Topology of communities for the collaborative recommendations to groups

Cédric Bernier, Armelle Brun, Armen Aghasaryan, Makram Bouzid, Jérome Picault, Christophe Senot

Topology of communities for the collaborative recommendations to groups

Cédric Bernier * **, Armelle Brun *, Armen Aghasaryan **, Makram Bouzid **, Jérôme Picault **, Christophe Senot **, Anne Boyer *

***Abstract*— More and more systems allow user personalization and provide item recommendations, intended to fit individual user interests. In a traditional VoD system, for example, the recommendations are oriented towards a single user even though he is not watching the video alone. Hence, there is a need to have recommendations for a set of users, a group. Collaborative filtering techniques are traditionally used to make a recommendation for a single user. Usage traces or user ratings are used to deduce their profile and to select an appropriate recommendation that way. Performing recommendation for groups is considerably more difficult because the retrieval of a group's traces of usage or ratings is complicated. As the individual profile for each member of the group is usually available, the recommendation for a group can be based on these individual profiles. This paper explores this approach and is the first step of the construction of a software toolkit for computing recommendations in function of the group composition and the chosen strategies.**

***Index Terms*— group profiling, group recommendation, user profiling**

I. Introduction

More and more systems allow user personalization and provide resource recommendations (movies, books, music, videos, etc.) intended to fit individual user interests [2], [9]. However, many human activities are social and involve several people. In a traditional VoD system, for example, the recommendations are oriented towards a single user even though he is not alone while using this system in the household. Hence, there is a need to have recommendations for a set of users (a group of friends, family members, or even strangers assembled in the same place), named a group, present at the same place (e.g. at home, in a pub) and that have the same activity (e.g. watching TV in the household or listening to music in a pub).

Different applications of such a group recommendation system can be used to provide music or videos in a pub, personalized advertisement to a group of friends or to propose a movie to a family. This work is conducted in the framework of the ITEA WellCom project [10]. The project is carried out in the context of a distributed home environment and elaborates advanced multimedia applications based on interactions between TV/STB (set-top box) and nearby mobile telephones featured with RFID and Bluetooth; these technologies allow identifying the users in front of the TV screens.

To make a recommendation for a single user, two main methods exist, collaborative filtering techniques [2] and content-based techniques [9]. The first method exploits usage traces or user ratings in order to deduce individual profiles and to select appropriate recommendations. In the second method, a user profile is composed of items or interest domains that can be associated with a numerical value that represents the importance of the given items or interest domains.

In the same way, the individual recommendations use the individual profile, the recommendation for a group needs a group profile. Nevertheless, getting a group's traces of usage or ratings is considerably more difficult. However, the individual profile (deduced from individual traces or introduced explicitly) for each member of the group is usually available. Thus, one way to cope with the lack of group traces to generate recommendations for a group is to use the individual traces. In this paper, we will explore this approach.

More precisely, we will detail in this paper the different approaches and strategies that are proposed in the state of the art to make recommendations for a group. Moreover, we propose a typology of groups to choose the good strategy to make an adequate recommendation for the group.

This paper is the first step of the construction of a software toolkit to compute resource recommendations in function of the group composition and the chosen strategies. We first describe related works in the area of group recommenders, in section 2. Next, in section 3, we detail the group analyzer that characterized define our proposition of the concept of presence group. Then, section 4 discusses how the recommender engine makes recommendations to a presence group. In section 5, we enlarge our vision to include our concept of presence community. Finally, section 6 presents our conclusions and the follow-up activities.

