

EXPLOITING SOCIAL TAGS TO OVERCOME COLD START RECOMMENDATION PROBLEM

Ghabayen, A.S. and Shahrul Azman Noah

Faculty of Information Science and Technology, Universiti Kebangsaan, Bangi, Selangor, Malaysia

Received 2013-12-09; Revised 2014-01-11; Accepted 2014-02-18

ABSTRACT

The practice and method of collaboratively creating and managing tags to annotate and categorize content has resulted in the creation of folksonomy. Folksonomies provide new opportunities and challenges in the field of recommender systems. Despite the considerable amount of researches done in the context of recommender systems, the specific problem of integrating tags into standard recommender system algorithms is less explored than the problem of recommending tags. Collaborative filtering is one of the popular approaches for providing recommendation. However, despite the popularity of collaborative filtering, to some extent, it could not recognize the preferences of users in cold-start scenarios, where insufficient preferences are associated to certain users or items, which leads to degraded recommendation quality. This study presents a collaborative filtering approach based on the expansion of users' tags. In this case, semantics between tags can be unveiled which subsequently resulted in the identification of semantically similar users. Experiment on real-life dataset shows that our approach outperforms the state-of-the-art tag-aware collaborative filtering approaches in terms of recommendation quality particularly in the cold-start situation.

Keywords: Collaborative Filtering, Recommender System, Folksonomy

1. INTRODUCTION

Social collaborative tagging systems, also known as folksonomies, act as a web-based system that allows users to easily annotate their content or any content that someone else has authored with arbitrary terms. These annotations or tags are made in many ways such as editing, rating, organizing and classifying. Nowadays, these web-based systems are becoming more popular among web users. For example, sites such as Flickr, Del.icio.us, last.fm and YouTube provide users the ability to freely annotate content of interest to them. Due to the growing popularity of social collaborative tagging systems, many researchers have recently focused on developing recommender systems that incorporate social collaborative tagging; (Lee and Brusilovsky, 2011; Xu *et al.*, 2011; Pan *et al.*, 2012) to mitigate the limitations of traditional collaborative filtering systems such as "cold-start" users and the data sparsity problem, which leads to degrade recommendation quality.

Tags co-occurrence has been widely used to find similarity between users. However, relying on co-occurrence information alone unable to unveil the semantic similarity of interests among users.

Furthermore, most existing studies on recommender systems with tags are limited to tag suggestions and recommendations in order to help users annotate a related item.

This study focuses on improving the recommendation quality in collaborative filtering recommender system, by tackling the cold start problem. We proposed a method that expand user tags to further unveil the semantic relation between these tags. The semantic relation between these tags are then used to find the similarity between users within a collaborative filtering environment. Recommendation of items are based on the social ranking of items, which is semantically related to the tags that the similar users have created.

The rest of this study is organized as follows. The next section presents an overview of the traditional collaborative

Corresponding Author: Ghabayen, A.S., Faculty of Information Science and Technology, Universiti Kebangsaan, Bangi, Selangor, Malaysia

filtering approach and related work. Section 3 describes the proposed approach. Section 4 and 5 provide description of the experiment conducted and discussion the experiments result respectively. Finally, our conclusion and directions for future work are detailed in section 6.

2. BACKGROUND AND RELATED WORK

Collaborative Filtering hereinafter (CF) is considered to be the best recommendation technique that automates the process of the “word-of-mouth” paradigm in estimating the utilization of unseen items by a user. The main two approaches in CF are the item-based approach and the user-based approach (Su and Khoshgoftaar, 2009). Usually, the recommendation process in both of these CF approaches involves three steps. The first step, is finding a similar pattern for the target user and other users holding similar preferences to form a group of “neighbourhood” users (“similar items” is the term used in the item-based approach). This step considered the most critical task in a CF recommender system, because differences in these like-minded neighbours lead to different recommendations, which influences the quality of the recommendation process (Sarwar *et al.*, 2000). The similarity can be calculated using a similarity calculation method such as cosine similarity (Su and Khoshgoftaar, 2009). Hence, in the traditional rating-based CF, the feature vectors elements are users rating on items. Once these neighbors have been identified, the next step is the process of estimating the prediction value of particular items, which influences the degree to which the target user is likely to prefer the recommended items. The greater the number of similar users found in the recommendation environment for the target user, the more influence the user has on the process of estimating the prediction value for unseen items. The last step in the recommendation process is the presentation of a list of top-N items with the highest predicted values, which are recommended to the target user. The target user then will decide whether they actually like these top-N items (Su and Khoshgoftaar, 2009).

