Third International Conference on Information Systems for Crisis Response And Management in Mediterranean Countries. October 26-28, 2016
University Carlos III of Madrid. Madrid, Spain

MADRID 2016
ISCRAM-med

POSTER PAPERS

Paloma Díaz
Narjès Bellamine Ben Saoud
Julie Dugdale
Chihab Hanachi (Eds.)

Paloma Díaz
Narjès Bellamine Ben Saoud
Julie Dugdale
Chihab Hanachi (Eds.)

ISCRAM-Med 2016

Third International Conference on Information Systems for Crisis Response and Management in Mediterranean Countries

Madrid, Oct, 26.28 2016

Organized by:
Interactive Systems Group – DEILab. Universidad Carlos III de Madrid

Sponsored by:
Vicerrectorado de Investigación y Transferencia. Universidad Carlos III de Madrid

Published by the Interactive Systems Group. Universidad Carlos III de Madrid.
Avda de la Universidad 30, 28911
Leganés (Spain), October 2016
ISBN: 978-84-16829-03-3

Table of Contents

From the social semantic web to recommendation

Firas Ben Kharrat[1], Aymen Elkhleifi[2] and Rim Faiz[1]

[1]LARODEC, University of Carthage-IHEC, 2016 Carthage Présidence, Tunisia
[2]LaLic, University Paris-Sorbonne, 75006 Paris, France
{Firas.BenKharrat,Rim.Faiz}@ihec.rnu.tn
Aymen.Elkhlifi@paris.sorbonne.f

Abstract. This paper present a new recommendation algorithm based on contextual analysis and new measurements. Social Network is one of the most popular Web 2.0 applications and related services, like Facebook, have evolved into a practical means for sharing opinions. Consequently, Social Network web sites have since become rich data sources for opinion mining. This paper proposes to introduce external resource from comments posted by users to predict recommendation and relieve the cold start problem. The novelty of the proposed approach is that posts are not simply characterized by an opinion score, as is the case with machine learning-based classifiers, but instead receive an opinion grade for each distinct notion in the post. Our approach has been implemented with Java and Lenskit framework; the study resulted in Movie dataset, we have shown positive results. We compared our algorithm to Slope One algorithm. We have obtained an improvement of 8% in precision and recall as well an improvement of 16% in RMSE and nDCG.

Keywords: Recommendation system, collaborative filtering, user profile, social network, user cold start.

1 Introduction

Recommender systems help users to identify their interests and sets of choices by predicting the usefulness degree of an item or group of items to these users. They are defined as a special type of information filtering that gives information about which items might be interesting to users. Recommender systems generate personalized recommendation for users based on a set of previously rating items. However, social networks and platforms promote the participation of users in many ways, stimulating the expression of opinions about the contents inserted by other users, by 'Like', 'Share' annotations, star-rating systems... This huge amount of data is a precious information source about perceptions, trends, and feelings, and a lot of research is being carried on to identify ways for extracting meaningful information from these data. Indeed, social networks could be an important source of recommendation.

We have improved the Slope One algorithm [4] by external knowledge and through some experiments results. The approach has been implemented in Java and tested on the Netflix MovieLens dataset.

2 Related works

In a social network context, the identification of the opinion of users is receiving an increasing attention. It is a striking indicator of the appraisal of topics, people, situations, re-sources and trends. Hence, the development of opinion Mining [1], of ontologies of emotions and of W3C markup language proposals like Emotion Markup Language [2]. However, there are few application that still use the most avanced results in web technology to deal with opinion. Moreover, most of the approaches use emotional ontologies where emotions are individual isolated units (e.g. WordAffect). Such considerations motivated our focus on the ontology of emotions in [3], an OWL ontology where emotions are structured and organized in levels, trying to integrate the results of the most recent psychological models.

3 Our approach

3.1 General architecture

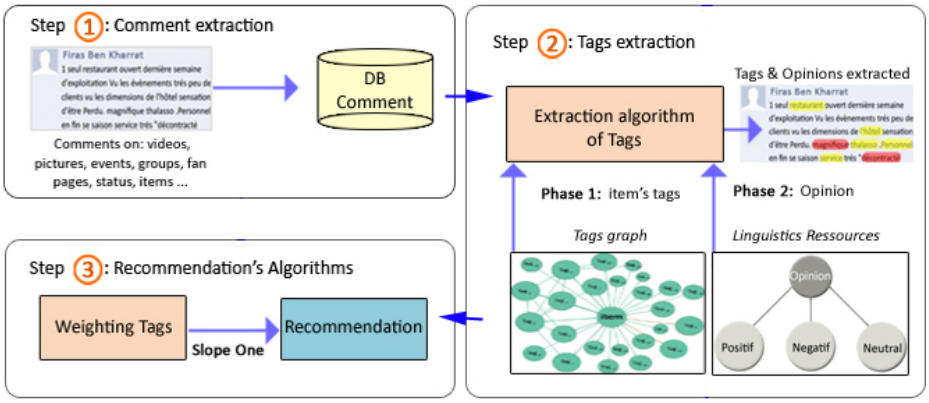

Fig. 1. Architecture of collaborative recommender systems based on contextual analysis.