* LORIA - Université Nancy2, 54500 Vandoeuvre-lès-Nancy, France
{Cedric.Bernier, Armelle.Brun,Anne.Boyer}@loria.fr

** Alcatel-Lucent Bell Labs. Centre de Villarceaux, Route de Villejust, 91620 Nozay, France
{Cedric.Bernier, Armen.Aghasaryan, Makram.Bouzid, Jerome.Picault, Christophe.Senot}@alcatel-lucent.com

II. Related Works

The Fig. 1 represents our vision of the recommendation process for a group. This recommendation process is initiated by the presence of a set of persons (noted (a) in Fig. 1), a presence group. The group analyzer (noted (b) in Fig. 1) uses the topology of groups to determine the characteristics of the presence group. This paper mainly focuses on this step that will be further detailed. From the result of the group analyzer, the recommender engine (noted (c) in Fig. 1) generates recommendations by using strategies of the state of the art. These strategies allow, for example, to privilege a part of the group or to make a consensus.

The rest of this section is dedicated to the presentation of the state of the art of the recommender engine.

Before computing the recommendation, the recommender applies aggregation of individual information: either from the individual profiles of the members of the group, or from the individual recommendations that are computed from the individual profiles. These two approaches are now detailed.

A. Aggregation

1) Group Model Based (GMB) Recommendation

This approach consists in making the aggregation before generating the recommendations. This approach can be performed in two ways:

- building the group profile using the individual profile of each member and generating the recommendations [3], [6] based on this group profile. For example, Mc Carthy *et al.* [3] make a recommendation of radio stations in a fitness center. They built a group profile by calculating the quadratic sum of notes; each note corresponds to a genre of music given explicitly by the user. In this approach, if the individual profiles are heterogeneous, some users may be unsatisfied of this recommendation. O'Connor *et al.* [6] make recommendations of films based on user profiles of MovieLens [11]. The authors keep from each style of film the lowest note of the group. This method is not adapted to large groups because the group profile tends to a minimum profile.
- making an intersection of the individual preferences to know the group's interests [4] where the authors present a system to initiate a conversation between several persons in a common area. From individual profiles, the system computes the intersection in order to discover common points and display on a screen items that satisfy all present persons. However, this system is not adapted to large groups because the larger is the group, the stronger is the probability to obtain an empty intersection.

2) Individual Recommendation Merging (IRM)

Making the aggregation after having computed individual recommendations consists in making an individual recommendation for every member of the group and then aggregating these results [1]. To our knowledge, few works have been interested in this approach.

Ardissono *et al.* [1] go a step further by combining both approaches and aggregating similar profiles to build homogeneous subgroup profiles. The recommendations are