Cold-start is the prevalent problem in recommender systems (Bobadilla *et al.*, 2013). It refers to a situation where no historical ratings on items or users are available. This situation usually leads to degraded recommendation quality.

Tagging is now becoming common in the Web 2.0 era. Social tagging systems allow users to assign content (items) with a freely chosen keyword (tag) and share them with other users (Golder and Huberman, 2006). Tags enable users to simply edit, rate, organize, categorize, classify and search for items. The result of user tagging in social tagging systems is a richly connected network of users, tags and

items which commonly referred to as “folksonomy” (Mathes, 2004). As a consequence, the huge amount of users contributed data available in folksonomies has attracted many researchers to propose novel methods for improving current recommender system.

Different methods for exploiting tags in a folksonomy have been proposed in the last years to improve the recommendation quality of current CF systems. For instance, Nakamoto *et al.* (2008) proposed a tag-based contextual CF recommender system; the main idea being to represent users profiles by using their created tags on items. Sen *et al.* (2009) propose a tag-based recommendation algorithm called “tagommenders” which can be used to predict user preferences for items based on their inferred tag preferences data. Wang *et al.* (2010) proposed a similarity update method that uses the user tags to retrieve nearest neighbors users for the target user. The underlying idea of the proposed similarity is based on using both usual users rating and tagging data to find the closest neighbors for the target user. Nakamoto *et al.* (2008) proposed to use tags and time information when predicting user preferences and consequently exploit such information to build an effective recommender system. Kim *et al.* (2010) exploited tags to enhance the recommendation quality in CF where the tags created by the users were employed to predict user preferences for certain items. Pan *et al.* (2012) use the data mining clustering techniques to cluster the users based on their tagging behavior instead of their similar rating behavior. However, the number of clusters has to be defined because, in the case of a huge tagging space, clustering can be computationally expensive. Furthermore, what score should be assigned to a user who does not fit nicely into the cluster remains a question. This approach still suffers from cold start and data sparsity problems. However, in social tagging systems users tend to annotate a small portion of the shared items that considered interested to them leaving most of other items without tagging. This lack of sufficient tagging (i.e., cold-start problem) leads to degrade recommendation quality.

In this study, we present an approach to overcome the cold start problem by exploiting social tags. We consider semantic tag expansion in order to discover like-minded users. We expect this will be more useful not only in overcoming the cold-start problem, but also in realizing better user perception of relevant items.

3. PROPOSED APPROACH

The inherent idea of exploiting social tags is to build up relations between existing items with new

ones. Theoretically, this mechanism will be able to solve the issue of cold-start mentioned earlier. The main idea behind our proposed approach is to expand users' tags by considering synonyms and word co-occurrence. As such similarity between users can be further enhanced and lead to better recommendation quality. An overview of our proposed approach is as shown in Fig. 1.

3.1. Representation of User Tags

We based our approach on the triple of ⟨user, item, tag⟩ representation which is widely adopted in the collaborative tagging community. A folksonomy can be seen as a set of triples. Each triple represents a user's annotation of an item with a tag. Technically, if there is a list of users $U = \{u_1, u_2, u_3, \dots, u_k\}$, a list of items $I = \{m_1, m_2, m_3, \dots, m_N\}$ and a list of tags $T = \{t_1, t_2, t_3, \dots, t_l\}$, the folksonomy is $F = \langle U, I, T, Y \rangle$, where Y is the user tag assigned to an item (Hu *et al.*, 2012). Note that the traditional CF approaches cannot be used directly on folksonomies because traditional CF generally operates on a binary dimensional relation between users and items. While in folksonomies there is a ternary dimensional relation between users, items and tags, therefore, in order to apply traditional CF in folksonomies the ternary relation has to be decomposed or reduced to a lower dimensional space this leads us to consider the user feature as a vector of tags posted by the user. For example, in Fig. 1, the user $user_1$ tagged $item_1$ with t_1 and t_2 which can be represented as $\langle u_1, m_1, (t_1, t_2) \rangle$.

