Mobile Networks and Applications (2018) 23:1111–1122
https://doi.org/10.1007/s11036-018-1076-1

Towards Social Big Data-Based Affective Group Recommendation

Minsung Hong[1] · Jason J. Jung[2]

Published online: 6 July 2018

Abstract
Social big data is currently an emergent issue, especially for recommender systems. In particular, with respect to social big data, various data mining techniques have been applied in group recommender systems. However, three social phenomena (i.e., social influence, emotional contagion, and conformity) have not been applied enough in existing studies. In this paper, a novel method for a group recommendation is proposed based upon the affective social network from social big data. In this regard, to explore and measure the social phenomena, a variety of similarity measures were applied in a content-based recommendation. Moreover, the proposed method has a computational complexity of $O(N^2)$, where N is the number of users in a group. Therefore, it is appropriate for big data environments, since the N is generally small for user groups. This study's results revealed that the Mahalanobis distance was suitable for the affective group recommendation. Moreover, the proposed method outperformed the other group recommender systems, those with large groups.

Keywords Social big data · Social influence · Emotional contagion · Conformity · Group recommendation

1 Introduction

The advance of Internet technology has led to an explosive increase in the availability of data. Big data refers to the difficulty in searching for useful information by users regarding their interests. Therefore, to provide appropriate information, various big data analytics have been applied in recommender systems [1–4]. Moreover, the emergence of social network services has encouraged the development of group activities, including watching a movie with friends, going to a restaurant with colleagues, and traveling with one's family [5, 6]. Thus, several studies focused on a group of users rather than an individual user, to recommend the proper items for the group [7–10]. In this regard, to obtain the preferences of the groups, several aggregation strategies have been introduced such as the Plurality Voting, Average, Borda Count, Copeland Rule, Least Misery, and Most Pleasure [11, 12]. However, it has been difficult to gather the preferences of every individual user among a group. Hence, social big data is an emergent issue currently, especially for the group recommender systems (GRS). In particular, to analyze the social networks and discover the relationships between the group users from social big data, like social media, various data mining techniques have been applied in group recommender systems [13–16]. Furthermore, in terms of organizational behavior and psychology, three social phenomena, which occur by user relationships when a user group selects an item, should be considered [17]. These phenomena are as follows:

- *social influences* that effect the decision-making of the groups [13, 18].
- *Emotional contagion* regarding the affective influences between users [12].
- *Conformity* that relates to the opinion expressions according to the user characteristics of the groups [12, 19].

In this regard, researchers have attempted to apply these phenomena in group recommendations using auxiliary information such as trust, satisfaction ratings, and the results of personality tests [5, 7, 8, 10]. However, these studies encounter several problems, such as the absence of a reflecting vertical relationship, usage of ratings as one emotion like satisfaction, and additional burden associated with testing [20, 21].

✉ Jason J. Jung
j2jung@gmail.com

Minsung Hong
msh@vestforsk.no

1 Western Norway Research Institute, Songdal, Norway

2 Department of Computer Engineering, Chung-Ang University, Seoul, Korea

In this paper, to consider the three phenomena for group recommendations, a novel method based on the analysis of social big data is proposed. First, an affective model is selected to represent the emotions and describe the definitions to discover and propagate the affective affinity (i.e., the emotional contagion) from the social big data. In this regard, a variety of similarity measures for emotions are taken into account to calculate the affective affinity between users. Then, the affective affinities are applied to the group recommendations by integrating with the social affinities that indicate the social influence in previous works. With respect to the final affinities for group users, conformity is also considered according to group characteristics, such as the majority of users and cohesion of the group.

1.1 Scenario and research questions

In this section, to clearly describe the motivation of this study and identify the research questions, a scenario of the movie domain is introduced. Figure 1 represents the list of movies watched by Alex, John, and Mary. In the diagram, the overlapping areas represent movies that had been seen by the users together. For instance, Alex and John viewed *Scream* and *[Rec]* together. From the Fig. 1, several phenomena regarding the *social influences* between the three users is visible. With respect to the movie logs of the individuals, Alex and Mary primarily watched movies from the horror and romance genres, respectively. In the case of John, he enjoyed the movies of various genres such as comedy, drama, and horror. According to the movies that the groups watched, Alex and John had primarily viewed horror movies, and John had typically watched romance movies with Mary. From the movie logs of individuals and groups, the following social influences are able to be inferred, when users watch a movie together:

- John follows the preferences of Alex.
- John and Mary prefer movies that are liked by Mary.

To illustrate the *emotional contagion*, Fig. 2 displays the affective reactions in the cases of watching alone and two users viewing together. In this regard, the inter-user relationships that are deeper than the user preferences might be inferred.

John felt fear when he watched the movie *[Rec]* alone. However, when John saw the *[Rec]* with Alex, the affective reaction of John changed from fear to surprise. Alex typically felt tension during horror movies. When Alex saw the movie *[Rec]* with John, he was more excited compared to when he viewed it alone. Alternatively, Mary and John respectively felt sadness and calmness from watching the movie *Titanic*. While, when John and Mary watched romance movies together, they felt romantic and loving emotions. In this regard, it is possible to presume affective relationships in terms of the emotional contagion, as follows:

- John relies on Alex when he watches a horror movie with Alex, since the affective reaction of John grows from his fear.
- John and Mary feel loving emotions when they watch the romance movies together. Thereby, it could be assumed that John and Mary are a couple.

Until now, a simple scenario has illustrated the social influence and emotional contagion according to user relationships. As an example, results from the movie *Avatar* that was watched by three users together is displayed in Fig. 1. In this case, it is not easy to determine the relationships using the logs of individuals and groups. The difficulty was due to the occurrence where Alex, John, and Mary watched the movie, *Avatar*, together which may have caused cohesion among the group. Particularly, the two aforementioned groups have stronger cohesion due to their relationships (i.e., friendship and love) than a group that consists of the average three users. Therefore, according to partner's preferences, movies are more easily selected for the groups. Meanwhile, for the three users, they may have chosen a movie which was generally liked by people [12]. In the field of social psychology, researchers have found that the group characteristics (e.g., majority and cohesion) are able to affect the opinions of the group users for group decision making, which is referred to as *conformity* [22].

