A Constrained Spreading Activation Approach to Collaborative Filtering

Josephine Griffith[1], Colm O'Riordan[1], and Humphrey Sorensen[2]

[1] Dept. of Information Technology,
National University of Ireland, Galway, Ireland
josephine.griffith@nuigalway.ie, colm.oriordan@nuigalway.ie
http://www.it.nuigalway.ie/cirg/
[2] Dept. of Computer Science, University College Cork, Cork, Ireland
h.sorensen@cs.ucc.ie

Abstract. In this paper, we describe a collaborative filtering approach that aims to use features of users and items to better represent the problem space and to provide better recommendations to users. The goal of the work is to show that a graph-based representation of the problem domain, and a constrained spreading activation approach to effect retrieval, has as good, or better, performance than a traditional collaborative filtering approach using Pearson Correlation. However, in addition, the representation and approach proposed can be easily extended to incorporate additional information.

1 Introduction

There is much evidence from various domains within Information Retrieval that the combination of sources of evidence leads to more effective retrieval [7]. With respect to recommender systems, Basu et al. claim that "there are many factors which may influence a person in making choices, and ideally, one would like to model as many of these factors as possible in a recommendation system" [3]. Although the types of information combined and the techniques used vary substantially across the domains, it has been shown that there are advantages to be gained from considering more than one source of information. Two of the primary issues to consider when combining multiple sources of information are that of the representation and the approach that can be used to model the problem.

The aims of this work are to show that not only does a graph representation and spreading activation approach provide good performance but furthermore that both the representation and approach can be easily extended to incorporate additional information. In this paper, the additional information considered is information that can be obtained from the collaborative filtering data set itself. Future work will consider the incorporation of other information.

The collaborative filtering problem space is often viewed as a matrix consisting of the ratings given by each user for items in a collection. Using this matrix, the aim of collaborative filtering is to predict the ratings of a particular user, i,

for one or more items previously not rated by that user. The problem space can equivalently be viewed as a network where users and items are represented by nodes and the links between user nodes and item nodes represents a rating for the item by the user. Several researchers have adopted graph representations in order to develop recommendation algorithms to deal with particular aspects of the recommendation problem (e.g. sparsity [10]). A variety of graphs have been used and a number of graph algorithm approaches have been adopted [1], [10].

In this work, a graph model and spreading activation approach is proposed for the incorporation of additional sources of information. Users and items are represented as nodes in the graph and user nodes and item nodes are connected by weighted edges. A spreading activation approach is used to provide item recommendations. The additional information used is extracted from the analysis of certain features of the dataset. Based on these features, certain nodes (or users and items) of the graph are identified prior to retrieval. These nodes are used to constrain the spreading activation approach thus resulting in only some of the edges being considered and only portions of the graph being traversed. The idea is that this representation and approach can be easily extended to incorporate additional information via more constraints on nodes and via additional weighted edges between nodes.

The paper outline is as follows: Section 2 presents related work and the approach taken is outlined in Section 3. Section 4 details the experiments performed. Results are presented in Section 5 and Section 6 discusses conclusions and future work.

2 Related Work

Given a set of users, a set of items, and a set of ratings, collaborative filtering systems attempt to recommend items for some active user using the ratings of other users, where these users have similar preferences to the active user. The collaborative filtering system essentially automates the "word of mouth" process [17]. It differs from traditional retrieval and filtering systems which return items to some user based on a comparison between the content contained in items (documents) and the content of a user query (information need).

In various situations, collaborative filtering, on its own, is not adequate, e.g. when no ratings exist for an item; when new users join a system; or when a user is not similar to any other users in the dataset. Many different models and approaches have been proposed to overcome these problems. One proposed solution involves utilising any available additional information or evidence (collaborative, content, link, etc.). Work on the combination of different types of information is not restricted to the collaborative filtering domain and within the information retrieval and filtering domains, many approaches have been developed for combining different types of information and combining results from different information retrieval systems [7]. Several authors suggest methods for combining content with collaborative information [2], [14]. There has also been some

work on extracting additional information (or features) from the collaborative filtering data set and using this information to aid in recommendation [15].

2.1 Graph-Based Approaches for Recommendation

Several researchers in Information Retrieval and Collaborative Filtering have adopted graph representations of the problem domain [5], [16], [13]. A variety of graphs have been used (e.g. directed and two-layer) and a number of graph algorithm approaches have been adopted.