3.2. Semantic Tag Expansion

WordNet is a large lexical database comprising of nouns, verbs, adjectives and adverbs that are grouped into sets of cognitive synonyms (synsets) (Fellbaum, 1998). Synsets are interlinked by means of conceptual-semantic and lexical relations. The terms which hold the same meanings are referred to as synonyms and those that belong to the same concept are in the same synset. Through these relations, WordNet provides a method to identify semantically similar words. When the system is given a word, semantically similar words can be identified by searching words in the synset of the given word. However, the exploitation of WordNet in social tagging has limitations, according to the nature of the tags because, for example, some keywords may not exist in WordNet. Also, WordNet does not cover domain-specific words. To overcome these limitations we proposed to exploit the internal structure of a folksonomy to identify semantically related tags on the

basis of the co-occurrence distribution between tags on the context of items in a folksonomy.

Budanitsky and Hirst (2006) pointed out that similarity can be considered as a special case of relatedness because both are semantic notations. Therefore, we propose using co-occurrence-based information to identify semantically related tags, where two tags are considered related if they frequently co-occur in specific resources. Wartena *et al.* (2009) proposed an approach to define the co-occurrence relation between tags in social tagging systems as follows.

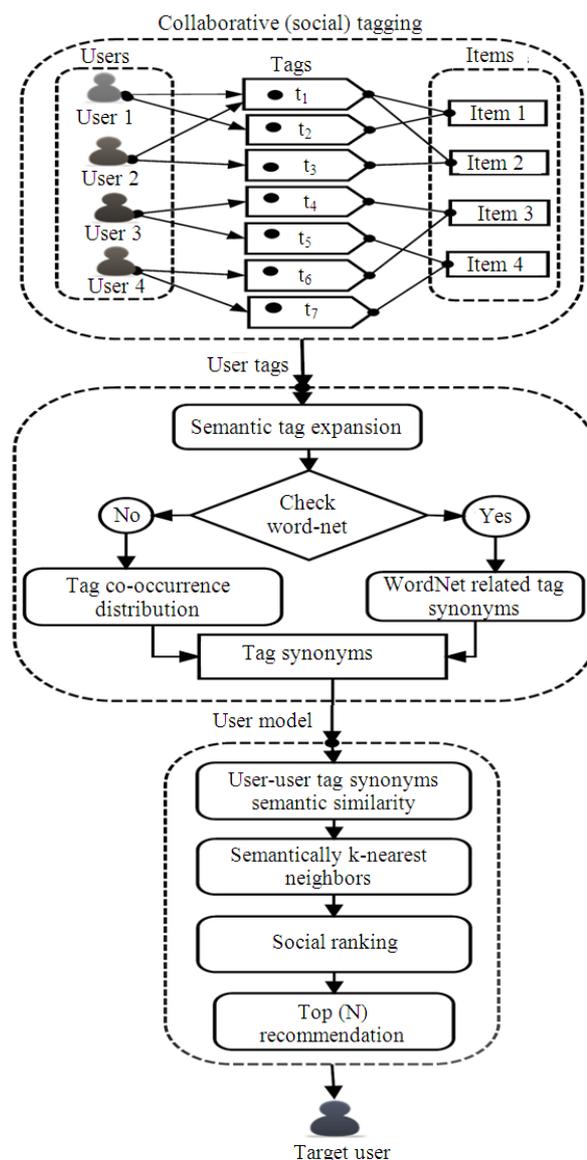


Fig. 1. System overview of our proposed approach

Let $n(m, t)$ be the number of occurrences of tag t on item m , let $n(t) = \sum_m n(m, t)$ be the number of occurrences of tag t and let $N(t) = \sum_t n(m, t)$ be the number of tag occurrences for item m , where m is an instance of I , $m \in I$. In (Wartena *et al.*, 2009), the similarity between two tags t_x and t_y is the weighted average of the tag distribution of items, which denotes the co-occurrence distribution between tags for such an item. The co-occurrence distribution between tags for all items in a social tagging system is calculated by Equation 1.