Figure 1 displays the overall architecture of the approach we propose in this paper. It is based on contextual analysis of user's comment. Our architecture connects the users via collaborative social network in order to use the opinion of their comments. This system combine the social network analysis comments with the opinion of item's tags in one social recommendation algorithm. The proposed system consists of three parts; The first one deals with comment posted by the user. This comment is stored in the data base of comment, described in section 3.2. We use these data to create a user profile. The other one is based on opinion analysis methods that deploy linguistics-based techniques. This consists of exploring the linguistics of different tags, for each user's comment, is described in section 3.3. Next, we calculate a new item and user weight from the extracted opinion. Finally, the recommendation layer, where new input is sent to our algorithm, described in section 3.4.

3.2 Data

The main confronted problem was the issue of gathering data from social network. We created a script in java to extract all available comment based on 'Facebook Graph API', the package is installed in online web site to collect profiles. The dataset used in training and test was a collection of 1k Facebook profiles, which included comments related to the topic of movies.

3.3 Contextual Analysis

Tags graph: This graph is defined as a collection of tags, organized in a hierarchical structure (depending on the domain), in order to describe objects of the items. Our system use tags graph to represent items descriptions. Figure 2 shows an example of tags graph for the item 'Aviator' from the domain of films.

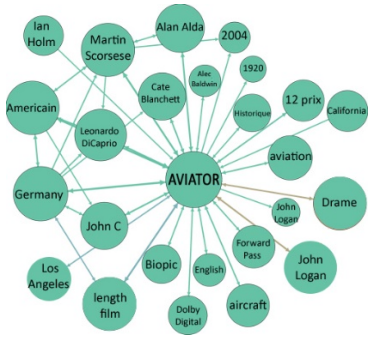

Fig. 2. Example of tags graph of the film 'Aviator'.

Fig. 3. Example of positive Tags.

Linguistic resources: is a lexical resource for opinion mining represented in tags form. We used WordNet the lexical database for the English language. It groups english words into sets of synonyms called synsets. We selected opinion words and we assigned to each synset of WordNet three opinion scores: positive, negative, neutral. An example of positive tags is described in details in figure 3.

Extraction algorithm of tags: used to annotate all tags in comments. We have two types of tags; Tags that described items and tags that described opinion. When, we found the two types of tags in the same sentence, then, we affect the option found to all tags described items in this sentence. Sometimes we found more than two opinions in the same sentence. In this case, we have defined a priority for the selection of opinion: Negative, then Neutral finally positive.

3.4 Recommendation algorithm

As already explained, the basic idea behind the proposed approach is to take advantage of contextual analysis for providing more elaborate opinion scores regarding the tags contained in a comment. The aim is to have a system that accepts as input a comment regarding a specific subject and provides opinion scores for every item's tags of this

subject. The step of recommendation is divided in two phases: (a) creation of opinion's score, and (b) the integration of opinion dimension into recommendation algorithms. These two phases are further described next.
The formalization of weight for the user *i* and item *j* is as follows:

$$w1_{i,j} = \frac{\sum_{k=1}^{n}(\alpha_k \times g_k)}{n} \quad (5)$$

With: $\alpha_k$is the weight of the tag's opinion *k*. *n* is the number of tag. And g_k is the grad of the tag *k;* $g_k \in]0,1[$ depend on tag's importance.
The formalization of weight for system users and item j is:

$$w2_j = \frac{\sum_{k=1}^{n}(\beta_k \times g_k)}{n} \quad (6)$$

With: β_k is the weight of the tag's opinion *k*, returned from all opinions of current users.
The proposed recommendation algorithm is based on the opinions information about users, the tags information about items and the social comments information.
The new algorithm is defined by an improved Slope One algorithm, Eq (7).

$$P^{SS1}(u)_j = \frac{\sum_{i \in R_j}(dev_{i,j}+u_i)}{card(R_j)} + \log\left(1 + \frac{w1_{u,j} \times w2_j}{1+w1_{u,j}+w2_j}\right) \quad (7)$$

With:
$$dev(i,j) = \sum_{u \in S_{i,j}(X)} \frac{(u_i - u_j) \times (1+w1_{u_i,j})}{card(S_{i,j}(X))} \quad (8)$$

4 Conclusion and future work

In this paper we have proposed a new approach for the recommendation systems based on users' profiles. The proposed approach is spread over three stages to recommend the items, in a first stage, by the preprocessing that consists in collecting users' profiles from MovieLens dataset. As for the future perspectives of this study, we suggest to elaborate user profile using a domain ontology. We aim to extend our algorithm to search several connected communities and we will try also to extract information from other resources such as discussion forums, blogs and search history to enrich knowledge about users.

5 References

1. Pang B. and Lee L. (2008) Opinion Mining and Sentiment Analysis. Now Publishers Inc, pages 112-135.
2. Schroder M., Baggia P., Burkhardt F., Oltramari A., Pelachaud C., Pe-ter C., and Zovato E. (2010). Emotion Markup Language (Emo-tionML) 1.0. W3C Working Draft, pages 56-64.
3. Francisco V., Gervas P., and Peinado F. (2007) Ontological reasoning to configure emotional voice synthesis. In Proc. of RR, volume 4524 of LNCS. Springer, pages 18-33.
4. Lemire, D., and Maclachlan. A., (2005). Slope One prediction for online rating-based collaborative filtering. In SIAM Data Mining, Newport Beach, California, pages 21-23.