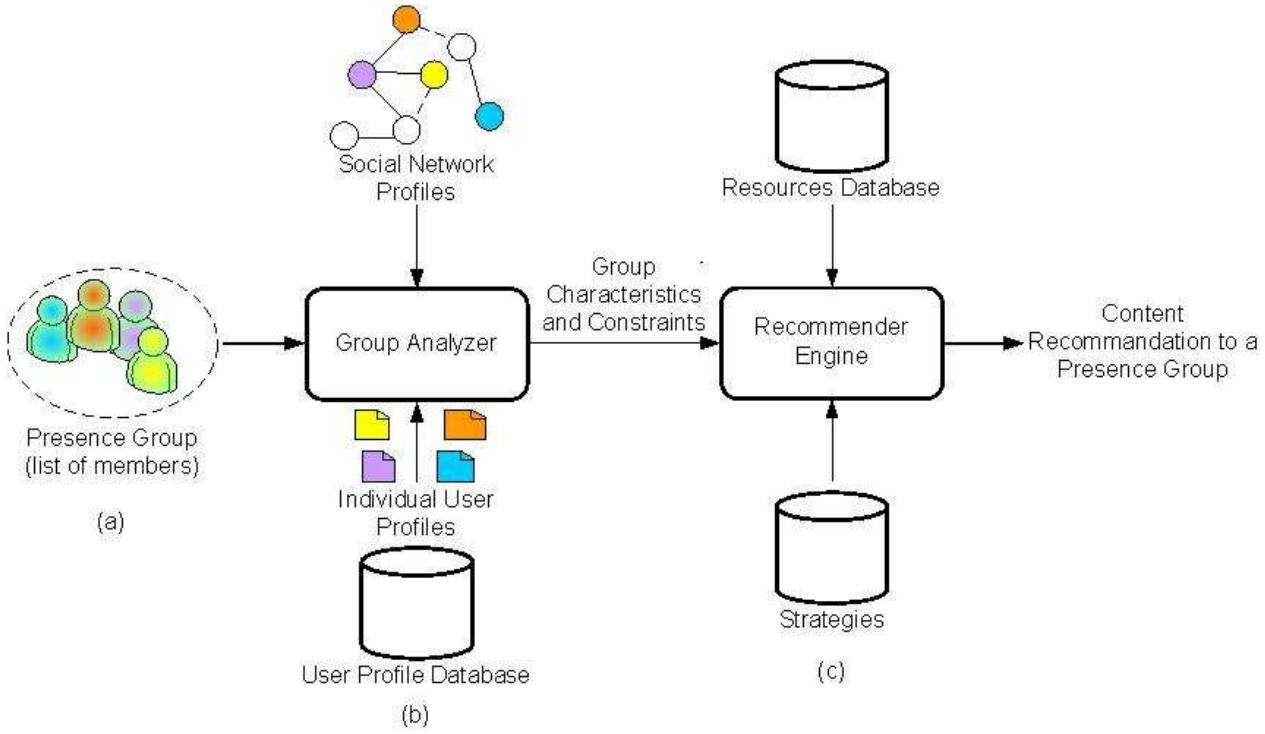


Fig. 1. Architecture to build a contents recommendation to a presence group by taking into account the group characteristics and the strategies

made for each subgroup and then aggregated for all the subgroups. This combination allows treating a large group in the same way as a small heterogeneous group.

B. Strategies

We can find in the state of the art strategies to generate different recommendations. We detail the strategies in this section.

Group recommendations can be improved by using a strategy that defines more precisely the aim of the system and the way the aggregation is made.

In [7], Judith Masthoff lists ten strategies for combining user profiles to generate a group profile. These strategies can also be used to aggregate individual recommendations. Contador *et al.* [5] take back these strategies to test them with a semantic user profile. We can note that they observed that the manual choice made by humans corresponds to *Borda count* and *Copeland Rule* strategies, whereas the automatic choice (with a larger distance between member profiles and items) use the *Average Without Misery* and *Plurality Voting*. Hence, applying only one strategy is not enough. We cite the most frequently mentioned strategies from these authors.

Group recommendation strategies can be divided into three main categories. They can be defined based on the principles adopted for the conciliation of individual preferences, and regardless of the aggregation approach that each strategy applies. Below, we use the term user preference to reflect both a preferred content item in the IRM and an interest category in the GMB approach.

1) Majority-based strategies

They use most popular/shared interest categories or preferred items between group members. Related works mention e.g. Plurality Voting, Borda Count, Copeland Rule, or Approval Voting. For example, with the Plurality Voting strategy, each member (implicitly) votes for his most preferred item/interest category and those with the highest votes are selected. Then, this method is reiterated for the rest of the preferences to obtain a ranked list.

The majority-based strategies allow satisfying the majority of the members of the group, even if the recommendation is extremely unsatisfying for the others.

2) Consensus-based strategies

They consider all group member preferences, such as averaging all users' preferences for each item/concept (Additive or Multiplicative Utilitarian strategy). Other examples include Average without Misery, Fairness, or Satisfaction alternated.

In addition, an Advanced Utilitarian strategy can be considered for highly heterogeneous groups: the computed average values are increased (diminished) if the respective standard deviations are small (large). This will prioritize the preferences where a better consensus is reached.

The consensus-based strategies are adequate for small groups because the users' opinions have a large impact. But the extreme users are not taken into account; these strategies are not adapted to heterogeneous groups.

