

EGF Procedure, executed by the g_i admin agent $\hat{a}_j$

Input: $K_j, K_{MAX}, a_n, w_j, Z = K_j \bigcup \{n\}$;

1: **for** $a_m \in K_j$ **do**
2: **if** $d_i \geq w_j$ **then**
3: ask to a_m its updated profile
4: **end if**
5: **end for**
6: **if** $(V(v, g_j, a_n) == 0)$ **then**
7: Send a reject message to a_n
8: **else**
9: **if** $|Z| \leq K_{MAX}$ **then**
10: Send an accept message to a_n
11: **else**
12: **for** $a_m \in Z$ **do**
13: compute $c_{g_j - \{a_m\}, a_m}$
14: **end for**
15: Let $S = \{s_1, s_2, \ldots, s_{K_{MAX}+1}\}$, with $s_i \in Z$ and $c_{g_j - \{s_i\}} \leq c_{g_j - \{s_k\}}$ iff $i \geq k$
16: **if** $S[K_{MAX} + 1] == a_n$ **then**
17: Send a reject message to a_n
18: **else**
19: Send a leave message to the node $S[K_{MAX} + 1]$
20: Send an accept message to a_n
21: **end if**
22: **end if**
23: **end if**

Fig. 2. EGF algorithm. Top A) - Member Agent. Bottom B) - Group Agent

Trust model	
A	The virtual community
N	The number of agents of the community
α	Weighting reliability vs reputation
γ	Weighting trust vs similarity in computing compactness
β	Scaling factor for the local reputation
$L(n,k)$	Local network of a_n with respect to a_k
$\omega_{n,k}$	Local reputation of a_n in $L(n,k)$.
$\tau_{n,k}$	Local trust of a_n about a_k
$\hat{\tau}_{n,k}$	Global trust of the agent a_n about the agent a_k
$\eta_{n,k}$	Compactness on the agent a_n and a_k
$\eta_{n,g}$	Compactness on the agenta_n and group g
$\eta_{g,n}$	Compactness on the group g and the agent a_n

Group formation	
N_{max}	Maximum number of agents a group is able to host
K_{max}	Maximum number of groups a agent can join with
C_1	Class of agents $h \leq 2$
C_2	Class of agents $2 < h \leq 3$
C_3	Class of agents $h > 3$
T_g	Trust threshold for the voting mechanisms
p_1, p_2, p_3	Probability that a agent of class C_i will join a group

TABLE I
SYMBOL TABLE

B. *The EGF algorithm running on the group agent.*

In Figure 2-B is listed the EGF pseudocode running on the group manager side, where K_j is the set of the agents affiliated with the group g_j belongs to and K_{MAX} is the maximum number of agents to join with the group g_j, with $||K_j|| \leq K_{MAX}$. Suppose that the group administrator $\hat{a}_{g_j}$ stores the profile P_i of each agent a_i and the timestamp d_i of its retrieval. Then, $\hat{a}_{g_j}$, fixed the time threshold w_j, actives the procedure each time that an agent a_n sends a request to join with g_j. In lines $1 - 5$, $\hat{a}_{g_j}$ asks the updated profile of the components of the group itself. By line 6 of the algorithm, a request is sent at all the agents belonging to the group g_j to send their preferences (i.e., a vote) about the possible joining of a_n with the group g_j by adopting the strategy specified in the previous section. After the voting, different choices exist, namely:

- if the agents of g_j voted to refuse a_n in their group, then the procedure ends with a_n out by the group g_j (line 7);
- if the agents of g_j voted to accept a_n in their group; if the number of agents already present in g_j is less than K_{MAX} then a_n is accepted into g_j, otherwise a_n is not accepted into g_j (line $8 - 10$);
- if the agents of g_j voted to accept a_n in their group and the number of agents into the group is already equal to K_{MAX} then for a_n and all the agents belonging to the group is updated the value of their compactness (lines $12 - 14$) then for the agent (denoted by m) having the worst value of the compactness c (line 16):
  - if m is the same agent a_n then it is not admitted into g_j (line 17);
  - if m is not the agent a_n, then a_n will take the place of m into g_j (lines 19 and 20).