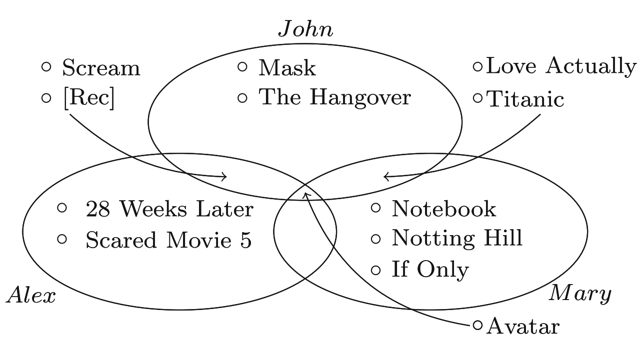

Fig. 1 Movie logs of Alex, John, and Mary

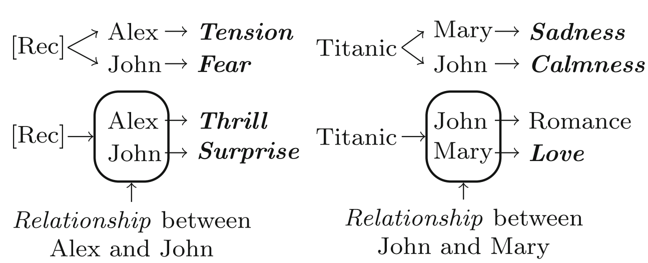

Fig. 2 Emotional reactions between two users

To apply these three phenomena in group recommendations, the following research questions are raised with respect to this scenario.

- How can affective affinity (i.e., emotional contagion) be obtained from affective reactions?
- How can the group characteristics be reflected among group users in relation to conformity?
- How can social affinity, affective affinity, and group characteristics be integrated to recommend optimal items for a group?

1.2 Outline

The outline of this paper is as follows. Section 2 reviews the existing studies that are related to the proposed method. In Section 3, the previous works that are related to the social affinity are introduced, and the descriptions of the basic definitions and a variety of similarity measurements are provided for affective affinity. Section 4 illustrates how the affective affinity, social affinity, and group characteristics integrated into an affective social network can be applied in group recommendations. The experiments were conducted to evaluate the similarity measures and group recommendations in Section 5. Finally, in Section 6, the conclusion and limitations are illustrated.

2 Related work

2.1 Emotion-based recommender system

In the field of individual recommender systems, several researchers have applied explicit or implicit emotions in modeling items (users) to improve the quality of recommendations [23, 24]. Zheng et al. used emotions for context-aware splitting and differential context modeling [21]. As a result, the emotional contexts improved the performance of the recommendation in terms of the Root Mean Square Error (RMSE). Furthermore, *EndEmo*, which denotes users' emotions after they view a movie, is a top factor for improving the recommendation quality. Thereby, the *EndEmo* from the subjects for our experiments were collected.

Similar results were also obtained regarding the content-based recommendations (CBR) [25]. To model users and items, Tkalcic et al. compared three data types as follows: generic metadata (genre and watching time), explicit emotions of the users, and implicit emotions that were tracked by the facial expressions of users. The best performance was derived from the explicit emotions. These results are the reason why the explicit affective reactions were obtained from subjects to generate affective relationships between the users.

Moshfeghi and Jose proposed a collaborative filtering (CF) method by integrating semantic and emotional information [26]. They extracted 24 emotions from plot summaries and reviews of movies using the natural language processing technique. Their results demonstrate that the cold-start problem, which is one of the drawbacks for CF, could be alleviated by using emotions. Additionally, the inter-relationships between the different features (i.e., movie feature, semantic, and emotion) were discovered in their work. In the proposed method, 30 emotions are used to find the inter-relationships between the users in a group. Chakraverty and Saraswat proposed a method for an emotion-based cross domain recommendation (CDR) [27]. They extracted six emotions from textual content such as reviews, blogs, and comments, using emotion lexicons. By using the emotions, emotion-profiles of items and users in both a movie and book domain was generated. The profiles were applied to the CF method to recommend movies or books. Contrary to their approach, the emotions from subjects are used to extract affective affinities.

In spite of the many positive results of emotions for individual recommendations, to the best of our knowledge, emotions have been passively applied to GRSs. Especially, as mentioned in Section 1, user rating has been considered as the emotional satisfaction of group users instead of their real emotions [11, 12]. There have been attempts to apply real emotions into GRSs. However, these approaches focused on the user interface rather than applying the emotions to the GRS [28, 29].

2.2 Group recommender system

The HappyMovie, which was implemented as a Facebook application, was proposed to recommend a movie to the user groups [7]. The application asked users about the Thomas-Killman Instrument (TKI) Test and ratings of the movies. Additionally, the trust between the users was calculated by the degree of joining the same events (i.e., watching a movie together). In our previous study, the Facebook application was also implemented to collect data. With respect to social influence, social affinity was calculated based upon the movie logs of users and features of the movies instead of the total number of same logs.

Sanchez et al. also introduced a delegation-based prediction method to consider the social behavior within a group [5]. This method uses various data (e.g., size and structure of the groups, user personalities, and trust between users) for social influence and conformity. However, the emotional contagion between the users was not considered, and the TKI test became a burden for users. Guo et al. proposed a group-recommendation method based on group-user modeling using relationships, personality, expertise, susceptibility, and intimacy, among other factors [30]. Although their method considered various aspects for the modeling group, emotional contagion was also not taken into account; so, they also used the TKI Test. Christensen and Schiaffino proposed a GRS based on the social influence [31]. They calculated the influence from three social factors as follows: trust as closeness, social similarity, and social centrality as reputation. Christensen et al. also introduced the Hermes system based on social relationships to recommend the items for a group in the tourism domain [32]. The system uses content-based and demographic filtering to solve the cold-start problem of the CF-based recommendations. In contrast, the proposed method considers not only the social influence but also the emotional contagion and conformity. The emotional contagion and conformity were applied with the social influence in a GRS by Masthoff and his colleagues [11, 12]. From the aspects of the sociological and psychological theories, they focused various factors (e.g., mood, retrospective emotions, expectations, and emotions) over time. These studies also demonstrated that user emotions are able to be influenced by other users in a group. In contrast, this study applies not only satisfaction but also other emotions. Furthermore, the group characteristics are reflected for the conformity.