3) Borderline strategies

They consider only a subset of interest categories/items, belonging to a (subgroup of) member(s) based on the roles or other criteria identified within the group. Related works mention Least Misery, Most Pleasure, Strongly/Weakly Support Grumpiest, Dictatorship or Privileged sub-group strategies. For example, the Least Misery strategy keeps for each preference the minimum weight among the group members; a single member can therefore impose his choices.

In addition, a No Misery strategy can be considered to eliminate items/interest categories disliked by at least one individual (preference value under a threshold). The group recommendations are then based on the remaining preferences.

The borderline strategies are generally adapted for a heterogeneous group because they can be used to allow satisfying each member alternatively.

Different strategies pursue different aims: the manager of the system can choose the policy that he prefers in function of these aims. Moreover, we can consider hybrid strategies obtained by combining different principles. Namely, the Advanced Utilitarian could be combined with the No Misery strategy to avoid recommending content displeasing to at least one group member.

To sum up, the recommender engine is frequently studied in literature, but how to make the good choice among these strategies to obtain a relevant recommendation for a group? We explain this in the following sections.

III. The Group Analyzer

To generate recommendations to a group, we assume that the system knows the individual profile of every member, the place profile, and the moment the people enter or exit the place. The place profile is built in the same way than a user profile: it contains the history of the items that are viewed in the place. For example, in a pub which has a television that plays some videos, the corresponding place profile contains the list of played videos, or derived from the latter, a list of domains (e.g. rock, pop) with a value defining the importance of every domain (it is supposed that for each video its characteristic domains are known in terms of content metadata).

The group analyzer (noted (b) in Fig. 1) aims at deducing a maximum of data in term of constraints and characteristics on the group in order to allow the recommender engine (noted (c) in Fig. 1) to use a relevant strategy on the adequate set of resources.

We detail now the constraints that the recommender engine has to consider and the characteristics of the set of users who compose the group.

A. The constraints

Some constraints related to the group composition have to be taken into account like the presence of children: some items like horror movies cannot be recommended to children, thus such items cannot be recommended to the whole group.

The recommendations have also to be adapted to the evolution of the group, i.e. when a person comes or leaves, recommendations may be recomputed. The process should also be incremental and continue from the previous step, and avoid recomputing the whole recommendation list, when the group evolves. Nevertheless, the system can decide to wait for a more significant modification of the group before calculating a new recommendation list.

Another information that can be taken into account is the group effect on the individuals' appreciation. Indeed, when a user watches a funny movie at the cinema, his perception is influenced by his close environment: if people around him are cheerful or gloomy, he lives this moment differently. Therefore, the group has an effect that can be positive or negative on the perception a user has. Questions are thus how to measure the group effect and how to take it into account to compute a recommendation?

B. The presence group

Our problem aims at making a resource recommendation to a set of persons who are at the same time in the same place. We call this set of persons a *presence group* (noted (a) in Fig.1). For example, in a household, all persons that are in front of the television screen form a presence group. Parents, children, neighbors, friends or others can belong to this presence group. We define the presence group with the following characteristics.

- The *nature of the relations:* it is measured by the proportion of each relationship type in the group (family, friends, colleague, acquaintance, stranger, *etc*). This implies the availability of a graph of social relations (called *social network*).
- The *cohesiveness:* it expresses the strength of the relationships within the group. It can be measured by the density of relations between group members as well as the frequency of their meetings.
- The *social structure:* it expresses the structure of the group. A group may have no structure (e.g. strangers present in a public place), or may have hierarchical, egalitarian or ambiguous structures. This characteristic can be obtained by using role (leaders, followers, etc) detection techniques, usually applied in social network analysis [8].
- The *profile diversity:* it expresses the diversity among individual profiles, e.g. statistical variation in terms of interests, goals, or past interactions. It can be computed globally or for some specific profile dimensions.
- The *size: it* is the number of individuals in the group.

With these characteristics and these constraints from the group analyzer, the recommender engine can generate recommendations for the presence group.

