Weighted Hybrid Recommendation for Heterogeneous Networks

Fatemeh Vahedian
Center for Web Intelligence, Depaul University, Chicago, IL 60604
fvahedia@cs.depaul.edu

ABSTRACT

Social media sites accumulate a wide variety of information about users: likes and ratings, friend and follower links, annotations, posts, media uploads, just to name a few. Key challenges for recommender systems research are (a) to synthesize of all of this data into an integrated recommendation model and (b) to support a wide variety of recommendation types simultaneously (items, friends, tags, etc.) One approach that has been explored in recent research is to view this multi-faceted data as a heterogeneous network and use network-based methods of generating recommendations. However, most such approaches involve computationally-intensive model generation resulting in a single-purpose recommender system. Our approach is to create a component-based hybrid model whose components can be reused for multiple recommendation tasks. In this paper, we show how this model can be applied to heterogeneous networks.

1. INTRODUCTION

Social media sites are an important element of today's Internet, drawing millions of users each day. The wealth of information found in these sites makes recommender systems essential. Such sites often must integrate recommendations of many types: for content, for like-minded users, for appropriate tags, etc. Our approach, called the Weighted Hybrid of Low-Dimensional Recommenders (WHyLDR) is designed to support the flexible creation and rapid deployment of a wide variety of recommenders in a heterogeneous environment. We have demonstrated its effectiveness in prior work focusing on social tagging systems [8, 7, 6, 3].

Compared to a homogeneous network of uniform node and edge types, a heterogeneous network is defined by a diversity of objects and relations, three examples of these networks can be seen in Figure 1. The diversity of nodes in heterogeneous networks means that different types of relations can be imagined between nodes.

The wide variety of relations in a social media site supports many recommendation types. One obvious use of the Yelp network is for recommending new businesses to users, but there are a variety of other possible recommendation tasks. Recommending other users to befriend, recommending locations, and recommending categories are all user-focused recommendation tasks. A site like Yelp may also be interested in recommending users to businesses for advertising or marketing purposes. In addition, a user may wish to constrain the recommendations in various ways: looking for a recommended business in a particular category or in a particular location, for example.

What is needed for complex social media sites is a recommendation technique that is responsive to both the complexity of heterogeneous data and the multiplicity of recommendation tasks. The WHyLDR model has been shown to meet both of these needs in the area of social tagging systems, performing basic recommendation tasks with accuracy surpassing that of single-purpose model-based techniques such as tensor factorization, and also supporting a wide variety of recommendation tasks [8].

As in other work with heterogeneous networks [10], the WHyLDR model views the network structure as a set of mappings, or projections from nodes to nodes: $proj_A(n) \rightarrow \{m_0, m_1, ..., m_i\}$ where A is a set of paths, n is a starting node and the ms are nodes reachable from n via some path in A. Sets of paths that pass from one type of node to another over specific categories of edges are known as meta-paths. Table 1 show example of some meta-paths in different heterogeneous networks. Each such projection induces a different profile for user. Using such profiles, we can build standard collaborative filtering components to select neighbors of users and generate recommendations. Individually such components may be relatively weak. However, as we have shown in prior work and confirm here, ensembles of such components can be effective.

It is not obvious that extended meta-paths will yield useful user profiles. However, our prior work with WHyLDR systems for social tagging has shown that, in some cases, components built from longer meta-paths may make a larger contribution to a recommendation hybrid than their more narrowly-focused subsets. Judicious exploration of the meta-path space is therefore important for hybrid construction.

A heterogeneous network is one where there are multiple object types and/or multiple edge types – typically both.The network schemas of the three datasets shown in Figure 1 give overviews of their respective heterogeneous networks by indicating the different object types and the relations that exist between them. A meta-path in a heterogeneous network is a path over the network schema, a sequenced

RecSys'14, October 6–10, 2014, Foster City, Silicon Valley, CA, USA.

ACM 978-1-4503-2668-1/14/10 ...$15.00.
http://dx.doi.org/10.1145/2645710.2653366 .

