DEWS2007 M5-6

Tag-Based Contextual Collaborative Filtering

Reyn NAKAMOTO†, Shinsuke NAKAJIMA†, Jun MIYAZAKI†, and Shunsuke UEMURA†

† Nara Institute of Science and Technology, 8916-5 Takayama-cho, Ikoma-shi, Nara-ken, 630-0101, Japan
E-mail: †{reyn-n,shin,miyazaki,uemura}@is.naist.jp

Abstract In this paper, we introduce a new Collaborative Filtering (CF) model which takes into consideration users' context based upon tagging information such as available from recently popular social tagging systems. In numerous implementations, traditional CF systems have been proven to work well under certain circumstances. However, CF systems still suffer a weakness: They do not take context into consideration. Yet recently, social tagging systems have become popular–these systems provide a well suited combination of context clues through tags as well as important social connectivity among users. Thus, we combine the features of these two systems to create a Tag-Based Contexual Collaborative Filtering model.

Key words collaborative filtering, tagging, recommendation systems, information retrieval

1. Introduction

As the Internet continues to mature and becomes more accessible to the common user, the amount of information available increases exponentially. Accordingly, finding useful and relevant information is becoming progressively difficult. Moreover, a lot of the information available–blogs, various types of reviews, and so forth–are highly subjective and thus, hard to evaluate purely through machine algorithms. Being subjective in nature, one person may absolutely love something while the next may loathe the same–no single authority exists. It is in these cases where people–more so than the current ability of machine algorithms–are greatly effective in evaluating and filtering this information. For this reason, the idea of Collaborative Filtering (CF) was started, extensively researched, and eventually deployed to relatively good amounts of success. Using the people and the community, recommendations of subjective information can be made through the matching of similar users.

However, CF systems suffer a weakness: They do not take into consideration the context in which a resource was liked. While two users may both like the same resource, they may like it for different reasons. For example, one user may like a resource because it is funny and interesting, while another user may like it because it was informative and written clearly. Moreover, context includes the subject as well: one user may like a resource because it is about his favorite baseball team, while another user may like something because it is about his favorite book series. Since traditional CF is based upon numerical ratings of such resources, determining why a user likes something–in other words, the context of the preference–is difficult, and thus, this is a weakness of the traditional CF model. Thus, the limitation is further exaggerated when crossing resource domains or using larger domains such as the internet: for example, while two users may be interested in sports–which would be the context in this case–the same preference may not hold for a vastly different domain like politics.

Yet the recent advent of social tagging systems and its characteristics justify another look at CF systems. Social tagging systems rely upon the similar concepts of using the community to sort and organize information: these systems allow users to attach tags–natural language keywords of their choosing–to describe resources. Subsequently, these tags are used for later retrieval and resource discovery not only by the original user, but by the entire community of users as well. Interesting enough, these tags are used for several different purposes, including denoting the subject itself, the category, or the refining characteristics of the resource [7]–for example, a picture of a dog would most likely be tagged something like 'dog', 'animal', or maybe 'cute'. Thus, tags seem to provide the missing link in CF: it provides the subject, category, or some refining trait of a resource–in other words, the context in which the user liked and subsequently bookmarked a resource. With these characteristics, social tagging systems seem to be a well-suited combination of social integration and context to fit together with CF systems.

That being said, combining these two seem like the next likely step in their evolution: Tag-Based Contextual CF. In this paper, we introduce two models for combining CF and tagging systems at different stages in the recommendation process. After this, we sum it together in terms of actually

recommending resources to users of such a system.

2. Related Work

2.1 Collaborative Filtering Systems

Collaborative Filtering (CF) is the process whereby the community of users is used to sort out relevant or important information from the non-relevant or non-important information. The process is based upon the idea that if users prefer the same item or items, then their preference will be similar for other items liked by the similar users. In other words, a user should like the same items that their similar users like. From a wider perspective, once users have recorded their preferences within the system, subsequent users can benefit from the previous users' knowledge, hence the collaborative aspect of the system.

CF has been proven to work well under certain domains–mainly entertainment domains–such as usenet recommendations [10], movie recommendations [6], product recommendations [1], and so forth. However, as noted before, traditional CF systems do not take context into consideration.

Many CF systems rely upon a matrix of numerical ratings of resources by users [10]. Once enough ratings are in place, a similarity score is calculated between the user and other users. These similarity scores are multiplied by the ratings other users recorded and then averaged. Those resources with an average score above a certain threshold are recommended. This is to be further explained in section 3.1 and 3.2.