IV. The Recommender Engine

The recommender engine (noted (c) on Fig. 1) uses the group characteristics and the constraints from the group analyzer to build the list of resources to recommend. In a first step, the recommender engine selects a subset of resources according to the constraints. For example, to recommend a film, if a child is present, only adequate films are kept to the generation of recommendation.

In a second step, with the strategy corresponding to the group characteristics, the recommender engine generates a list of relevant resources by using a relevant strategy. This one is chosen in function of the characteristics of the presence group provided by the group analyzer. We detail in the following the influence of each group characteristic on the choice of the strategy.

The cohesiveness, a group characteristic, shows the potential of the group effect. Indeed, if a group has a weak cohesiveness, the group effect will be weak.

Moreover, the nature of the relation and the social structure are deduced of the social network profiles. If a leader appears, the system will privilege him by using borderline strategies because the leader influences the other member with this viewpoint. However, if the social structure is more egalitarian, the system will privilege a consensus strategy.

Likewise, if the profile diversity is homogeneous, a consensus strategy is adapted, like a simple average because all members are taken into account.

In addition, the size of the group is important for the choice of the policy: some policies cannot be used on small or large groups. Furthermore, when the composition of a group evolves over the time and if this evolution is not significant, the history of this evolving group can be maintained and used because the main preferences are supposed to be the same.

V. From Groups to Communities

To improve the quality of the recommendations made to the presence group, we suggest using an additional group characteristic: the *presence community*. To build presence communities, we get inspired from collaborative approaches that use information about similar people to deduce information of a person (for example, estimate the tastes of that person). Here, we search the similar groups (similar in terms of their habits) to estimate the tastes of the current presence group. By exploiting these similar groups, we deduce a *presence community*.

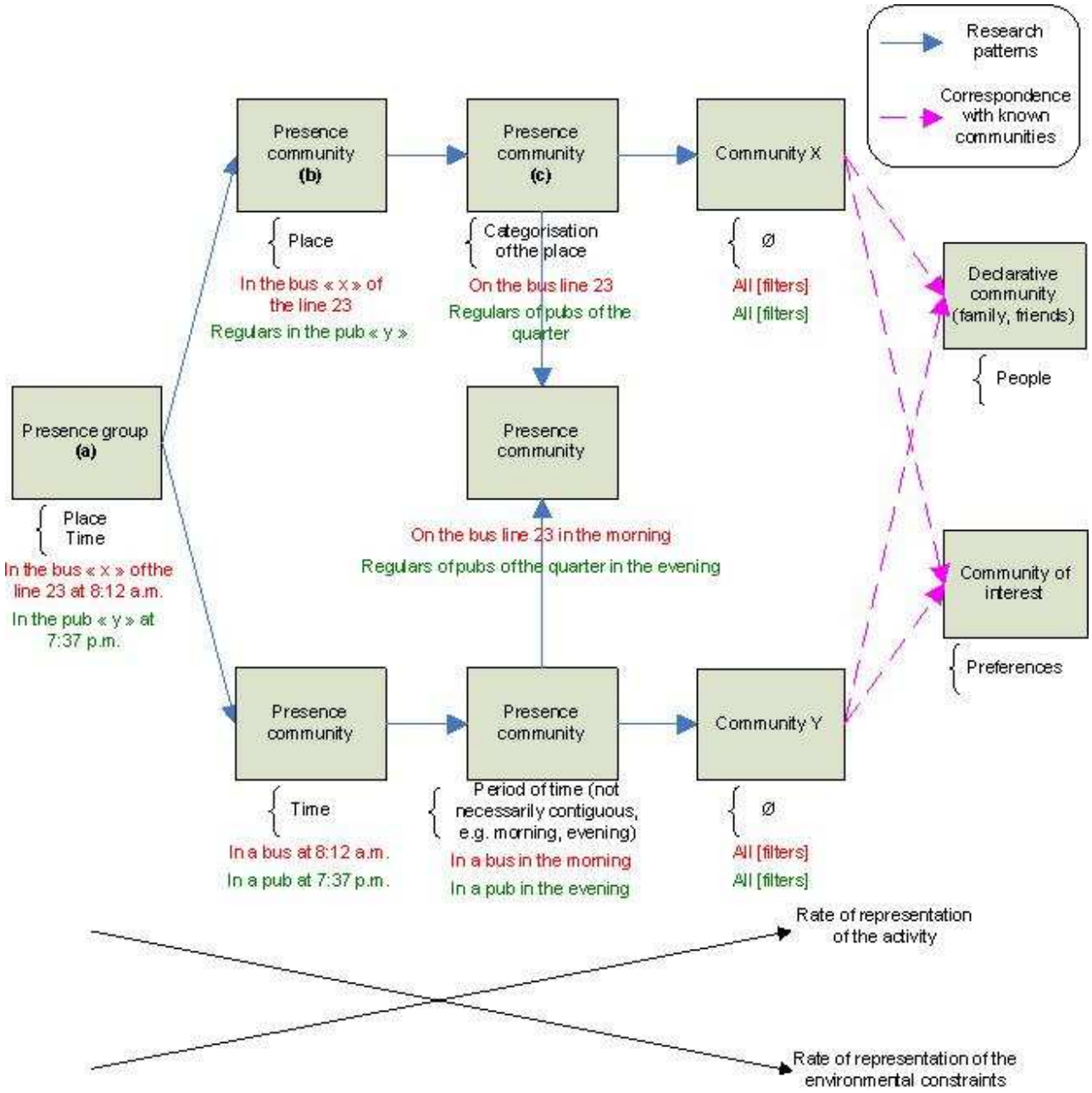
In Fig. 2, (a) is a presence group composed, for example, of all the people that are in pub "y" at 7:37 p.m. (in green) or people in the bus "x" of line 23 at 8:12 am (in red). These people have in common *a priori* only the fact that they are at the same time