2.3 Affective model

Emotions are primarily represented by three main models as follows: (i) the categorical model, (ii) the dimensional model, and (iii) the circumplex model [21]. The categorical model assumes the existence of a limited set of distinct emotional categories. There is no unanimity for universal emotions; however, the categories proposed by Ekman (i.e., happiness, anger, sadness, fear, disgust, and surprise) has been popularly used [33]. Contrarily, the dimensional model describes each emotion as a point in a continuous multidimensional space. Each dimension also represents an emotional degree. The dimensions that are used most frequently are valence (i.e., pleasure-displeasure and positive-negative) and arousal (i.e., arousal-sleepiness and high-low activation). Additionally, the circumplex model maps the basic emotions into the Valence, Arousal, and Dominance (VAD) space. The model in which the inter-correlations among the variables are represented by a circle has a shorter history than the others. Nevertheless, it has obtained empirical support [34].

Recently, a 12-point affect circumplex model of the "core affect" was introduced to represent emotion and mood [35]. Furthermore, the circumplex model is based on 30 emotions and moods which were summarized from seven articles that are frequently cited by researchers. The role of a trustworthy affective model is important in the proposed approach, since the affective affinity for the emotional contagion is calculated based on the model. Therefore, the affective models are carefully reviewed and the circumplex model proposed by Yik et al. is selected. The circumplex model provides the confidence values of the radius and the length of the affective states. Various elements of the model allow its application in a variety of similarity measures. For a similarity measure according to angle, the radius of the emotions is directly used. In addition, the radius and length of affective states are converted to two-dimensional coordinates to apply to the other distance measures.

3 Affinity between group users

In this section, the definitions for social and affective affinity are introduced. The affinities respectively represent the social influence and emotional contagion between users, as mentioned in Section 1.1.

3.1 Social affinity

In our previous studies, as shown in Fig. 1, it was assumed that preferences of a user are changed according to a person who watches a movie together [18, 36]. For instance, a wife might enjoy a romantic movie with her husband while she usually watches animated movies with her child. This type of social influence is known as social affinity. The social affinity between two users is calculated by the movie similarities based on the Overlap coefficient. In this regard, a social affinity network which contains the user nodes and affinity edges are generated. Moreover, to alleviate the sparsity of the social network, affinities are propagated based on additional information such as the distances between the nodes and edge directions. Lastly, maximum affinities are selected as the final affinities for their group, in order to avoid information loss. The differences between the previous study and this paper are as follows:

- a variety of similarity measures between the two affective reactions for the emotional contagion,
- a consideration of the group characteristics to address the conformity, and
- the integration of the social and affective affinities into an affective social network.

This section starts with a description of several definitions from the previous work, since the basic assumptions are similar to the previous works.

Definition 1 (Social Affinity between Two Users) Given average similarities $S^s_{i,k}$ and $S^s_{j,k}$ of the movies which are belong to individual and common histories of the users u_i and u_j, the social affinity $A^s_{i,j}$ of the user u_i for the user u_j is defined as follows:

$$A^s_{i,j} = \begin{cases} S^s_{i,k}/(S^s_{i,k} + S^s_{j,k}) & \text{if } k \neq 0, \\ 0.5 & \text{otherwise}, \end{cases} \tag{1}$$

where k is the number of movies in common history.

For instance, as mentioned in Section 1.1, the social similarity of John for Mary is calculated as the average similarity between three movies (i.e., *Titanic*, *The Hangover*, and *[Rec]*) and two movies (i.e., *Love Actually* and *Titanic*) which belong to John's individual history and John and Mary's common history, respectively. As a result, the range of social affinity is from 0 to 1, and the sum of the affinities between the two users is 1. Particularly, their social network is able to be represented as a directed graph $G(u, a)$ in which u and a are users and their affinities, respectively. In this regard, the network should be propagated by using enriched social data (i.e., indirect connection and direction) on the network, since the initial social network $G(u, a)$, which consists of direct connections, is sparse.

Definition 2 (Propagation of Affinity Network) Given a group's social network $G(u, a)$, the affinity $A'_{i,j}$ of user u_i for user u_j is refined as follows:

$$A'_{i,j} = A_{i,j} + \sum\nolimits_{p=2}^{P} \prod\nolimits_{h=1}^{H} A_h / P, \tag{2}$$

where p and h represent the number of edges and an edge's index in paths to link from user u_i to user u_j, respectively.

Although a density of the initial social network is scant, it becomes a complete directed graph by Definition 2. It indicates that final user affinities for the group should be chosen.

Definition 3 (Selection of an User Affinity for Their Group) Given a set of the social affinities $A'_{i,l}$ of user u_i for the other users u_l in their group, an final affinity A''_i of the user u_i for the group is defined.

$$A''_i = \max_{A'_{i,l} \in A'} A'_{i,l}, \tag{3}$$

where A' indicates a set of the edges (i.e., social affinities) in the group $G(u, a)$.

3.2 Affective affinity

According to the scenario in Section 1.1, we describe basic definitions for the affective affinity.

Definition 4 (History) History $\mathcal{H}$ includes a set $\mathcal{I}$ of items and a set $\mathcal{U}$ of users who saw the items. The history is expressed by:

$$\mathcal{H} = \langle \mathcal{U}, \mathcal{I} \rangle. \tag{4}$$

In this regard, items are able to be consumed by a single user or several users at one time. Therefore, the histories should be separated for the individual user and group users. A shared history represents that several users watched items together. For instance, in Fig. 1, the shared history of Alex and John contains the *Scream* and *[Rec]* movies. In this regard, an individual history of John includes *Titanic*, *The Hangover*, and *[Rec]* movies.

Property 1 (Shared History) Given the two users $u_i, u_j \in \mathcal{U}$ and their histories $h_i, h_j \in \mathcal{H}$, a shared history $sh_{i,j} \in \mathcal{H}$ consist of items $i = \{x | x \in h_i \cap h_j \wedge x \in \mathcal{I}\}$ that are watched by both the users u_i and u_j together.