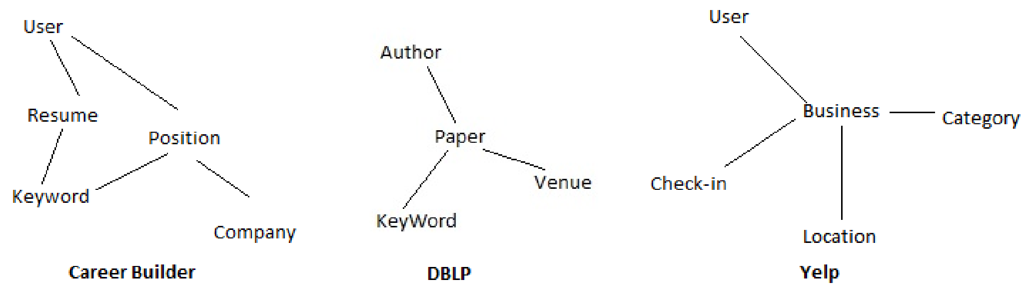

Figure 1: dataset network schema

composition of relations between two object types. Table 1 shows meta-path examples in these networks.

Table 1: Meta-path example in heterogeneous networks

dataset	Meta-paths
DBLP	paper→paper, paper→venue, paper→author, paper→author→paper, author→paper→venue
Career Builder	user→position, user→position→company, user→resume→keyword
Yelp	user→business, user→business→category, user→business→location, user→business→category→business

- DBLP: The Digital Bibliography and Library Project which is a collection of bibliographic information on major computer science journals and proceedings.There are a number of recommendation tasks relevant to this data set. One obvious recommendation task is co-author recommendation, also we can recommend publication venue to authors, paper citation recommendation could be another task is this network.

- Career-Builder: This dataset contains the profiles for job hunter, including education history, previous positions, and location. the main task in this network is to recommend job to user, we can add recommending users to companies that have open positions.

- Yelp: The academic version of Yelp dataset [1]. A variety of types of recommendation can be performed in such a network. Obviously, the recommendation of businesses to users is one, but one can also imagine recommending users to businesses for marketing purposes, recommending categories, even recommending particularly informative reviews. In this paper, all results and experimentations are based on this dataset.

2. CONSTRUCTING THE HYBRID

A weighted hybrid recommender is a system comprised of multiple recommendation components, each of which returns a real-valued score for a combination of user and item. The scores from all the components are combined in a weighted sum [2]. More formally, $s(u,i) = \sum_j \alpha_j s_j(u,i)$ where $s(u,i)$ is the score computed for a user-item combination, $s_j(u,i)$ is the score computed by the jth component, and α_j is the weight associated with the jth component.

The components and the weights are therefore the building blocks needed. The weights are learned through an optimization procedure as discussed below. The components are a function of the recommendation task and the structure of the network.

2.1 Components

WHyLDR components are built on two-dimensional matrices familiar to users of collaborative recommendation [5]. A user-based matrix is one in which the rows are users and the columns are information about user interests. Users are compared on the basis of their profiles and peer users form a neighborhood from which information about unknown items can be extrapolated.[1]

Three types of recommendation component can be considered in a our model. First, user-based KNN models which are constructed based on meta-paths starting from user node. An item-based version of this idea is also well-known. We can follow the meta-path starting from item to create a profile of each item in terms of the users who have rated it then applying a KNN method to make item-based recommendation.

A third type of low-dimensional recommender can also be constructed in which there are two matrices: one for users and one for item with the same columns. Users and item can therefore be compared directly using any one of a number of metrics. We use cosine similarity, so these are the cosine metrics.

Table 2: Recommendation components based on meta-paths

Type	Meta-paths
User-based	UB, UBC, UBL, UBH, UBCB, UBLB, UBHB
Item-based	BU, BL, BC, BH, BLBC, BLBU, BUBU, BUBL
Cosine	UBC, UBH

2.2 Hybrids

One outcome of our prior research on social tagging systems was some surprising trends in the behavior of components formed from long meta-paths. We expected that a component with a long meta-path (for example, ABCD) would be a less effective contributor to the overall recommender than its prefix meta-paths (AB or ABC). Usually

[1] All of the optimizations that have been applied to collaborative recommenders can therefore be applied to the individual WHyLDR components, for example, matrix factorization.

this was the case, but there were important exceptions [4]. Here we investigated this phenomenon by building a number of hybrids, incorporating successively longer meta-paths.

We divided the components into user-based and item-based and by path length (1, 2, and 3) and assembled the following hybrids:

- HM-1: User-based and item-based, paths of length 1 plus popularity
- HM-2: HM-1 plus user-based and item-based, paths of lengths 2
- HM-3: HM-2 plus cosine, paths of length 2
- HM-4: HM-3 plus user-based, paths of length 3
- HM-5: HM-4 plus item-based, paths of length 3

Table 3 gives the component breakdown of each hybrid by meta-path.