$$P_{t_y}(t_x) = \sum_{m \in I} q(t_x | m) Q(m | t_y) \tag{1}$$

Where:

$$Q(m | t_y) = \frac{n(m, t_y)}{n(t_y)} \text{ on } I, \quad q(t_x | m) = \frac{n(m, t_x)}{N(m)} \text{ on } m \in I$$

3.3. Generation of a Semantic-Based Neighborhood

One of the critical tasks of a user-based CF recommender system is the generation of a set of like-minded or nearest neighbor users who have similar tastes to the target user. Consider two users, u and v , where $u, v \in U$. First, we obtain items $m_{u,v} \in I$ which are sharable in terms of tagging behavior between the related users. For each item $m \in m_{u,v}$ we present each user with the tags that have been posted by user u and v . The tag posted by user u for item m can be identified by $T_u^m := \{t_u \in T \mid t_u \times (m)\}$. All tags posted by user u on item m are identified by $T_u^* := \{t_u \in T \mid \exists m \in I : T_u^m \in T\}$. Each tag $t_u \in T_u^*$ is expanded to its corresponding synonyms. Consequently, the set of user u tags on items are expanded to a set of tag synonyms Ts_u^m such that $Ts_u^m \cap T_u^m = \emptyset$, which means the tag is not considered as a synonym for itself.

The similarity between two users based on tag synonyms is denoted as the UTSSim. User tag synonyms similarity between two users for a specific item is calculated using dice similarity as follows Equation 2:

$$UTSSim(u, v) = \sum_{m \in m_{u,v}} 2 \times \frac{|Ts_u^m \cap Ts_v^m|}{|Ts_u^m \cup Ts_v^m|} \tag{2}$$

Where:

- Ts_u^m = The synonyms of tags posted by user u on item m
- Ts_v^m = The synonyms of the tags posted by user v on item m
- $m \in m_{u,v}$ = Which is the co-tagged item between user u and v

The higher the UTSSim value between the two users the more similar they are. Finally, for a given user u we determine the top k users with the highest UTSSim for user u . We denote this set as a set of Semantically Nearest Users (SNU) and define it as follows Equation 3:

$$SNU_k(u) = \operatorname{argmax}_{v \in U - \{u\}} (UTSSim(u, v)) \tag{3}$$

3.4. Item Recommendation

When the set of k semantically similar users has been identified, the last step consists of the actual prediction for each item and the generation of a top N list of recommended items. In our approach, the basic idea of estimating relevant unseen items for the target user starts from the assumption that the user prefers items that have been tagged by like-minded users. We describe this assumption as a Semantic Social Rank (SSR) from the set of SNU and defined as:

$$SSR(u, r) = \sum_{v \in SNU_k(u)} UTSSim(u, v) \times \text{social rank}(v) \tag{4}$$

Where:

- r = The items that have not seen by user u ;
- UTSSim = User tag synonym similarity value between two users' u and v
- Social rank = The rank value from semantically like-minded users

The social rank is equal to 1 if the item has been annotated by semantically neighborhood users; otherwise the value is 0. Finally, a set of top- N ranked items that obtained higher SSR scores are recommended to user u .

4. EXPERIMENTAL EVALUATION

4.1. Dataset

We conducted our experiments using HetRec2011-movielens-2k dataset (<http://www.grouplens.org>). The dataset is an extension of MovieLens10M dataset, published by Group lens research group. It links the movies of MovieLens dataset with their corresponding web pages at Internet Movie Database (IMDb) and

Rotten Tomatoes movie review systems. It contains rating and tagging information of users and items.

One of the major issues when dealing with tagging data is the quality of the tags because tags are words or combinations of words that are freely assigned by users. In order to ensure the quality of our experiment and its findings, it is necessary to remove meaningless data by filtering the dataset. Since our proposed approach depends on co-occurrence distribution, we removed meaningless tags, i.e., tags that had not been assigned to at least two items and items that had not been annotated by at least two tags, because this would lead to a zero co-occurrence distribution score with other tags. The final pruned dataset used in our study consists of the following: 2,077 users, 4,480 items, 46,528 tagging records and 8,866 tags.

The main problem when trying to map tags in the datasets to WordNet lexicon is that not all the tags in the dataset are recognized by the lexicon. For MovieLens data set, 37% of the tags in the dataset were not in WordNet lexicon.