Property 2 (Individual History) Given the two users $u_i, u_j \in \mathcal{U}$ and their histories $h_i, h_j \in \mathcal{H}$, an individual history of the user u_i for the user u_j has items $ih_i = \{x | x \in h_i \wedge x \notin sh_{i,j}\}$.

Users are sometimes able to log their emotion (e.g., label, tag, and feedback) that they feel after consuming the items. In this paper, the explicit emotion, which is represented by users, is named an affective reaction.

Definition 5 (Affective Reaction) Given the user $u_i \in \mathcal{U}$ and the item $i_m = \{x | x \in h_i \wedge \in \mathcal{I}\}$, the affective reaction of user u_i about the item i_m is expressed as $e_{i,m} \in \mathcal{E}$. Also, a set of affective reactions $\mathcal{E}$ is commonly included in $\mathcal{H}$ for emotion-based recommender systems.

Based on the affective models, the similarities between the affective reactions between the two users are able to be calculated using a variety of measures. The various similarity measurements are mentioned in Section 3.3 and evaluated for affective reactions in Section 5.2.

Definition 6 (Affective Similarity) Given the two affective reactions $e_n, e_m \in \mathcal{E}$ regarding the items i_n and i_m, the affective similarity between the reactions e_n and e_m is expressed as $S^a_{n,m}$.

From the similarities between the affective reactions, an affective affinity between the two users is able to be calculated according to Definition 1, like social affinity.

Definition 7 (Affective Affinity) The affective affinity expresses the emotional contagion which indicates the influence of the emotion toward an item by the relationship between users in their group $\mathcal{G}$. Therefore, group information is also contained with a set $\mathcal{E}$ of affective reactions in history. In this regard, the history $\mathcal{H}$ is represented as follows:

$$\mathcal{H} = \langle \mathcal{U}, \mathcal{I}, \mathcal{E}, \mathcal{G} \rangle \tag{5}$$

The affective affinity has a non-commutative property like the social affinity.

Property 3 (Non-commutative) Given the affective affinity $A^a_{i,j}$ and $A^a_{j,i}$ which represent the affinity of the user u_i to the user u_j and the affinity of the user u_j to the user u_i, respectively, the two affinities share the following property:

$$A^a_{i,j} + A^a_{j,i} = 1. \tag{6}$$

Additionally, the affective affinity of the user u_i is larger for the user u_j when the $A^a_{i,j}$ is closer to 1, and vice versa. If the affinity of the value is 0.5, the effects of the users u_i and u_j are the same in terms of the affective reaction. According to Definition 2, other networks between the users are also able to be constructed and propagated with the affective affinities as edges. The next steps for group recommendations will be introduced in Section 4.

3.3 Measuring the similarity between affective reaction

In this section, several conventional measurements are introduced for calculating a similarity between two affective reactions on the selected affective model. The Cosine method measures a similarity between two of the non-zero vectors, and it uses an angle between the vectors. If the two vectors share the same orientations, their cosine similarity (COS) is 1. Two vectors at an angle of 90° have a similarity of 0, and a similarity between two diametrically opposed vectors is -1.

Alternatively, the selected circumplex model is able to be converted into a coordinate system. In this regard, the Euclidean distance (EUD) calculates the ordinary length between two points in the Euclidean space.

Meanwhile, the Mahalanobis distance (MAD) relates to a point P and a distribution D. It is a multidimensional generalization of the idea for measuring the multiple standard deviations between the P and the mean of D. The MAD is defined as a dissimilarity measure between the two vectors $\mathbf{A} = (a_1, a_2, ..., a_n)^T$ and $\mathbf{B} = (b_1, b_2, ..., b_n)^T$ of the same distribution where the covariance-matrix S is $\sqrt{(\mathbf{A} - \mathbf{B})^T S^{-1} (\mathbf{A} - \mathbf{B})}$.

The Euclidean Norm (EUN) is a natural extension of the notion of a vector norm to matrices. On an n-dimensional Euclidean space R^n, the intuitive notion of the length of the vector $\mathbf{x} = (x_1, x_2, ..., x_n)$ is captured by $\|\mathbf{x}\| := \sqrt{x_1^2 + x_2^2 + ... + x_n^2}$. Thus, the distance between the vectors $\mathbf{A} = (a_1, a_2, ..., a_n)$ and $\mathbf{B} = (b_1, b_2, ..., b_n)$ is expressed as $\big| \|\mathbf{A}\| - \|\mathbf{B}\| \big|$.

To generate the affective affinity, the distance D should be converted into the similarity S using the following equation:

$$S(D) = 1/(1 + D). \tag{7}$$

The similarity is normalized to a value between 0 and 1. However, the range of the Cosine similarity is from -1 to 1. Therefore, the value S_C of the cosine similarity is normalized as follows:

$$S(S_C) = (1 + S_C)/2. \tag{8}$$

The results of the various similarity measures are used to generate the affective affinities between two users by using Eq. 1. Then, the affinities are combined with the social affinity for the weights of users in their group. In next section, we explain the approach for integrating the social affinity and affective affinity for a group recommendation.

4 Affective social network for group recommendation

This section demonstrates that the social and affective affinities are integrated to form an affective social network for users in a group and describes how the characteristics of a group are applied to a GRS.

4.1 Generating affective social network for group

The social affinity of two users represents the social influence and is calculated by their social network and descriptive similarities of the movies in their histories [36]. Similarly, the affective affinity between the two users indicates the emotional contagion and is generated based on their social network and affective similarities for movies in their histories. As a result, the social affinities generate a social network; likewise, the affective affinity between the users is included with an affective network. Then, the graphs are then propagated by Eq. 2 to develop a complete directed graph, which is represented as an adjacent matrix of the affective affinities between the users. To simplify the process, two matrices for the social and affective affinities are integrated into an affective social network. From the network, given the social affinity $A^{s''}_{i,j}$ and affective affinity $A^{a''}_{i,j}$ between the users u_i to u_j, the unified affinities $A^*_{i,j}$ of the users u_i to u_j is calculated as follows:

$$A^*_{i,j} = \alpha \times A^{s''}_{i,j} + (1 - \alpha) \times A^{a''}_{i,j}, \tag{9}$$