3. CONTROLLING COMPONENT GENERATION

There is no requirement that meta-paths be simple: nodes and edges can be revisited, as seen in components like kNN_{UBLB}. Component generation is therefore in principle an unbounded process. There are significant computational costs in generating components and in optimizing a hybrid with a large number of components. In addition, some components may make only a minor contribution to performance. It is therefore important to control this process – ideally, we would like to be able to estimate in advance what components are likely to make a substantial contribution and be able to trade off expected accuracy against the computational costs of adding another component.

As part of our experiments with this data, we used mutual information related to each meta-path to estimate the utility of recommendation components. For a given two-dimensional projection AB, the mutual information can be calculated as

$$I(A, B) = H(A) - H(A|B)$$

where $H(A)$ is entropy of dimension A and $H(A|B)$ is the conditional entropy.Entropy is calculated as

$$H(A) = -\sum_i p(a_i) log(p(a_i))$$

The entropy is therefore a function of the probability of occurrence of nodes in each dimension. In our networks, we define probability of node a_i from dimension A as

$$p(a_i) = \frac{Degree(a_i)}{\sum_i Degree(a_i)}$$

Conditional entropy measures the uncertainty of one dimension given another dimension. Considering an AB projection of a network, we make use of a two-dimension matrix to calculate probability of dimension B given A $P(B|A)$ as follows. The likelihood of reaching node b in dimension B is considered to be the fraction of paths from node a leading to node b out of all possible paths from node a. The conditional probability is therefore calculated as

$$P(b_i|a_j) = \frac{\#path(a_j \to b_i)}{\sum_i \#path(a_j \to b_i)}$$

For example, consider the user-business meta-path and associated recommendation component. The values for $H(U)$ and $H(U|B)$ can be calculated using the formulas above. If these values were roughly the same then the $I(U, B)$ will be around zero. This suggests that the meta-path does not add much information beyond what is already contained in the U dimension and that the UB meta-path is unlikely to give any additional contribution. The same principle can be applied to any user-based or item-based component.

To test this hypothesis, we measure the correlation between mutual information and the weight of components in the hybrid. If our information metric is a good measure of component contribution, we should expect a positive correlation between it and the weight of the corresponding component in the hybrid model.

4. RESULTS

For the experiments reported here, we used the **Yelp** dataset, Following the methodology [4] to build recommender models. Also we make use of Particle Swarm Optimization (PSO) [9] to learn α values for the weighted hybrid model. Figure 2 shows the recall-precision curve for the top individual recommendation component (kNN_{UB}) and the five hybrids. The HM-1 and HM-2 models do not show significant improvement over kNN_{UB} method, but the other three larger hybrids do show significant improvements over the simple collaborative component. The HM-5 hybrid, which incorporates all the components, is dominant. The effect of each

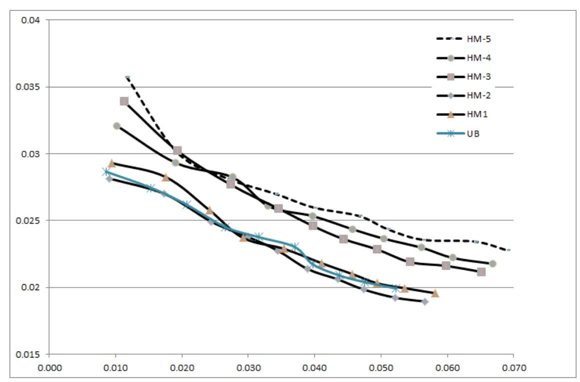

Figure 2: Recall vs. Precision

dimension of the network can be seen in the average learned α weights shown in Figure 3. Here the components with the largest contribution are colored red, fading to pink for those with a lower α value. The white cells are components not present in a particular hybrid. Unlike some of the our other datasets where overall popularity had minimal utility, the non-personalized popularity method contributes relatively strongly in HM-1, on par with location (kNN_{BL} component). However, we can see its contribution decreases in the models including longer meta-paths, as this popularity comes to be represented through particular paths. kNN_{BU} is the strongest contributor in all hybrid models, with kNN_{UB} placing second. This may be because part of the effectiveness of kNN_{UB} is due to a popularity effect. As noted above, location has a strong effect in the hybrid models. The kNN_{UBL} component makes a good contribution in HM-3. Its contribution decreases in HM-4 and HM-5, no doubt because these

Table 3: Component composition of the hybrid models

	UB + BU + BC BH + BL	UBC + UBL + UBH	UBCB + UBHB + UBLB	Cos-UBC + Cos-UBH	BLBC + BLBU + BUBL + BUBU
HM-1	X				
HM-2	X	X			
HM-3	X	X		X	
HM-4	X	X	X		
HM-5	X	X	X	X	X

models include kNN_{UBLB} a meta-path-based component that also incorporates location information.