at the same place.

To find similar groups, we propose to observe groups by removing one dimension of the presence of the group. For example, find people in that pub "y" over a time period (upper part of Fig. 2) or to observe groups in a pub at this moment (lower part of Fig. 2). Similar groups can thus be either those at the same place at a different time or those at the same time at a given place.

- **In that pub over a time period**.

This point of view allows us to detect the regulars, the clientele of this pub and deduce preference data. By relaxing the time dimension, a set of presence groups is available: the presence group in that pub at 8:00pm, the presence group in that pub at 9:00pm, etc. Given the set of presence groups, we search the frequent occurrences from the members who compose these groups; we thus obtain the people who are more or less regularly together. We can thus place these people in a representation space in function of their involvement for this given place and know people who are often together. In the center of this space, we can find the regulars and people who have come only once are borderline. This partially ordered set of people is called a *presence community*. The more the members are in the center of this space, the stronger is their affiliation, consciously or not, to this community. That is why the profiles of the members who compose the center of this community (the regulars) are nearly



equivalent to the place profile. For example, in a pub, the presence community represents this clientele and the fact to be a good client or not represents their involvement.

- **In a pub at this time**.

By observing groups in pubs at 7:37p.m., we can deduce the persons who go regularly in a pub at 7:37 pm. Similarly to the previous point of view, the resulting presence community is composed of people who practice an activity at the same time (not necessarily at the same place). We obtain in the centre of the space people who practice this activity regularly at this given time.

In consequence, given a presence group, we can compute its corresponding presence community, where a presence community is an abstraction of presence groups. Given this presence community, we have information about its preferences and what activity its members are used to share. The presence community will be exploited as an additional characteristic of the presence group to compute the recommendations.

We can go further in the search of similar groups, by relaxing some additional constraints of presence communities. For example in Fig. 2 (c), the presence community is made of persons regularly in pubs, the place constraint has been relaxed (in (b) people had to be in **that** pub, in (c) they have to be in **a** pub).