where the unification constant α is empirically set as 0.7. The affinities are relationships of inter-users rather than regarding their group. Therefore, to apply the affinity into GRS, it should be converted for the group. In the field of social psychology, researchers have shown that a small group is more cohesive than large groups [37]. Regarding this cohesion, we consider three methods (i.e., maximum, average, and minimum selections) according to the group size. Accordingly, given an adjacent matrix A^* of the unified affinities $A^*_{i,j}$ between the users u_i, u_j, the final affinity of the i^{th} user $A^{*\prime\prime}_i$ is calculated as follows:

$$A^{*\prime\prime}_i = \begin{cases} \min_{A^*_{i,l} \in A^*} A^*_{i,l} & \text{if } |G| < 4 \\ \max_{A^*_{i,l} \in A^*} A^*_{i,l} & \text{if } |G| > 5 \\ \mathrm{avg}_{A^*_{i,l} \in A^*} A^*_{i,l} & \text{otherwise,} \end{cases} \tag{10}$$

where $|G|$ indicates the group size, and l represents the other l^{th} user in the group. The final affinities are normalized and applied to a CBR as user weights. Moreover, in terms of the GRS, a weight-based aggregation strategy is used for our approach. The recommendation approach is considered because of complexity. The proposed method assumes that the similarities between movies (and emotions) are already saved into a database. Therefore, to calculate the affinities, the computational complexity of the proposed method is $O(N^2)$, where N indicates the number of users in a group. Meanwhile, the computational complexity of the memory-based CF with overall N users and M items is $O(NM)$. Therefore, the proposed method is more appropriate than combining with CF for big data environments, since the N is significantly smaller than the numbers of overall users.

According to [12], existing GRSs have tried to use two influences (i.e., normative influence and informational influence) with respect to the conformity. With aspects of social psychology, Stasser and Davis proposed the influences based on the majority [38]. Thereby, in this paper, the majority influence that indicates the conformity for group users is reflected. To obtain the majority between the similarities, a standard deviation is used. Given the average similarity $\langle S_k \rangle$ of the group users for an k^{th} candidate movie and the standard deviation σ_s of the users' similarities, the range of the majority is decided by $\langle S_k \rangle \pm \sigma_s$. In this regard, the majority size N is the number of users, and the final similarity $S^g_{i,k}$ of the user u_i for a candidate movie m_k is calculated as follows:

$$S^g_{i,k} = (\langle S_k \rangle - S_{k,i}) \times \beta(N/(1+N)), \tag{11}$$

where $S_{k,i}$ indicates the similarity between the movie m_k and the movies in the history h_i of the user u_i, and β is an influence constant with the empirical value of 0.44. The first term represents the difference between the average similarity of the group users and the i^{th} user's one as their opinions, and the second term denotes the effects of the conformity according to the majority size.

4.2 Group recommendation process

Algorithm 1 displays the processes of applying three phenomena (i.e., social influence, emotional contagion, and conformity) to recommendation for group users.

Algorithm 1 Group Recommendation Process

input : $\mathcal{H}$ of k users on a group g_l, n candidate movies
output: List M of movie to recommend for group g_l

for $i \leftarrow 1$ **to** k `// k is the number of group users`
do
  Set adjacent matrices A^s and A^a between user u_i and u_j by Eqs. 1 and 5;
  for $j \leftarrow 1$ **to** k **do**
    if $sh_{i,j} \neq \emptyset$ *and* $i \neq j$ **then** Calculating $A^s_{i,j}$ and $A^a_{i,j}$;
    else $A^s_{i,j}$and$A^a_{i,j} \leftarrow 0.5$;
  end
end
Interpolating matrices A^sandA^a by Eq. 2;
Calculating unified affinity matrix $A^*_{i,j}$ by Eq. 9;
Selecting final affinities of users by Eq. 10;
for $i \leftarrow 1$ **to** k **do**
  Set a matrix $M_{i,j}$ of similarity between candidate movie for movies in h_i ;
  for $j \leftarrow 1$ **to** n **do**
    $M_{i,j} \leftarrow \sum_{k=1}^{p} S(m_i, m_q)/p$; `// p is the number of movies in` h_i
  end
end
Generating similarity matrix $S^g_{i,n}$ by Eq. 11;
A vector $\mathbf{M} \leftarrow A^{*\prime\prime} \times {S^g_{i,n}}^T$; `// Score of n candidate movie for the group` g_l
Descending Sort Order M;
Return M;

Given the histories of the users in a group g_l and n candidate movies, the algorithm outputs a list M of the movies for the group g_l. First, the adjacent matrices A^s, A^a of the social affinity $A^s_{i,j}$ and the affective affinity $A^a_{i,j}$ between the pair of the users u_i and u_i of the group g_l are generated using Eqs. 1 and 5. And the square matrices A^s and A^a then are interpolated using Eq. 2. The unified matrix of the affective social network A^* is calculated using Eq. 9, and the final affinity $A^{*\prime\prime}_i$ of user u_i should be selected by Eq. 10 according to the group size. Next, the average similarities between the candidate set and the movies on user histories are calculated. The similarities

Table 1 Steps and activities for the experiments

Step	Activity
1	Selecting movies and discovering groups from users who had registered in our website.
2	Collecting affective reactions of the users for the movies using the website.
3	Calculating affective affinities using various similarity measurements.
4	Asking influences between the users when they watch movies.
5	Selecting ten candidate films and ranking them by the users for their group.
6	Generating a film lists with ranks though methods which will be assessed.
7	Comparing the methods by using similarity between the two film lists ordered.

are manipulated to reflect the conformity using Eq. 11. To rank the candidate movies, a multiplication of the adjusted similarity and the affinities of each user of the group are performed. As a result, a movie list M that is sorted in descending order is obtained.

5 Experimentation

5.1 Experimental environment

In this section, the experimental environment such as the dataset, evaluation protocol, measure, baseline methods, survey, and recommendation is explained. With respect to the research questions, several measurements were compared in terms of the influences between the users. In this regard, a survey was conducted with the 16 pairs of users to evaluate the proposed affective affinity between two users. Moreover, the similarities between the two movie lists, which were made by subjects, and the proposed method were analyzed by comparing them with recent GRSs. To mitigate any uncertainty in the survey, the experiment is described according to the following Table 1.