2.2 Social Tagging Systems

Tagging has been around for sometime, albeit known by other terms such as metadata, categorization, labels, and so forth. Tagging is the process of attaching natural language words as metadata to describe some resource like a movie, photo, book, etc. Tagging vocabulary is usually uncontrolled, whereby the user themselves can decide what word or combination of words are appropriate.

The current main use of tagging is for the purpose of retrieval [9], whereby users can search for a tag and the resources with that tag attached will be returned to the user. The user who added the tag can use tags for later retrieval. For other users, tags serve as a way to discover new resources by searching for whatever tag they are interested in.

In recent years, the advent of Social Tagging Systems have brought tagging back into the limelight. Currently, there are several online social tagging systems that are popular and are the subject of continuing research [9]: they range from website bookmarking such as del.icio.us [4], photo sharing [5], research paper searching [2], to even people rating [3]! All of these sites use tagging for many purposes, but in addition to that, they focus on the social networking aspects of tagging to enhance the experience for end users. However, in their present form, tags are generally used for tag searching; user profile matching and subsequent recommendations are yet to be implemented. As mentioned before, tags provide the clues as to why a user liked something. Because of this, as well as the similar use of social networking, social tagging systems provide an ideal choice for combination with CF systems.

3. Tag-Based Contextual Collaborative Filtering

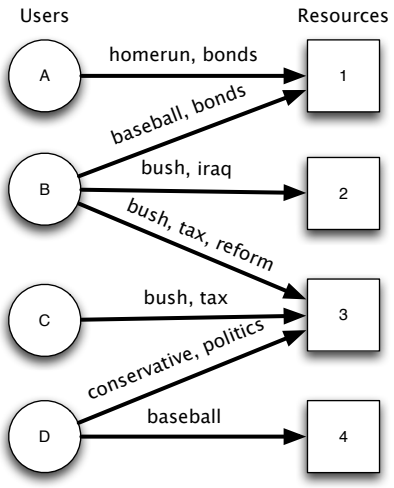

Figure 1: Contextual CF Model

Tag-Based Contextual Collaborative Filtering (TCCF) is the combination of traditional CF systems and social tagging systems to allow for accurate resource recommendation that takes into consideration the context of the preference. An example system is shown in figure 1. In this case, users are attaching tags to resource that they like and wish to access later. A resource may be anything including music, videos, etc.–but in terms of using our new TCCF model, larger domains such as website bookmarking would show its full potential. An example system that would use our TCCF model would follow the following process:

(1) The user evaluates a resource such as a website.

(2) If the user likes a resource, they bookmark it with whatever tags are appropriate. These tags that explain what the resource meant to them. These bookmarks are for later retrieval.

(3) The system calculates user similarity to other users based upon their common bookmarks and tags. This is to be explained in section 3.1.

(4) The system then calculates the predicted scores for yet unrated resources based upon the user similarities calculated in the step before and also other users' bookmarks. This is to be explained in section 3.2.

(5) The system recommends new resources to the user based upon both step 3 and 4. This is further explained in section 3.3.

Unlike traditional CF models which use numeric ratings, this TCCF model uses tags as the indicator that a user likes something: if a user bookmarks something, the system sees it as the user liking it. Thus, in this sense, it is a simpler CF model in that it uses a boolean rating as opposed to a numeric scale like in most CF systems. The intended scale would be a from zero to one. A bookmark would correspond to a high rating or one, and the lack of one would be analogous to a non-rating or zero.

However, tags also provide the key distinguishing factor from traditional CF systems–the tags attached to the resource can be seen as the context in which the user likes the resource. Usually the user will use tags to describe the resource as the user themselves see it, and in most cases, it would be the 'context'. From this assumption, we build upon incorporating using this context to modify the CF model.

Based upon this basic process, we now introduce two submodels in which the traditional CF model is modified in the user similarity stage and the resource recommendation stage.

3.1 Contextual CF User Similarity Model

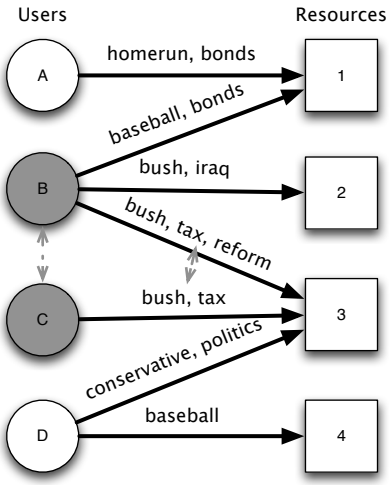

Figure 2: Contextual CF User Similarity Model

This model focuses upon considering the context when calculating the user similarity between users. In traditional CF, the similarity rating is based upon the numerical ratings such as shown in the following table:

	1	2	3	4
A	1	-	-	-
B	1	1	1	-
C	-	-	1	-
D	-	-	1	1

Table 1: Traditional CF Ratings

where A,B,C,D are users and 1-4 are resources. For simplicity's sake when comparing this to our new model, a one corresponds to high rating and negative or low ratings have been omitted. Additionally, a non-rating is denoted by a '-'. User similarity is the cosine similarity of the vectors of the scores users rated for each resource:

	A	B	C	D
A	-	0.58	0	0
B	0.58	-	0.58	0.41
C	0	0.58	-	0.71
D	0	0.41	0.71	-

Table 2: Traditional CF User Similarity Scores

If users have enough similar ratings over the same resources, the system will give them a high similarity rating. If the users have dissimilar ratings, or did not rate the same resources, the system does not give them high similarity rating. However, using only numbers does not tell as to why the user likes something–or in other words, in what context the resource was liked. Tags, however, provide more insight as to why the user may have liked it.

Consider two users tagging a blog on politics. One user whose views are in line with the blog's may tag the resource 'informative' or 'insightful'; however, another user whose views do not agree may tag the blog as 'funny' or 'entertainment'. As can be seen here, the users may follow the same blog, but follow it for different reasons.

Thus, the first model modifies the user similarity calculation. User similarity between a user A and a user B is calculated using the following equation:

$$sim_{ccf}(A,B) = \frac{1}{2n}\sum_{k=1}^{n}\{sim(T_{A\to k}, T_{B\to k}) + 1\} \qquad (1)$$

where n is the number of commonly tagged resources between the users A and B. Commonly tagged means both users bookmarked the same resource with any, possibly differing tag vectors. Also:

$$sim(T_{A\to k}, T_{B\to k}) = \frac{T_{A\to k}\cdot T_{A\to k}}{|T_{A\to k}||T_{B\to k}|} \qquad (2)$$

where $T_{A\to k}$ is the tag vector that user A used for commonly tagged resource k and $T_{B\to k}$ is the tag vector that user B used for commonly tagged resource k. Similarity of the two tag vectors is computed through cosine similarity.

Essentially, in this user similarity model, the cosine similarities of the tag vectors of all commonly tagged resources between user A and user B are averaged. Moreover, one is added to the cosine similarity value to give value to a commonly tagged resource. Regardless of mismatching tag sets, a commonly tagged resource is worth more than none at all.

For example, for the system shown in figure 2, we can calculate user B and user C's similarity. Since user B and user C only have one commonly tagged link–resource 3–$n = 1$ and their similarity score is based entirely on this. User B and C tag vectors on resource 3 would be:

	bush	tax	reform
$T_{B\to3}$	1	1	1
$T_{C\to3}$	1	1	0

The cosine similarity between the two tag vectors is 0.5 and thus, $sim_{ccf}(B, C) = 0.75$. The resultant user similarities for the rest of system in figure 2 would be:

	A	B	C	D
A	-	0.75	0	0
B	0.75	-	0.91	0.5
C	0	0.91	-	0.5
D	0	0.5	0.5	-

Table 3: Contextual CF User Similarity Scores

As can be seen here, whereas before users C and D had a high similarity (data sparsity is also a factor), now the similarity score between the two is lower due to dissimilarity between the tags used on the commonly tagged resource. Oppositely, between users A and B as well as between users B and C, their similarity scores are now high because both pairs bookmarked the same resource with similar tag vectors.

With this new user similarity model, context is considered when calculating user similarity. Users that have bookmarked the same resource are still considered similar; however, the similarity is higher if the tags used to describe the resource are similar. Thus, matching context is pushed higher than when the context does not match.

However, weaknesses of this model still exist. In this system, there is only the option of bookmarking or not bookmarking. Thus, most users would only bookmark a resource if and only if they like the resource. However, the lack of a bookmark does not necessarily mean dislike, but also that they may just not have evaluated it. Given this, the model is only dependent on commonly tagged resources–this may be an issue when commonly tagged links are few and far between. This issue is common to other CF-based systems as well [8].

Additionally, there are the natural language issues that exist with tagging sites [9]. Issues like synonymy and polysemy may have to be accounted for–and in that case a method for linking semantically related words must be considered when implementing this model. Moreover, there is the issue of whether different users would use tags for the same purpose–categorizing, naming, etc.–and whether those purposes would match between users.