V. EXPERIMENTS

The results of some tests on the EGF algorithm are shown in this section. To this aim, a publicly available

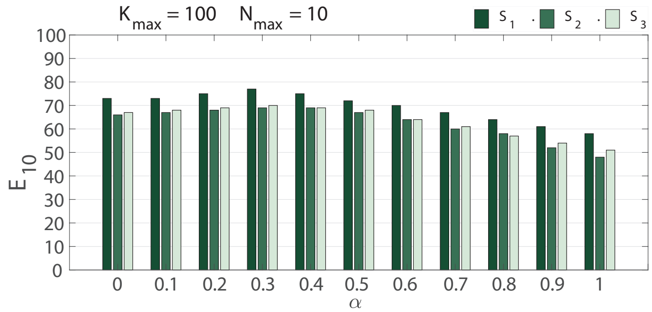

Fig. 3. Configurations S_1, S_2 and S_3 with $N_{max} = 10$ and $K_{max} = 100$

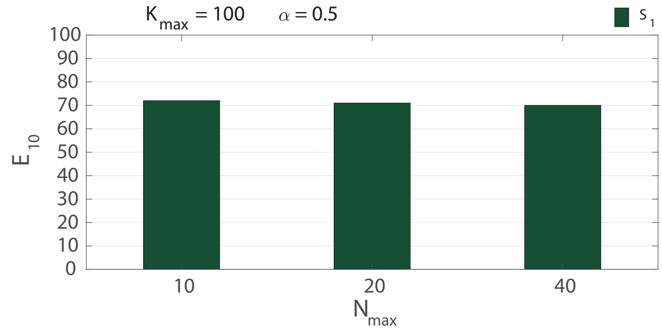

Fig. 4. Configuration S_1 with $K_{max} = 100$ and $\alpha = 0.5$

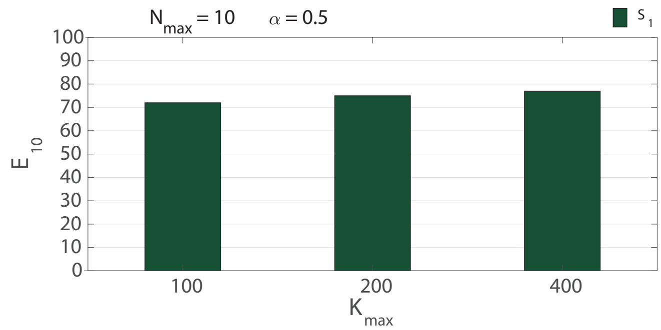

Fig. 5. Configuration S_1 with $N_{max} = 10$ and $\alpha = 0.5$

dataset extracted from the social network CIAO [8] has been used. It stores data of reviewed items, referred to $12,375$ users, about user-item ratings and user-trust relationships stored into the matrices EM and TM. In particular, each EM row consists of $\langle userID, productID, categoryID, rating, helpfulness, timestamp\rangle$ data, where the first three terms identify a user, the product category and the rated product itself, the fourth term is the rate (i.e., *helpfulness*) assigned to the review by the other members and, finally, the last term is the review publishing data (unused in these tests). Individual trust networks are built using the helpfulness values.

In our experiments we assumed the helpfulness as a social value and the group formation activity has been addressed to obtain different groups configurations in terms of distribution of social values. To this aim, CIAO members have been partitioned into three classes (C) characterized by the following helpfulness values $C_1 : \ h_1 \leq 2$, $C_2 : \ < h_2 \leq 3$ and $C_3 : \ h_3 > 3$, and defined the three scenarios described in Table II.