For the experiments, the users were asked to provide their affective reactions to movies, state the influences between them, and rank movie lists for their group. To invite subjects, an e-mail was sent to users who had registered in our "MyMovieHistory" website[1]. From the users who had positively replied to the e-mail, 34 groups of five types, which were composed of 2, 3, 4, 5, and 6 persons, resulted, and the numbers of the groups were 16, 5, 4, 5, and 4, respectively. In our experiment, the groups were formed based upon the friend information that was located on Facebook and the shared history between the users. The 167 subjects were undergraduate and graduate students of Chung-Ang University and members of society. It was possible for the subjects to be included in several groups. Then, films that had recorded high rankings at the US box office from the IMDb[2] were selected. From 22 main genres, 440 movies that had been released during 2010-2016 were chosen with the consideration of the subjects' ages which were from 20 to 30 years. Overlapping movies were removed, and, finally, list of 50 movies were selected. Therefore, the conducted survey was composed of the 50 movies and subjects were asked to input their multi-affective reactions for the movies as step 2. The reaction degrees were collected with the integers from 1 to 5 (i.e., 5 means strong feeling). The descriptive features of the movies with posters and definitions of the 30 emotions with emoticons were provided to help the respondents while completing the survey. As a result, 2591 histories of the 167 users for the 50 movies were collected in total. From the history data, the affective affinities between two users for their group were generated based on various similarity measurements (i.e., similarity of the Cosine, Euclidean, Mahalanobis, and Euclidean Norm) and were propagated by according to their group, as step 3. Using the influences which had been provided by the subjects in step 4, the affective affinities were evaluated in terms of the first research question. Next, in step 5, a new list that included 10 movies was made as the recommendation candidate, and the subjects were asked to rank the list with their preferences with the consideration of their group. In step 6, the movie sequence of the list was also arranged by various methods including the proposed method. With respect to the second research question, the similarity between the two sorted lists was calculated using Kendall's τ-b correlation with basic aggregation strategies for a GRS. Additionally, regarding the last research question, the baseline methods were compared with the proposed method using the correlation in terms of the social influence, emotional contagion, and conformity.

The Kendall's tau coefficient is commonly used as a rank-correlation coefficient; it measures the similarity between the ordinal association of the two lists. There are same-ranked movies in the lists that were made by the users. Therefore, the Kendall's τ-b was chosen instead of Kendall's τ-a, since the Kendall's τ-b is able to reflect the ties from the same ranked items in a list. The τ-b range is from -1, which means a perfect inversion, to 1 (i.e., 100% agreement). A value of zero indicates the absence

[1]MyMovieHistory, https://ke.mymoviehistory.kr

[2]IMDb, https://www.imdb.com/genre/

of a relation. Given a ranked list of n items containing n_c concordant pairs and n_d discordant pairs, the τ_b coefficient is defined as follows:

$$n_o = n(n-1)/2,\ n_1 = \sum_i t_i(t_i-1)/2,$$
$$n_2 = \sum_j u_j(u_j-1)/2,\ \tau_b = \sqrt{\frac{(n_0-n_1)(n_0-n_2)}{(n_c-n_d)}}, \quad (12)$$

where t_i and u_j represent the number of tied values in the i^{th} group of ties for the first quantity and the number of tied values in the j^{th} group of ties for the second quantity, respectively.

5.2 RQ1: affective affinity with similarity measure

In this section, to evaluate various measurements for affective similarity, the affective affinities between them were compared. From the 16 groups that consisted of two users, there were 10 couples, and whereas, five of the groups included two females, the last group contained two males. To obtain the real user influences, an average of the numerical values, which were collected from two users, had been calculated. In this regard, the mean squared error (MSE) which measures the average of the squares of the errors between the real values and observations was used. Table 2 provides the results of the comparison regarding the MSE values of the various similarity methods.

As a result, the COS and the MAD measurements performed the best in this experiment. Therefore, the two measures have been determined as measurements to use in the next experiments.

To evaluate the effect of affective affinity for a group that consists of two users, the COS and the MAD measures also were compared with other methods. Additionally, as a realization from the previous paper, we identified that it was a necessity to consider balanced and unbalanced relationships for a group recommendation. The former relationship means that one person considerably relies on another. On the contrary, the latter relationship denotes that two users share a similar value. If the influence between two users is greater than or equal to 0.7, their relationships were considered unbalanced. There were four balanced groups and 12 unbalanced groups among the 16 groups that consisted of two users.

Table 2 Comparison of similarity measurements

	COS	EUD	MAD	EUN
MSE	0.007†	0.041	0.008†	0.081

†indicates that a method statistically outperforms the others based on the paired-sample t-test (0.01 confidence coefficient)

Therefore, the proposed methods using the COS and MAD were compared with both of the representative recommendation strategies. The aggregation strategies were based upon the average strategy (AS) and least-misery strategy (LMS). Another case which only considered the social affinity (SOA) was also compared together. For the two aggregation strategies, ten movies were sorted using the rating scores that were collected by the survey. Table 3 displays the comparison results of the Kendall's τ-b coefficient between the two ranked lists which were obtained by the user groups and the methods.

In the case of the balanced group, the MAD and the LMS are superior to other methods. In terms of whether or not the affective affinity was applied, the COS and MAD outperformed the SOC, which indicates that the consideration of the affective affinity was able to improve the recommendation quality for a group. In particular, as mentioned in Section 1.1, it demonstrates that the affective affinity was able to reflect a deeper relationship than social affinity, since the balanced groups had similar influences between users in terms of sociality. With respect to the unbalanced group, the MAD was still the best, and the other proposed methods were superior to the AS and LMS. These findings indicate that the proposed methods were able to efficiently reflect the influences between the users in a group in order to recommend items. Especially, the MAD was the best for both the balanced and unbalanced groups. From these findings, it is demonstrated that the affective affinity was able to support the social affinity for the group recommendation.

5.3 RQ2: group characteristic for conformity

In this section, with respect to conformity, the proposed approaches and two aggregation methods are compared in terms of group size. The affective affinity was calculated based on the MAD, since its performance was the best in previous experiments. The proposed methods are separated into three approaches as follows: (i) applying only unified affinity without majority and cohesion (AOS), (ii) applying the unified affinity and the cohesion without majority (AOC), and (iii) applying all the factors (i.e., the unified affinity, the cohesion, and the majority (ACM).