4.2. Evaluation Metrics

To evaluate the performance of our proposed approach, we adopted two famous metrics from the information retrieval field, namely, precision and recall, which judge how relevant the recommended items are to the target users (Herlocker *et al.*, 2004). Precision measures the ratio of the number of relevant items in a list of recommendations out of all items retrieved. Recall measures the ratio of the number of relevant items retrieved to the total number of relevant items in the test set. In our experiment we withheld the items that had been previously tagged by the target user and then calculated the precision and recall for the target user u , as follows; where $\text{Test}(u)$ is the relevant items for user u in test set and $\text{Top } M(u)$ is the top M recommended items for user u Equation 5 and 6.

$$\text{Precision}(u) = \frac{|\text{Test}(u) \cap \text{Top}N(u)|}{|\text{Top}N(u)|} \quad (5)$$

$$\text{Recall}(u) = \frac{|\text{Test}(u) \cap \text{Top}N(u)|}{|\text{Test}(u)|} \quad (6)$$

Another effectiveness measure used is the F1 measure which is defined as the harmonic mean of recall and precision as follows Equation 7:

$$\text{F1} = \frac{2 \times \text{recall} \times \text{precision}}{\text{recall} + \text{precision}} \quad (7)$$

The measure has an advantage of summarizing effectiveness in a single number instead of individual measure of recall and precision.

4.3. Comparisons with Baseline Approaches

We refer our proposed approach as Semantic Tag Expansion Collaborative Filtering (STECF) and compare with two baseline approaches in a cold start situation. The two compared baseline approaches are:

- **User-based Collaborative Filtering (UBCF).** UBCF is a conventional user-based collaborative filtering approach based on a *user-item* relationship, or a *user-tag* relationship. In our experiment the similarity between two users is calculated in three different ways: (i) using the classical cosine similarity of users tags set (denoted as *UBCF-Tag-cosine-sim*) (Sarwar *et al.*, 2000); (ii) using the overlapping principle of users items sets (i.e., the User-item mapping) (denoted as *UBCF-Item-overlap-sim*); and (iii) using the overlap of users tags sets (i.e., the User-tag mapping) (denoted as *UBCF-Tag-overlap-sim*) (Jäschke *et al.*, 2008)
- **Item-based Collaborative Filtering (IBCF).** IBCF is the conventional item-based collaborative filtering approach based on the *item-user* relationship, or *item-tag* relationship. The same similarity measures in UBCF is implemented in UBCF and respectively referred them as: *IBCF-Tag-cosine-sim*, *IBCF-Item-overlap-sim* and *IBCF-Tag-overlap-sim*.

4.4. Evaluation Protocol and Settings

To evaluate the performance of our proposed approach, we used the so-called back testing approach, which is common in recommender system evaluations (Baluja *et al.*, 2008). The dataset was divided into five portions and each fold was used once as a test set. In this way, 20% of the dataset was used as the test set and 80% was used as the training set. For each user in the test set, 20% of items in the user profile were used as the test set while the remaining 80% of items of the user profile formed the training set to generate the recommendations for the 20% of items in the test set. Finally, the values for our evaluation metrics of the validation items were then averaged over the five folds. We adopted this approach to ensure that the result would not be biased toward a particular test/training portion for each user with profiles that might be easier to predict. This means that our

proposed approach can make equal recommendations for all users, not just for the most active ones.

We test the performance of the proposed approach STECF in a cold-start situation, where traditional CF recommender systems is generally unable to provide a high recommendation quality, due to insufficient users or items preferences in case of user-based and item-based CF. In order to simulate the cold-start situation, we considered users who had 5 tagging records or less on the training set as cold start users. Similarly for item-based CF, items which annotated with less than 5 different tags or less in the training set considered as unpopular items.

5. EXPERIMENTAL RESULTS

In this section, we present the results of our experiment with respect to the average improvement of recall, precision and F1-measure achieved by the proposed approach compared with the baseline approaches. The experiment to find the top N recommended items was performed with various numbers of recommended items; we considered N from 5 to 20 with an increment of 5 for each user in the test set. Science in real environment users inclined to choose items with higher ranks. Furthermore, because the size of the k nearest neighbours significantly influence the recommendation quality in CF recommenders systems (Herlocker *et al.*, 2004). Based on a subsequent experiment most of the presented baseline approaches obtains a best performance at k value equal to 10 with respect to F-measure as a judgment metrics. Hence, a value of k is equal 10 has been considered in the experiments.