3.2 Contextual CF Score Prediction Model

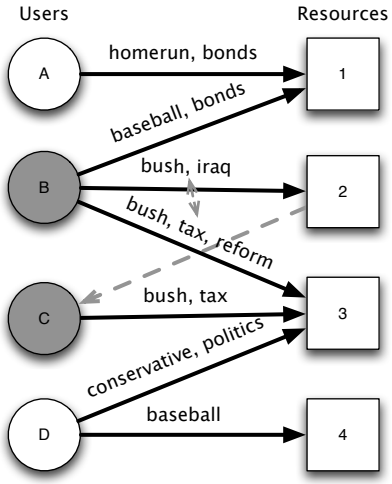

Figure 3: Contextual CF Score Prediction Model

This model modifies the score prediction calculation. In traditional CF, once user similarity has been calculated, the predicted score for some user A for an unevaluated resource x is calculated by the following:

$$score(A, x)_{pred} = \frac{\sum_{k=1}^{n}\{sim(A, S_k) * score(S_k, x)\}}{\sum_{k=1}^{n} sim(A, S_k)} \quad (3)$$

where n is the number of other users. Essentially, all other users' scores, weighted by their similarity, are averaged to predict A's score for x.

For example, for the ratings data provided in table 1 from section 3.1, calculating user B's predicted score for resource 4 would be 0.26. The rest of the prediction would be as shown in table 4.

	1	2	3	4
A	-	1.0	1.0	0.0
B	-	-	-	0.26
C	0.45	0.45	-	0.55
D	0.37	0.37	-	-

Table 4: Traditional CF Score Predictions

Traditional CF works well when recommending resources in the same domain–say only comedy movies or only soccer videos. However, in the case of website bookmarking and the internet, many domains and many contexts are covered. While a user may have the same preference for one context, the preference may not carry over to another: For example, although two users may both like reading Harry Potter books, this does not necessarily mean that if one likes soccer, the other user will too. Context filtering is needed to produce more accurate recommendations.

Thus, this model focuses upon considering context when

$$score_{pred}(A,x) = \frac{\frac{1}{2}\sum_{k=1}^{n}\{sim_{ccf}(A,S_k) * (\max(sim(T_{S_k\to 1}, T_{S_k\to x}), ..., sim(T_{S_k\to m}, T_{S_k\to x})) + 1)\}}{\sum_{k=1}^{n} sim_{ccf}(A,S_k)} \quad (4)$$

predicting a score for a user. This takes place after user similarity has been calculated. Once this is done, resource score prediction occurs. Instead of predicting scores for resources based upon just the average score as shown in equation 3, this model considers if the contexts of the resources are similar. Thus, the Contextual CF Score Prediction of resource x for a user A is as shown in equation 4.

In this equation, S_k is a user in the set of all users with a similarity score with user A above a certain similarity threshold and n is the number of users in this set. Also, in

$$\max(sim(T_{S_k\to 1}, T_{S_k\to x}), ..., sim(T_{S_k\to m}, T_{S_k\to x}))$$

m is the number of commonly tagged resources that user S_k has with user A. This part returns the tag vector of a commonly tagged resource which has the highest similarity to the tag vector of the target resource, $T_{S_k\to x}$. This is done in order to only use the most appropriate (highest tag vector similarity) context for the score prediction. Obviously, if user S_k has no commonly tagged resources with A, the max score would be zero. A value of one is added to the max to give value to the existence of tagged resource, regardless of tag vector similarity. Lastly, tag vector similarity is calculated as shown in equation 2 from section 3. 1.

Overall, this new model shown in equation 4 functions similarly to the traditional CF score prediction model shown in equation 3, except that the tag vector similarities are averaged instead.

In the case of figure 3, we want to predict user C's scores for resource 2, a resource in which he has not yet tagged. Using the CCF user similarity scores shown in table 3 in section 3. 1, user B has a high similarity to user C.

For user B, only resource 3 is commonly tagged with user C. Thus, the tag vectors that B attached to resource 2 and 3 are as follows:

	bush	tax	reform	iraq
$T_{B\to 3}$	1	1	1	0
$T_{B\to 2}$	1	0	0	1

Therefore $sim(T_{B\to 3}, T_{B\to 2}) = 0.41$. If user B had more commonly tagged resources with user C, similarity with that tag vector $T_{B\to k}$ and $T_{B\to 2}$ would also be calculated and the highest similarity used in score calculation. Given the user similarity of B to C is 0.91, the final predicted score is: $score(C,2) = 0.71$. Conversely, say we wanted to predict user C's score for resource 1. While user B's similarity to user C is high, tag set $T_{B\to 1}$ is totally different from B's commonly tagged resource's tag set, $T_{B\to 3}$, therefore meaning that the context of the preference is different, and consequently, its predicted score is lower: $score(C,1) = 0.5$. Resource 2 would be recommended over resource 1. Previously, resource 1 and resource 2 had the same score as shown in table 4. Using this model, if the tags of the target resource match the tags used on a commonly tagged resource, then its score is higher. If it is not, the score will be lower.