Table 3 Comparison for Balanced and Unbalanced relationships

Type	SOC	COS	MAD	AS	LMS
Balance	0.612	0.647	0.690†	0.602	0.694†
Unbalance	0.657	0.643	0.633	0.481	0.527

†indicates that a method statistically outperforms the others based on the paired-sample t-test (0.05 confidence coefficient)

Table 4 Comparison of recommendation results for group size

Size	AOS	AOC	ACM	AS	LMS
2	0.642	0.511	0.662	0.527	0.562
3	0.631	0.552	0.660	0.501	0.551
4	0.625	0.627	0.659	0.471	0.547
5	0.610	0.626	0.663	0.430	0.542
6	0.613	0.627	0.658	0.401	0.543

Table 4 represents the results of the Kendall's τ-b coefficient between the methods.

As a result, the AOS and ACM were superior to those of the aggregation methods. In particular, the ACM outperformed the others. In terms of the group size (i.e., cohesion), the performance of the AOS was better than the AOC in the case of small groups. The AOS decreased performances as the group size increased, while the performance of the AOC steeply increased. However, the increase due to the AOC ceased over four persons. These results are similar to Asch's experiment, which demonstrated that conformity pressures peak once the majority reaches approximately four or five persons [39]. For the aggregation methods, the performances of the AS significantly decreased with more users in a group; while, the decline degree of the LMS was slowly small and then stopped for large groups. The difference between the LMS and AS is that the LMS removed the worst items for recommendation. Therefore, the LMS was able to relatively satisfy the users by avoiding the worst cases even among large groups. Similar to the comparison with the AOS and AOC, the ACM outperformed, which signifies that the majority and cohesion contribute to the group recommendation for the proposed method in terms of conformity. Even if the group sizes increase, the ACM sustains similar performances. These results demonstrates that the proposed method can overcome the decrease in group sizes through three factors (i.e., the social influence, emotional contagion, and conformity).

5.4 RQ3: three influences for group recommendation

In this section, to evaluate the effect of the proposed method for the group recommendation, two social-based and one emotion-based GRSs were compared [11, 31, 32]. These systems were marked as SGR, SIGR, and EGR, respectively. The emotion variation over time was not considered, since the recommended lists were not sequences of items for continuous viewing. Therefore, for the conformity, informational and normative influences were applied in the EGR. For emotional contagion, assimilation and contrast were also used with respect to the expectations of the users.

Table 5 Comparison of recommendation results

Size	ACM	SGR	SIGR	EGR
2	0.662†	0.634	0.657	0.663†
3	0.660	0.601	0.610	0.631
4	0.659	0.589	0.591	0.620
5	0.663	0.558	0.588	0.624
6	0.658	0.538	0.570	0.622

†indicates that a method statistically outperforms the others based on the paired-sample t-test (0.01 confidence coefficient)

As displayed in Table 5, the proposed method generally demonstrated a higher performance than the SGR and SIGR. With respect to the social-based GRSs, the drop width of the SIGR was smaller than the SGR. It indicates that the consideration of personality alleviate the decline in the recommendation quality according to group sizes, in terms of the conformity. In addition, although the results of the EGR declined according to size increase, for groups containing two users, the performance of the EGR was similar to the ACM. Moreover, the shape of the degree of decrease followed a parabola. On the contrary, the performances of the ACM were retained even with increasing group size, which indicates that the proposed method can handle the growth in group size.

6 Conclusion and future work

Social big data is an emergent issue currently, especially for recommender systems. In particular, with respect to social big data, various data mining techniques have been applied in group recommender systems. However, three social phenomena (i.e., social influence, emotional contagion, and conformity) have been not applied enough in existing studies. In this paper, a novel method for a group recommendation based on the affective social network which is extracted from social big data is proposed. To evaluate the proposed method, the 2591 histories of the 167 users for 34 groups of various sizes were obtained using the website "MyMovieHistory". With respect to the emotional contagion, the proposed methods based on several similarity measures were compared with two aggregation strategies. The experiment results demonstrated that the proposed method was best with the MAD, especially for the unbalanced group. Regarding the comparison with the state-of-the-art GRSs, the proposed method exhibited superior performance. Contrary to other methods which have a performance decline with the growth of group sizes, the proposed method maintained its performance. In particular, the method was able to be applied well for recommending items which contained emotional content (e.g., movie, music, and art) because of the consideration of the affective affinity for a variety of the emotions.

However, in this work, the limitations are as follows.

- Only one affective model was applied as the circumplex model, which was composed of the angles and lengths of 30 emotions and moods.
- In terms of the experimental magnitude, a small number of groups were evaluated.
- Additionally, explicit emotions that were manually generated by subjects were used for the experiments.

As mentioned in [25], although explicit emotions are able to improve recommendation quality more than implicit emotions, we believe that there are additional opportunities to obtain the implicit emotions of high quality from enriched social big data. Therefore, a future study will be focused on discovering the implicit emotions from various contexts in a big data environment. Moreover, an experiment for which various affective models are used will be fulfilled to obtain a proper model for the group recommendation with a larger dataset.

Acknowledgements This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (NRF-2017R1A2B4010774).