In the collaborative filtering domain, Aggarwal et al. present *horting*, a graph-based technique where nodes represent users and directed edges between nodes correspond to the notion of predictability [1]. Predictions are produced by traversing the graph to nearby nodes and combining the ratings of the nearby users. Huang et al. present a number of graph-based representations of the collaborative filtering space using binary ratings. They test a number of associated algorithms, including spreading activation and link prediction approaches [10], [11]. In [10] the goal of the work was to deal with the sparsity problem in collaborative filtering by investigating transitive relationships using a graph representation and spreading activation approaches. Users and items are represented using a bi-partite graph representation where the transactions of users and user feedback are modelled as binary weighted edges connecting the nodes between the two sets of nodes. The authors noticed the effect of over-activation but did not constrain the spreading activation approach.

2.2 Spreading Activation

The idea of spreading activation originated from the field of Psychology and was first used in Computer Science in the area of Artificial Intelligence to process semantic networks. Spreading activation approaches have been used in many Information Retrieval applications [5] and more recently in the domain of collaborative filtering [10]. Spreading activation approaches have also been used to integrate sources of evidence and information [18], [6].

The spreading activation model involves a number of iterations where each iteration consists of one or more “hops” or “pulses” where a hop involves the spreading of activation from one node to all other nodes connected to it. A hop may also include a pre-adjustment stage and a post-adjustment stage which usually involves a decay factor. The activation of a node is calculated by summing the output from each of the nodes connected to it by the weight on the link connecting the two nodes. The output or activation of a node is usually a function of the activation value where many different functions can be used (e.g., a threshold function, a sigmoid function, etc.). A number of problems with this basic spreading activation approach have been identified, one of which is that activation can spread to all nodes in just a few hops. This problem can be overcome by the use of constraints in the pre- and post- adjustment stages. These constraints will specify which portions of the network should be traversed for some hop.

3 Methodology

In this paper, a graph representation is used consisting of a set of user nodes and a set of item nodes. User and item nodes are connected via weighted edges where the weights on the edges represent ratings given to items by users (See Fig. 1). Each user node and item node has an associated activity, output, constrained flag and threshold. The activity of a user or item node a, for N nodes connected to the node a with non-zero weight, is calculated by:

$$activity_a = \sum_{i=1}^{N} x_i w_i$$

where x_i is the output of the node i that is connected to node a and w_i is the weight on the edge connecting node i to node a. The output, $output_a$, of a user or item node is calculated using a threshold function:

$$output_a = \begin{cases} activity_a & \text{if } (constrained_a \neq 0) \text{ and } (activity_a > \tau) \\ 0 & \text{otherwise.} \end{cases}$$

The threshold function uses the node's activity, the constrained flag value (which can be on (1) or off (0)), and a threshold value, τ, where each node may have its own threshold value depending on its importance.

Spreading activation involves moving activation from one set of nodes to a second set of nodes. The terminology of a hop is used in this paper to define the activation spreading from one set of nodes to a second set of nodes. A hop involves the calculation of all node outputs in either the user set or item set, updating the associated activities and outputs of the nodes. The steps for three hops are:

1. hop1: Calculate the activities of all item nodes connected, with non-zero weight, to the current active user node. For each activated item node, calculate the output of the node using the threshold function.
2. hop2: Calculate the activities of all user nodes connected, with non-zero weight, to item nodes where the item nodes have non-zero output. For each activated user node, calculate the output of the node using the threshold function.
3. hop3: Calculate the activities of all item nodes connected, with non-zero weight, to user nodes where the user nodes have non-zero output. For each activated item node, calculate the output of the node using the threshold function.
4. Following three hops, items with the top-20 highest positive activities are recommended to the active user.
5. Steps 2 and 3 can be repeated any number of times before recommendations are given (step 4).

Two hops result in activating a set of user nodes constituting a user neighbourhood of the original active user node. The third hop, from user nodes to item nodes, provides item recommendations for the active user.