Figure 2 to 4 respectively presents the average recall, precision and F1 measures achieved by our proposed approach (STECF) in a cold start situation as compared to the baseline approaches.

6. DISCUSSION

The result in Fig. 2 to 4 show that the proposed STECF approach achieves the best performance with an average improvement of 51.5, 6.8 and 11.7% as compared to the baseline approaches in terms of recall, precision and F1-measure respectively.

The previous results illustrate that the baseline approaches are unable to achieve good recommendation, in a cold-start situation as too little tagging preferences are available about the users or items. Hence, the little available preference (tags), result in a mismatch problem

between users/items profiles vectors which makes the calculation of the similarity between users/items inefficient.

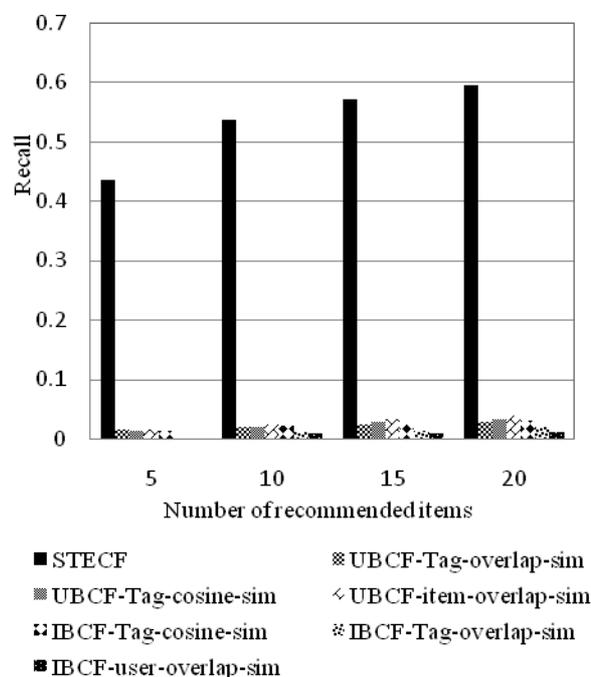


Fig. 2. Top-N Recall comparison with increment of recommended items

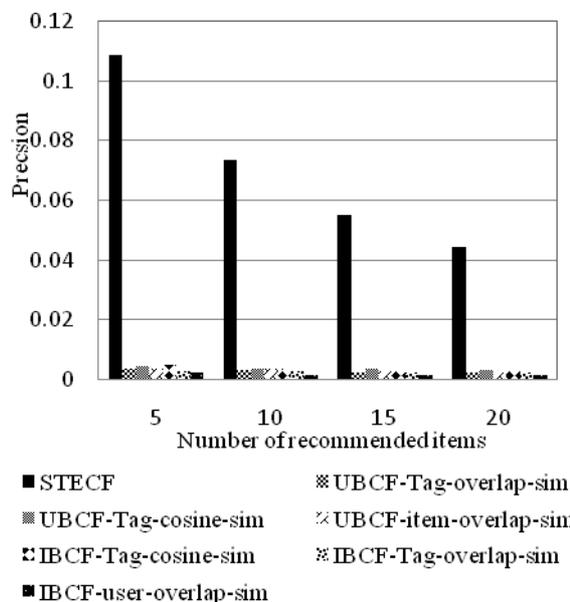


Fig. 3. Top-N precision comparison with increment of recommended items

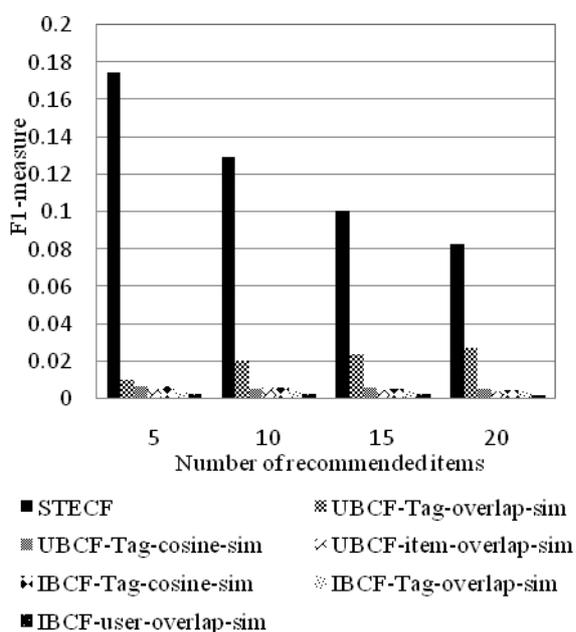


Fig. 4. Top-N F1-measure comparison with increment of recommended items

As a result inaccurate recommendation is produced. On the other hand the proposed approach was able to exploit the little information available and overcome the mismatch problem by enriching the users profile vectors with more semantically related preferences, resulting in relevant items which do not tagged with the same tags but have been tagged with semantically relevant tags will be recommended.

From Fig. 2 and 3 we can see that the recommendation performance improved with the increment of the Top-N recommended items in terms of recall, however the precision is gradually decreases. One possible explanation for this result is that with the increment of N recommended item, will result in more false positives recommendation, thereby resulting in low precision, but true positives are likely to be returned resulting in higher recall values. This pattern of findings is very common in information retrieval research.

The overall results illustrate that STECF shows a significant gain in recall, precision and F1-measure compared with other baseline approaches, which conform that the STECF approach can deal better with the cold start situation and recommend better semantically related items with higher rank compared with the baseline approaches.

7. CONCLUSION