The results obtained by EGF, in term of the E_{10} index, for different α, N_{max}, and K_{max} values are depicted in Figures 3, 4 and 5, where: *i*) Figure 3 shows the results for the three configurations for different values of α, fixed $N_{max} = 10$ and $K_{max} = 100$; *ii*) Figure 4 shows the results for S_1 and for different values of N_{max}, fixed $\alpha = 0.5$ and $K_{max} = 100$; *iii*) Figure 5 presents the results for S_1 and for different values of K_{max}, fixed $\alpha = 0.5$ and $N_{max} = 10$.

In detail, Figures 3 highlights that for $\alpha > 0.3$ the E_{10} value decreases, i.e. a greater relevance of reliability with respect to reputation in forming groups. Moreover, the EFG performance for the configurations S_2 and S_3 decreases with respect to S_1 for a higher difficulty to obtain the desired configurations. The influence of K_{max} and N_{max} on E_{10} is not significant and this confirms the trend shown by the results obtained for S_1.

	C_1	C_2	C_3
S_1	0.33	0.33	0.33
S_2	0.10	0.30	0.60
S_3	0.00	0.00	1.00

TABLE II
THE RATIO OF MEMBERS FOR THE THREE CLASSES C_1, C_2 AND C_3

VI. CONCLUSIONS

Group formation in social communities requires the computation of the mutual trustworthiness among their members on the basis of their reputation, a form of social information provided by the community. We observed that the effectiveness of a group depends on the capability of its members to satisfy mutual expectancies. Then, we proposed an index called E_k in order to measure the groups effectiveness with respect to both the desired composition computed and a specific objective. We also proposed a distributed algorithm to form groups in virtual communities by improving the effectiveness of the group formation activity in terms of E_{10} index by a weighted voting mechanism, where each vote is based on a combination of reliability and local reputation. The approach has been tested on real data extracted by the social network CIAO.

ACKNOWLEDGMENT

This study has been supported by **NeCS** Laboratory (DICEAM, University Mediterranea of Reggio Calabria).

REFERENCES

[1] J. M. Leimeister, P. Sidiras, and H. Krcmar, “Success factors of virtual communities from the perspective of members and operators: An empirical study,” in *System Sciences, 2004. Proc. of the 37th Annual Hawaii Int. Conf. on*. IEEE, 2004, pp. 10–pp.

[2] M. Paldam, “Social capital: one or many? Definition and measurement,” *Journal of Economic Surveys*, vol. 14, no. 5, pp. 629–653, 2000.

[3] P. De Meo, F. Messina, D. Rosaci, and G. M. L. Sarnè, “Recommending users in social networks by integrating local and global reputation,” in *Proc. of the 7th Int. Conf. on Internet and Distributed Information Systems*, ser. LNCS, vol. 8729. Springer, 2014, pp. 437–446.

[4] A. Comi, L. Fotia, F. Messina, G. Pappalardo, D. Rosaci, and G. M. Sarné, "A reputation-based approach to improve qos in cloud service composition," in *Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE), 2015 IEEE 24th International Conference on*. IEEE, 2015, pp. 108–113.

[5] M. Eirinaki, M. Louta, and I. Varlamis, "A trust-aware system for personalized user recommendations in social networks," *Systems, Man, and Cybernetics: Systems, IEEE Trans. on*, vol. 44, no. 4, pp. 409–421, 2014.

[6] D. M. Kilgour and C. Eden, *Handbook of group decision and negotiation*. Springer Science & Business Media, 2010, vol. 4.

[7] P. Fishburn, "Arrow's impossibility theorem: Concise proof and infinite voters," *J. of Economic Theory*, vol. 2, no. 1, pp. 103–106, 1970.

[8] Ciao, "http://www.public.asu.edu/∼tang20/datasetcode/truststudy.htm," 2014.

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

[10] D. Rosaci and G. Sarné, "Matching users with groups in social networks," in *Intelligent Distributed Computing VII*. Springer, 2014, pp. 45–54.

[11] T. Snijders, "Network dynamics," *The Handbook of Rational Choice Social Research. Stanford Univ. Press, Palo Alto*, pp. 252–279, 2013.