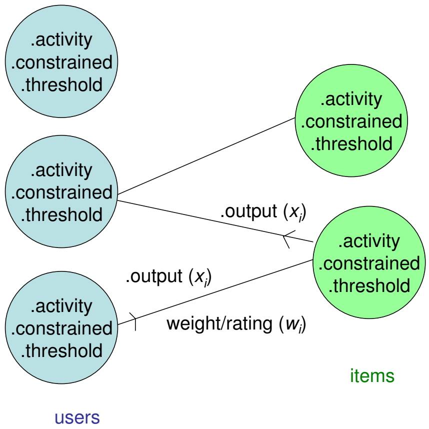

Fig. 1. Graph Representation of Users, Items and Ratings

Much implicit information about users and groups can be extracted from the collaborative filtering data set and can be represented in this graph model via the threshold value, the constrained flag, and the values of the weighted edges. The initial experimental work in this paper only considers two of the simpler user and item features and incorporates information relating to these via the constrained flag. The user and item features considered are:

- *rated*: the number of items rated by some user in comparison to the maximum and minimum number of items rated by users.
- *avg-item-popularity*: the number of ratings received by some item in comparison to the maximum and minimum ratings received by items.

Based on the values of these features certain user and item nodes are identified prior to filtering and the constrained flag value of these nodes is set to 1. This results in only some of the edges being considered and only portions of the graph being traversed.

4 Experiments

The experiments involve the comparison of two collaborative filtering approaches: a constrained spreading activation approach and a traditional memory-based collaborative filtering approach. The reason for choosing a traditional memory-based collaborative filtering approach is that it has been shown to perform well in comparison to many other collaborative filtering techniques [4], [8]. Also it is quite similar to the spreading activation approach outlined.

The aim is to show that the constrained spreading activation approach can perform as well, or better, than a traditional memory-based collaborative filtering approach. The important difference between the representations and approaches is the flexibility of the graph-based and spreading activation approach

in allowing the incorporation of additional information. In this work some simple constraining of nodes is performed but future work will hope to show improved performance over a traditional memory-based collaborative filtering approach, and other collaborative filtering approaches, by the incorporation of additional information via further more complex constraints, varying values for τ (the threshold), and further links between all sets of nodes.

A standard subset of the Movie Lens dataset is considered that contains 943 users and 1682 movies. Weights on the network edges indicate the strength of like/dislike for an item where "dislike" can be viewed as an inhibitory or negative rating and "like" can be viewed as an excitatory or positive rating. Given that the original rating values in the Movie Lens dataset are all positive numbers, the approach adopted maps the ratings to positive and negative values to indicate positive and negative influences. The mapping chosen is to subtract 2.5 from all non-zero values which will give ratings around 0, giving:

$\{0, 1, 2, 3, 4, 5\} \rightarrow \{0, -1.5, -0.5, 0.5, 1.5, 2.5\}$

A proportion of the dataset is removed for testing and the metric of precision is used to compare the performance of the two approaches at different recall points. Precision is used because the spreading activation approach returns a ranking of recommended items, not prediction values that can be compared with actual values. In addition, work by Herlocker et al. [9] suggests that distinguishing between exact actual and predicted rating scores is not as important as measuring if the system correctly or incorrectly identifies "good" or "liked" items.

The traditional collaborative filtering approach uses Pearson Correlation to find correlated (similar) users. An adjustment is used in the Pearson Correlation calculation based on the number of items that users have rated in common (co-rated items). The adjustment ensures that users must have similar preferences over more than a few items to be considered highly correlated [12]. The "nearest" neighbours of a user are selected using a low neighbour selection threshold, with any correlation value greater than 0.01 being considered. Work by Herlocker [8] using the Movie Lens data set found that experiments with no threshold (using all correlation values > 0) always outperformed experiments with higher thresholds.

In the spreading activation approach to collaborative filtering, three stages corresponding to the three stages in the traditional memory-based collaborative filtering approach are used. Neighbours of some active user are found after 2 hops of the approach, at which stage user nodes that have non-zero activity are the neighbours of the active user. When activation is spread again, from user nodes to item nodes, items not rated by the active user will be highlighted. These items are recommended to the user if the activity is sufficiently high. At any stage, a node may be constrained in spreading its activation if it has been previously identified based on the analysis of user and item features. The threshold value used in these experiments is 0 for all nodes, i.e. all positive activities will result in a node outputting a value.

5 Results

Figure 2 illustrates the precision recall graph for the two approaches. Results were averaged over 100 runs. It can be seen that even with the limited sources of additional information included in this representation, the constrained spreading activation approach outperforms the traditional memory-based approach at all recall points other than at the first returned recommendation. These results were shown to be statistically significant using a 2-tailed paired t-test at p-value $<$ 0.05.