To date, most CF recommender systems suffer from the cold-start and data sparsity problems which are the case of insufficient availability of users or item tagging data. In folksonomies, users tend to annotate a small portion of the shared items that considered interested to them and leave the remaining items untagged. This lack of sufficient tagging can significantly impact recommendation quality. Therefore, this study presented an approach for deriving the semantic similarity between users by exploiting the semantic expansion of user tags. Such semantic expansion similarity helps to enhance the recommendation quality of CF recommender system and alleviate the cold-start problem. The idea stems from the believe that, ‘similar’ tags provided by different users might indicates their relatedness and potential input for recommender systems. However, most tagging activities are with little to no control in terms of terms and vocabulary used. Therefore, different tags can semantically mean the same. Therefore, in order to overcome such a situation, we used WordNet to assist in measuring the semantic relatedness between tags. In the case of non-existence words in the WordNet database, the co-occurrence distribution measure is used.

Evaluation on the MovieLens dataset has shown interesting and promising results. The proposed approach outperforms the conventional CF approaches in terms of precision, recall and F1 measures in a cold-start situation, which indicate that exploiting semantic tags information, can improve the quality of item recommendation and alleviate the cold start problem in conventional CF recommender systems. However, no doubt that the complexity and extra processing required to perform the semantic processing might be the setback of this approach. Potential future research works include integrating user’s tags with other user’s preferences such as user rating, blogs and reviews to improve the recommendation quality. Furthermore, with the emergence of Semantic Web and particularly Linked Open Data (LOD) (Heath and Bizer, 2011), expansion of tags to such open data is another potential work in this area.

8. REFERENCES

Baluja, S., R. Seth, D. Sivakumar, Y. Jing and J. Yagnik *et al.*, 2008. Video suggestion and discovery for youtube: Taking random walks through the view graph. Proceedings of the 17th International Conference World Wide Web, Apr. 21-25, New York, NY, USA, pp: 895-904. DOI: 10.1145/1367497.1367618