References

1. Lee OJ, Jung JJ, You E (2015) Predictive clustering for performance stability in collaborative filtering techniques. In: Proceedings of the 2nd IEEE international conference on cybernetics, CYBCONF. Gdynia, Poland, pp 48–55
2. Gandomi A, Haider M (2015) Beyond the hype: big data concepts, methods, and analytics. Int J Inf Manag 35(2):137–144
3. Lee OJ, Hong MS, Jung JJ, Shin J, Kim P (2016) Adaptive collaborative filtering based on scalable clustering for big recommender systems. Acta Polytechnica Hungarica 13(2):179–194
4. Saranya KG, Sadasivam GS (2017) Personalized news article recommendation with novelty using collaborative filtering based rough set theory. Mobile Networks and Applications 22(4):719–729
5. Sánchez LQ, Recio-garcía JA, Díaz-Agudo B (2014) An architecture and functional description to integrate social behaviour knowledge into group recommender systems. Appl Intell 40(4):732–748
6. Piccialli F, Jung JE (2017) Understanding customer experience diffusion on social networking services by big data analytics. Mobile Networks and Applications 22(4):605–612
7. Sánchez LQ, Recio-garcía JA, Díaz-agudo B (2011) User satisfaction in long term group recommendations. In: Proceedings of the 19th international conference on case-based reasoning, ICCBR, vol 6880. Springer, London, Lecture Notes in Computer Science, pp 211–225
8. Piliponyte A, Ricci F, Koschwitz J (2013) Sequential music recommendations for groups by balancing user satisfaction. In: Proceedings of the 21st conference on user modeling, adaptation and personalization, Rome
9. Tommasel A, Corbellini A, Godoy D, Schiaffino SN (2015) Exploring the role of personality traits in followee recommendation. Online Inf Rev 39(6):812–830
10. Zhang Y (2016) Grorec: A, group-centric intelligent recommender system integrating social, mobile and big data technologies. IEEE Trans Serv Comput 9(5):786–795
11. Masthoff J, Gatt A (2006) In pursuit of satisfaction and the prevention of embarrassment: affective state in group recommender systems. User Model User-Adap Inter 16(3-4):281–319
12. Masthoff J (2011) Group recommender systems: Combining individual models. In: Recommender systems handbook, Springer, pp 677–702
13. Chen M, Mao S, Liu Y (2014) Big data: a survey. Mobile Networks and Applications 19(2):171–209
14. Wu X, Zhu X, Wu G, Ding W (2014) Data mining with big data. IEEE Trans Knowl Data Eng 26(1):97–107
15. Orgaz GB, Jung JJ, Camacho D (2016) Social big data: recent achievements and new challenges. Information Fusion 28:45–59
16. Zhang Y, Tu Z, Wang Q (2017) Temporec: temporal-topic based recommender for social network services. Mobile Networks and Applications 22(6):1182–1191
17. Kompan M, Bieliková M (2014) Social structure and personality enhanced group recommendation. In: Proceedings of the 22nd conference on user modeling, adaptation and personalization, Aalborg
18. Hong M, Jung JJ, Piccialli F, Chianese A (2017) Social recommendation service for cultural heritage. Pers Ubiquit Comput 21(2):191–201
19. Kompan M, Bielikova M (2014) Group recommendations: Survey and perspectives. Computing and Informatics 33(2):446–476
20. Tkalcic M, Chen L (2015) Personality and recommender systems. In: Recommender Systems Handbook, Springer, pp 715–739
21. Zheng Y, Mobasher B, Burke R (2017) Emotions in context-aware recommender systems. In: Emotions and personality in personalized services - models, evaluation and applications, Human-Computer Interaction Series, Springer, pp 311–326
22. Bond R (2005) Group size and conformity. Group Process Intergroup Relat 8(4):331–354
23. Han B, Rho S, Jun S, Hwang E (2010) Music emotion classification and context-based music recommendation. Multimedia Tools and Applications 47(3):433–460
24. Tkalcic M, Carolis BD, de Gemmis M, Odic A, Kosir A (2017) Introduction to emotions and personality in personalized systems. In: Emotions and personality in personalized services - models, evaluation and applications, Human-Computer Interaction Series, Springer, pp 3–11
25. Tkalcic M, Odic A, Kosir A, Tasic JF (2013) Affective labeling in a content-based recommender system for images. IEEE Trans Multimedia 15(2):391–400
26. Moshfeghi Y, Jose JM (2011) Role of emotional features in collaborative recommendation. In: Proceedings of the 33rd European conference on IR research, ECIR 2011, vol 6611. Springer, Dublin, Lecture Notes in Computer Science. pp 738–742
27. Chakraverty S, Mala S (2017) Review based emotion profiles for cross domain recommendation. Multimedia Tools and Applications 76(24):25827–25850
28. Chen Y, Pu P (2014) Designing emotion awareness interface for group recommender systems. In: Paolini P, Garzotto F (eds) Proceedings of the International Working Conference on Advanced Visual Interfaces, AVI 2014, ACM, Como, pp 347–348
29. Chen Y, Ma X, Cerezo A, Pu P (2014) Empatheticons: Designing emotion awareness tools for group recommenders. In: Proceedings of the 14 international conference on human computer interaction,

Interacción ’14, Puerto de la Cruz, Tenerife, Spain, September 10-12, 2014, ACM, pp 16:1–16:8
30. Guo J, Zhu Y, Li A, Wang Q, Han W (2016) A social influence approach for group user modeling in group recommendation systems. IEEE Intell Syst 31(5):40–48
31. Christensen IA, Schiaffino SN (2014) Social influence in group recommender systems. Online Inf Rev 38(4):524–542
32. Christensen IA, Schiaffino SN, Armentano M (2016) Social group recommendation in the tourism domain. J Intell Inf Syst 47(2):209–231
33. Ekman P, Friesen WV, O’Sullivan M, Chan A, Diacoyanni-Tarlatzis I, Heider K, Krause R, LeCompte WA, Pitcairn T, Ricci-Bitti PE, Scherer K, Tomita M, Tzavaras A (1987) Universals and cultural differences in the judgments of facial expressions of emotion. J Pers Soc Psychol 53(4):712–717
34. Posner J, Russell JA, Peterson BS (2005) The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. Dev Psychopathol 17(03):715–734
35. Yik M, Russell JA, Steiger JH (2011) A 12-point circumplex structure of core affect. Emotion 11(4):705
36. Hong M, Jung JJ (2016) Mymoviehistory: social recommender system by discovering social affinities among users. Cybern Syst 47(1-2):88–110
37. Piezon SL, Ferree WD (2008) Perceptions of social loafing in online learning groups: a study of public university and us naval war college students. The International Review of Research in Open and Distributed Learning 9(2):1034–1054
38. Stasser G, Davis JH (1981) Group decision making and social influence: a social interaction sequence model. Psychol Rev 88(6):523–551
39. Bond R, Smith PB (1996) Culture and conformity: a meta-analysis of studies using asch’s (1952b, 1956) line judgment task. Psychol Bull 119(1):111