Fig. 2. Comparing Constrained Spreading Activation and Traditional Memory-based Approaches to Collaborative Filtering

6 Conclusions

We have presented a graph representation of the collaborative filtering space and a spreading activation approach to collaborative filtering using features of the dataset to constrain the activation approach. Such a flexible representation and approach allow for the incorporation of additional information. Results show that the performance of the approach is better than a traditional memory based approach. Future work will consider more complex user features, as well as group features, and will show that the graph representation and algorithm outlined and demonstrated in this paper can easily be extended to incorporate additional sources of evidence.

References

1. C.C. Aggarwal, J.L. Wolf, K.-L. Wu, and P.S. Yu. Horting hatches an egg: A new graph-theoretic approach to collaborative filtering. In *Proceedings of the Fifth ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'99), San Diego, CA*, pages 201–212, 1999.

2. M. Balabanovic and Y. Shoham. Fab: Content-based, collaborative recommendation. *Communications of the ACM*, 40(3):66–72, 1997.
3. C. Basu, H. Hirsh, and W. Cohen. Recommendation as classification: Using social and content-based information in recommendation. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence (AAAI-98)*, pages 714–721, 1998.
4. M.L. Calderon-Benavides, C.N. Gonzalez-Caro, J. Perez-Alcazar, J.C. Garcia-Diaz, and J. Delgado. A comparison of several predictive algorithms for collaborative filtering on multi-valued ratings. In *Proceedings of the 2004 ACM Symposium on Applied Computing*, pages 1033–1039, 2004.
5. P. Cohen and R. Kjeldsen. Information retrieval by constrained spreading activation on semantic networks. *Information Processing and Management*, 23(4):255–268, 1987.
6. F. Crestani and P.L. Lee. Searching the web by constrained spreading activation. *Information Processing and Management*, 36:585–605, 2000.
7. W.B. Croft. Combining approaches to information retrieval. In *Advances in Information Retrieval*, pages 1–36. Kluwer Academic Publishers, 2000.
8. J.L. Herlocker. *Understanding and Improving Automated Collaborative Filtering Systems*. Phd thesis, University of Minnesota, 2000.
9. J.L. Herlocker, J.A. Konstan, L.G. Terveen, and J.T. Riedl. Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems (TOIS)*, 22:5–53, 2004.
10. Z. Huang, H. Chen, and D. Zeng. Applying associative retrieval techniques to alleviate the sparsity problem in collaborative filtering. *ACM Transactions on Information Systems*, 22(1):116–142, 2004.
11. Z. Huang, W. Chung, and H. Chen. A graph model for e-commerce recommender systems. *Journal of the American Society for Information Science and Technology*, 55(3):259–274, 2004.
12. M.R. McLaughlin and J.L. Herlocker. A collaborative filtering algorithm and evaluation metric that accurately model the user experience. In *Proceedings of the 27th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 329–336, 2004.
13. B. Mirza, B. Keller, and N. Ramakrishnan. Studying recommendation algorithms by graph analysis. *Journal of Intelligent Information Systems*, 20:131–160, 2003.
14. C. O'Riordan and H. Sorensen. Multi-agent based collaborative filtering. In M. Klusch et al., editor, *Cooperative Information Agents 99, Lecture Notes in Artificial Intelligence*, 1999.
15. J. Palau, M. Montaner, and B. Lopez. Collaboration analysis in recommender systems using social networks. In *Cooperative Information Agents VIII: 8th International Workshop, CIA 2004*, pages 137–151, 2004.
16. M.F. Schwartz and C.M. Wood. Discovering shared interests using graph analysis. *Communications of the ACM*, 36:78–89, 1993.
17. U. Shardanand and P. Maes. Social information filtering: Algorithms for automating word of mouth. In *Proceedings of the Annual ACM SIGCHI on Human Factors in Computing Systems (CHI '95)*, pages 210–217, 1995.
18. G.-R. Xue, S. Huang, Y. Y., H.-J. Zeng, Z. Chen, and W.-Y. Ma. Optimizing web search using spreading activation on the clickthrough data. In *Proceedings of the 5th International Conference on Web Information Systems*, 2004.