- Budanitsky, A. and G. Hirst, 2006. Evaluating wordnet-based measures of lexical semantic relatedness. *Comput. Linguist.*, 32: 13-47. DOI: 10.1162/coli.2006.32.1.13
- Bobadilla, J., F. Ortega, A. Hernando and A. Gutiérrez, 2013. Recommender systems survey. *Knowl. Based Syst.*, 46: 109-132. DOI: 10.1016/j.knsys.2013.03.012
- Fellbaum, C., 1998. *WordNet: An Electronic Lexical Database*. 1st Edn., MIT Press, Cambridge, ISBN-10: 026206197X, pp: 423.
- Golder, S.A. and B.A. Huberman, 2006. Usage patterns of collaborative tagging systems. *J. Inf. Sci.*, 32: 198-208. DOI: 10.1177/0165551506062337
- Heath, T. and C. Bizer, 2011. *Linked Data: Evolving the Web Into a Global Data Space*. 1st Edn., Morgan and Claypool Publishers, ISBN-10: 1608454304, pp: 122.
- Herlocker, J.L., J.A. Konstan, L.G. Terveen and J.T. Riedl, 2004. Evaluating collaborative filtering recommender systems. *ACM Trans. Inform. Syst.*, 22: 5-53. DOI: 10.1145/963770.963772
- Hu, J., B. Wang, Y. Liu and D.Y. Li, 2012. Personalized tag recommendation using social influence. *J. Comput. Sci. Technol.*, 27: 527-540. DOI: 10.1007/s11390-012-1241-0
- Jäschke, R., L. Marinho, A. Hotho, L. Schmidt-Thieme and G. Stumme, 2008. Tag recommendations in social bookmarking systems. *AI Commun.*, 21: 231-247.
- Kim, H.N., A.T. Ji, I. Ha and G.S. Jo, 2010. Collaborative filtering based on collaborative tagging for enhancing the quality of recommendation. *Electr. Commerce Res. Applic.*, 9: 73-83. DOI: 10.1016/j.eierap.2009.08.004
- Lee, D.H. and P. Brusilovsky, 2011. Improving recommendations using watching networks in a social tagging system. *Proceedings of the Conference Seattle*, Feb. 08-11, New York, NY, USA., pp: 33-39. DOI: 10.1145/1940761.1940766
- Mathes, A., 2004. Folksonomies-cooperative classification and communication through shared metadata. *Comput. Med. Commun.*
- Nakamoto, R.Y., S. Nakajima, J. Miyazaki, S. Uemura and H. Kato *et al.*, 2008. Reasonable tag-based collaborative filtering for social tagging systems. *Proceedings of the 2nd ACM Workshop on Information Credibility on the Web, (ICW' 08)*, ACM USA., Napa Valley, California, pp: 11-18 DOI: 10.1145/1458527.1458533
- Pan, R., G. Xu and P. Dolog, 2012. Improving Recommendations. In: *Tag-Based Systems with Spectral Clustering of Tag Neighbors*, Park, J.J., H.C. Chao, M.S. Obaidat and J. Kim, (Eds.,) Springer Netherlands, ISBN-10: 978-94-007-2792-2, pp: 355-364.
- Sarwar, B., G. Karypis, J. Konstan and J. Riedl, 2000. Analysis of recommendation algorithms for e-commerce. *Proceedings of the 2nd ACM Conference Electronic Commerce*, Oct. 17-20, New York, USA, pp: 158-167. DOI: 10.1145/352871.352887
- Su, X. and T.M. Khoshgoftaar, 2009. A survey of collaborative filtering techniques. *Adv. Artif. Intell.*, 2009: 19-19. DOI: 10.1155/2009/421425
- Sen, S., J. Vig and J. Riedl, 2009. Tagommenders: connecting users to items through tags. *Proceedings of the 18th international conference on World Wide Web, (WWW' 09)*, ACM, Madrid, Spain., pp: 671-680. DOI: 10.1145/1526709.1526800
- Wartena, C., R. Brussee and M. Wibbels, 2009. Using tag co-occurrence for recommendation. *Proceedings of the 9th International Conference Intelligent Systems Design and Applications*, Nov. 30-Dec. 2, IEEE Xplore Press, Pisa, pp: 273-278. DOI: 10.1109/ISDA.2009.130
- Wang, Z., Y. Wang and H. Wu, 2010. Tags meet ratings: Improving collaborative filtering with tag-based neighborhood method. *Proceedings of the Workshop on Social Recommender Systems, (SRS' 10)*, Hong Kong, China, pp: 15-23.
- Xu, G., Y. Gu, Y. Zhang, Z. Yang and M. Kitsuregawa, 2011. TOAST: A topic-oriented tag-based recommender system. *Proceedings of the 12th International Conference Web InformationSystem Engineering*, Oct. 13-14, Australia, pp: 158-171. DOI: 10.1007/978-3-642-24434-6_12