[12] M. Granovetter, "Economic action and social structure: The problem of embeddedness," *American journal of sociology*, pp. 481–510, 1985.

[13] P. De Meo, E. Ferrara, D. Rosaci, and G. M. L. Sarné, "Trust and compactness in social network groups," *IEEE Tran. on Cybernetics*, vol. 45, no. 2, pp. 205–216, 2015.

[14] V. Agarwal and K. Bharadwaj, "A collaborative filtering framework for friends recommendation in social networks based on interaction intensity and adaptive user similarity," *Social Network Analysis and Mining*, vol. 3, no. 3, pp. 359–379, 2013.

[15] R. Lande, "Statistics and partitioning of species diversity, and similarity among multiple communities," *Oikos*, pp. 5–13, 1996.

[16] Y. Wang and J. Vassileva, "Trust-based community formation in peer-to-peer file sharing networks," in *Proc. of the 2004 IEEE/WIC/ACM Int. Conf. on Web Intelligence*. IEEE Computer Society, 2004, pp. 341–348.

[17] A. Jøsang, R. Ismail, and C. Boyd, "A survey of trust and reputation systems for online service provision," *Decision support systems*, vol. 43, no. 2, pp. 618–644, 2007.

[18] D. Rosaci, G. M. L. Sarné, and S. Garruzzo, "TRR: An integrated reliability-reputation model for agent societies," in *Proceedings of the 12th Workshop from "Objects to Agents", WOA 2014*, ser. CEUR Workshop Proceedings, vol. 741. CEUR-WS.org, 2011.

[19] D. Rosaci, G. M. L. Sarnè, and S. Garruzzo, "Integrating trust measures in multiagent systems," *International Journal of Intelligent Systems*, vol. 27, no. 1, pp. 1–15, 2012.

[20] S. J. Brams and P. C. Fishburn, "Voting procedures," *Handbook of social choice and welfare*, vol. 1, pp. 173–236, 2002.

[21] F. Buccafurri, L. Fotia, and G. Lax, "Social signature: Signing by tweeting," in *International Conference on Electronic Government and the Information Systems Perspective*. Springer, 2014, pp. 1–14.

[22] T. C. Beierle and J. Cayford, *Democracy in practice: Public participation in environmental decisions*. Resources for the Future, 2002.

[23] L. Xia, "Computational voting theory: game-theoretic and combinatorial aspects," Ph.D. dissertation, Duke University, 2011.

[24] Y. Chevaleyre, U. Endriss, J. Lang, and N. Maudet, "A short introduction to computational social choice," *SOFSEM 2007: Theory and Practice of Computer Science*, pp. 51–69, 2007.

[25] V. Conitzer and T. Sandholm, "Universal voting protocol tweaks to make manipulation hard," *arXiv preprint cs/0307018*, 2003.

[26] K. J. Arrow, *Social choice and individual values*. Yale university press, 2012, vol. 12.

[27] J. Pitt, L. Kamara, M. Sergot, and A. Artikis, "Formalization of a voting protocol for virtual organizations," in *Proc. of the 4th Int joint Conf. on Autonomous Agents and Multiagent Systems*. ACM, 2005, pp. 373–380.

[28] T. Jiang and J. S. Baras, "Trust evaluation in anarchy: A case study on autonomous networks." in *INFOCOM*, 2006.

[29] G. Theodorakopoulos and J. S. Baras, "On trust models and trust evaluation metrics for ad hoc networks," *IEEE J. on selected areas in Comm.*, vol. 24, no. 2, pp. 318–328, 2006.

[30] S. E. Williams, L. P. Shoo, J. L. Isaac, A. A. Hoffmann, and G. Langham, "Towards an integrated framework for assessing the vulnerability of species to climate change," *PLoS biology*, vol. 6, no. 12, p. e325, 2008.

[31] L. S. Lai and E. Turban, "Groups formation and operations in the web 2.0 environment and social networks," *Group Decision and negotiation*, vol. 17, no. 5, pp. 387–402, 2008.