As shown in the Fig. 2 below, the granularity of the representation of the activity increases as one goes along the presence community, i.e. the more we relax the remaining dimension (time in the lower part or place in the upper part), the less we have representation of the context, the more the activity is represented. This presence community contains also information about the context.

Hence, with this new concept of presence community, we retrieve new information from the presence group that we can use to improve the knowledge of it by considering the presence community like another characteristic of the presence group.

VI. Conclusion and Future Work

In this paper, we discussed how to generate recommendations to groups. We identified the presence group notion, a set of people who are in the same place at the same time. We have characterized these presence groups and with this, we suggest a first architecture to build a resource recommendation for a presence group.

We presented a classification of several strategies. Given a presence group, the system analyzes its characteristics and, according to these group characteristics, the recommender engine chooses a strategy to generate relevant recommendations.

From several presence groups, we can obtain presence communities which contain more information about the initial group and that can be used as a new group characteristic in order to improve recommendations. A further work will allow testing this concept.

This work is a first step of the construction of a software toolkit to compute recommendations in function of the configuration of the group and the adapted strategy. The next steps are the implementation of aggregation algorithms, the definition and the implementation of different algorithms to recommend items and their tests. In this paper, we ask a question: how to measure the effect of the group? This question will also be studied in future works.

To compute recommendations for a user or a group, the system has to use personal data. Then it has to manage the privacy to secure these data. Therefore, the user privacy will also be investigated.

References

[1] L. Ardissono, A. Goy, G. Petrone, M. Segnan, P. Torasso, "INTRIGUE: Personalized Recommendation of Tourist Attractions for Desktop and Handset Devices". Applied Artificial Intellingence: Special Issue on Artificial Intelligence for Cultural Heritage and Digital, Libraries 17(8-9), pp. 687-714 (2003)

[2] J.S. Breese, D. Heckerman, C. Kadie, "Empirical analysis of predictive algorithms for collaborative filtering". in Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intel ligence, ed., San Francisco Morgan Kauf-mann, pp. 43–52, (1998).

[3] J.F. McCarthy, T.D. Anagnost, "MusicFX: An Arbiter of Group Preferences for Computer Supported Collaborative Workouts". ACM 1998 Conference on CSCW (1998)

[4] J.F. McCarthy, T.J. Costa, E.S. Liongosari, "UniCast, OutCast & GroupCast: Three Steps Toward Ubiquitous, Peripheral Displays". UBICOMP (2001)

[5] I. Cantador, P. Castells, D. Vallet, "Enriching Group Profiles with Ontologies for Knowledge-Driven Collaborative Content Retrieval," Proceedings of the 1st International Workshop on Semantic Technologies in Collaborative Applications (STICA 2006), at the 15th IEEE International Workshops on Enabling Technologies: Infrastructures for Collaborative Enterprises (WETICE 2006), (Manchester, UK, June 2006) pp. 358-363. IEEE Computer Society Press, ISBN 0-7695-2623-3.

[6] M. O'Connor, D. Cosley, J.A. Konstan, J. Riedl, "PolyLens: A Recommender System for Groups of Users". Proceedings of ECSCW (2001)

[7] J. Masthoff, "Group Modeling: Selecting a Sequence of Television Items to Suit a Group of Viewers". User Modeling and User-Adapted Interaction 14: 37-85 (2004)

[8] M.E.J. Newman, "The structure and function of complex networks," SIAM Review 45, (2003) pp.167-256.

[9] M. Pazzani and D. Billsus. "The Adaptive Web, chapter Content-Based Recommendation Systems", pages 325–341. Springer Berlin / Heidelberg, 2007

[10] ITEA WellCom project, http://www.itea-wellcom.org/, http://en.wikipedia.org/wiki/Eureka_ITEA2_WellCom_Project

[11] MovieLens, http://www.movielens.org/