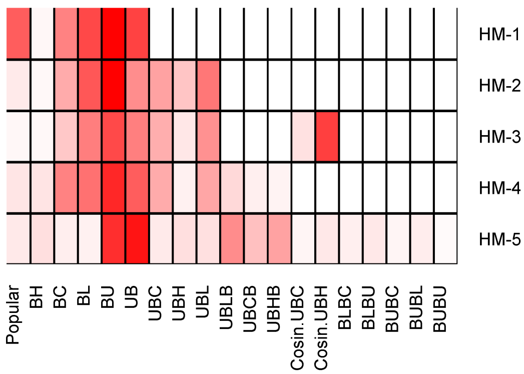

Figure 3: Heatmap of α values

Table 4: Correlation between mutual information and α values

Model	HM-1	HM-2	HM-3	HM-4	HM-5
Correlation	0.788	0.523	0.587	0.90	0.627

Table 4 shows the correlation between the mutual information measure described above and the learned α values in the various hybrids.[2] The strong correlations (all greater than 0.50) suggest that this metric may be a useful way to estimate the value of different components in a hybrid. We intend to make use of this finding in future work.

5. CONCLUSIONS

A key challenge in social media recommendation is to integrate the different types of information available in such systems to enhance recommendations and to offer recommendations of multiple types. In this work, we describe the WHyLDR approach that constructs a linear-weighted hybrid model from simple two-dimensional components based on meta-path traversals of a social media network. These components are then combined using weights learned through optimization. WHyLDR hybrids have been shown to be successful in social tagging applications. In addition to its effectiveness and simplicity, this approach has the benefit of creating re-usable components that can be applied to multiple recommendation tasks (e.g. friend recommendation or user recommendation for marketing).

As there are an unbounded number of possible components, our work raises the question of how to choose components for a hybrid and when to stop adding to it. Our study correlating mutual information with the α values optimized for each hybrid suggest that we may be able to use this measure or a similar one to control hybrid construction. Open questions include how to predict the contribution of components involving multiple paths (such as the cosine components) and how to factor in the influence of the recommendation task.

[2]This calculation did not include the cosine-based components.

6. ACKNOWLEDGMENTS

I would like to thank my advisor Professor Robin Burke for his support, guidance and valuable feedback.

7. REFERENCES

[1] https://www.yelp.com/academic _dataset.

[2] R. Burke. Hybrid recommender systems: Survey and experiments. *User Modeling and User-Adapted Interaction*, 12(4):331–370, 2002.

[3] R. Burke and F. Vahedian. Social web recommendation using metapaths. In *RSWeb@RecSys*, 2013.

[4] R. Burke, F. Vahedian, and B. Mobasher. Hybrid recommendation in heterogeneous networks. In *Proceedings of the 22nd Conference on User Modeling, Adaptation and Personalization*, 2014, to appear.

[5] C. Desrosiers and G. Karypis. A comprehensive survey of neighborhood-based recommendation methods. In F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, editors, *Recommender Systems Handbook*, pages 107–144. Springer, 2011.

[6] J. Gemmell, T. Schimoler, B. Mobasher, and R. Burke. Recommendation by example in social annotation systems. *E-Commerce and Web Technologies*, pages 209–220, 2011.

[7] J. Gemmell, T. Schimoler, B. Mobasher, and R. Burke. Tag-based resource recommendation in social annotation applications. *User Modeling, Adaption and Personalization*, pages 111–122, 2011.

[8] J. Gemmell, T. Schimoler, B. Mobasher, and R. Burke. Resource recommendation in social annotation systems: A linear-weighted hybrid approach. *Journal of Computer and System Sciences*, 78(4):1160–1174, 2012.

[9] J. Kennedy and R. C. Eberhart. Particle swarm optimization. In *Proceedings of the IEEE International Conference on Neural Networks*, pages 1942–1948, 1995.

[10] Y. Sun and J. Han. *Mining Heterogeneous Information Networks: Principles and Methodologies*. Synthesis Lectures on Data Mining and Knowledge Discovery. Morgan & Claypool Publishers, 2012.