Through this, context will be considered when making score predictions. Such systems will be more successful on bigger or cross domain recommendations. This would apply well to website bookmarking sites such as del.icio.us and so forth. Another benefit is that unlike CCF user similarity calculation, this model does not suffer as much from natural language issues. Since tag matching is done only within a single user's tag space, it is more likely that the user will use the same tags when describing another resource. Moreover, they will most likely use the same tag organization structure [9]. Therefore, it is not as necessary to use semantic matching between words.

As recommendation is dependent on tag matching, if there isn't sufficient reuse of tags for users, the system may have problems producing recommendations. In this case, the threshold for recommendation would have to be modified.

3. 3 Contextual CF Recommendation

We have explained the both Contextual CF User Similarity Model and the Contextual CF Recommendation model. They combine to form the formula as shown in equation 4. Finally after that, there is the actual recommendations. There are two recommendation ways: first, as in traditional CF models, any resource with a predicted score above a threshold can be recommended to the user. Thus, depending on the set threshold as well as the user similarity threshold, there will be varying amounts of recommendations.

However, this method ignores that the system has tags available for further filtering. Thus, instead the predicted scores can be used more in a ranking fashion, and that the domain of resource to be searched would only be on the context or tag that the user is interested in. For example, in the system depicted by figure 3, user B comes back to the system. He has an interest in politics and baseball, but today he wants to see some baseball related resources. He searches for 'baseball': Assuming there are users similar to B, he can get not only general tag search, but also personalized tag search with results ranked according to his predicted score calculations based upon this new Contextual CF model. Us-

ing this, the possibilities and potential for personalized and relevant recommendations are great.

4. Conclusions and Future Work

This paper describes a new contextual collaborative filtering model based upon tagging information available from recently popular online social tagging systems. In this model, two areas of traditional CF have been changed: user similarity calculation and score prediction calculation. Together, this new model effectively considers the context in which a user likes a resource. Given that the context is considered, effective recommendations can be made. Moreover, score predictions can be used in different ways from traditional CF, such as ones mentioned in section 3. 3. Additionally, they can be made over a larger domain, not just movies or just music, but rather a domain such as the entire internet. Even users who have differing interests can be effectively recommended for.

This new model has great potential in sites like website tagging site del.icio.us, due to its large domain of the internet itself. Other larger domains would similarly be well suited for such a system. Movies, for example, cover many genres and consequently, not considering the context of why a user likes a resource is not very effective. Rather, with the power of tags and collaborative filtering, one could get relevant personalized recommendations–and only recommendations with the context they are interested in at the time.

From here, we will implement such a model to a more applicable online application to more correctly gauge the strengths and weaknesses of this model. Additionally, further research into tag expansion through natural language processing methods is needed.

Acknowledgments

This research was partly supported by CREST of JST, JSPS (Grant-in-Aid for Scientific Research (A) #15200010), and MEXT (Grant-in-Aid for Young Scientists (B) #17700132).

References

[1] Amazon.com. http://www.amazon.com/.

[2] Citeulike. http://www.citeulike.org/.

[3] Consumating.com. http://www.consumating.com/.

[4] del.icio.us. http://del.icio.us/.

[5] Flickr. http://www.flickr.com/.

[6] movielens. http://www.movielens.umn.edu/.

[7] Scott A. Golder and Bernardo A. Huberman. The structure of collaborative tagging systems. Available from: `http://www.hpl.hp.com/research/idl/papers/tags/tags.pdf`.

[8] Joseph A. Konstan, Bradley N. Miller, David Maltz, Jonathan L. Herlocker, Lee R. Gordon, and John Riedl. GroupLens: Applying collaborative filtering to Usenet news. *Communications of the ACM*, 40(3):77–87, 1997. Available from: `http://citeseer.ist.psu.edu/konstan97grouplens.html`.

[9] C. Marlow, M Naarman, D. Boyd, and M. Davis. Position paper, tagging, taxonomy, flickr, article, toread. 2006. Available from: `http://www.rawsugar.com/www2006/taggingworkshopschedule.html`.

[10] P. Resnick, N. Iacovou, M. Suchak, P. Bergstorm, and J. Riedl. GroupLens: An Open Architecture for Collaborative Filtering of Netnews. In *Proceedings of ACM 1994 Conference on Computer Supported Cooperative Work*, pages 175–186, Chapel Hill, North Carolina, 1994. ACM. Available from: `http://citeseer.ist.psu.edu/resnick94grouplens.html`.