Weighted Random Walks for Meta-Path Expansion in Heterogeneous Networks

Fatemeh Vahedian, Robin Burke & Bamshad Mobasher
Center for Web Intelligence, DePaul University, Chicago, IL 60604
{fvahedia, rburke, mobasher}@cs.depaul.edu

ABSTRACT

In social networks, users and items are joined in a complex web of relations, which can be modeled as heterogeneous information networks. Such networks include a variety of object types and the rich relations among them. Recent research has shown that a hybrid recommendation approach combining components built from extended meta-paths in the network can improve the accuracy of recommendations in such networks. However most of this recent work is focused on unweighted heterogeneous networks, and simplifying relations by ignoring weight (including user ratings) loses important information. We propose a random walk based model to generate meta-paths in weighted heterogeneous network in which the frequency of edge sampling is a function of edge weight, and demonstrate that performance is improved using this method.

1. INTRODUCTION

Much research has been focused on recommender systems in heterogeneous networks using extended relations based on meta-paths, showing that accuracy can be improved over the use of simple direct relations. This benefit has been demonstrated with several different recommendation algorithms including multi-relational matrix factorization [7, 8], weighted hybrid of low-dimensional components [2, 1, 6, 5] and non-negative matrix factorization [9, 10].

However, these models all assume a uniform preference associated with all relations in the network. In many networks, however, users provide rating values that can provide useful information. The lack of user rating values in generating meta-paths can be misleading. For example, in a movie recommender, if user u gave movie $m1$ the lowest possible rating of 1, then it would be unproductive to treat this connection as having the same value as another movie $m2$ to which the user gave a higher rating. So, where such information is available, it is essential to rank differently the paths starting from highly rated movies and the ones that are not as interesting to the user.

In general, if we know that some edges represent weaker connections and some edges represent stronger connections, it is sensible to treat the weaker connections differently to avoid adding noise to the recommendation model. The only previous work on weighed heterogeneous networks breaks each meta-path into various similar rated paths [10] then combines all of them together. Since generating those meta-paths is computationally expensive in general, this method is fairly time consuming. In this work, we propose a random walk sampling method in a heterogeneous weighted graph in which meta-paths are generated using exponential sampling that prefers highly rated edges, and build a recommendation model from the resulting collection of paths.

2. RANDOM WALK SAMPLING

Christoffel et al.[3] introduced a random walk sampling algorithms to calculate the transition probability in a random walk model to rank items and generate recommendation model. This model works from an unweighted bipartite graph which represents the binary relations among user and items. Building on this approach, we propose a method for generating meta-paths in a heterogeneous network using biased random walk sampling. This method has the advantage of creating greater efficiency in meta-path generation and allowing for sensitivity to user ratings.

The goal of meta-path generation is to create a relation based on paths through the network. For example, the extended $user - movie - actor - movie - user$ meta-path enables the system to start with a given user and find other users that have watched movies containing actors in common with the user's movies. The semantics of this operation of meta-path expansion is that the end result is a set of destination nodes (in this case, users) weighted by how many of the expanded meta-paths reach that node. A random-walk version of this process chooses edges from the next relation in the meta-path randomly instead of following all possible paths. This is more efficient than generating all paths and the number of samples can be chosen to be large enough to provide a good estimate of what a full expansion would provide [3]. In this work, we look specifically at networks involving a single "rating" edge from user to item. In other words, the first connection from a user is assumed to be to an item and is assumed to have a weight that represents the user rating with higher rated items being more preferred. This construction is common in recommendation contexts where users' quantitative preferences can be gathered.

Random walk meta-path expansion therefore uses the process shown in Algorithm 1. The algorithm takes as input a

Algorithm 1 Random walk meta-path generation

Require: $l \leftarrow [u]$ // Initialize path with starting node: user
Require: $m \leftarrow$ metapath // Queue of edge types
  function GENERATE(l,m)
    if $m \neq \{\}$ **then**
      $me \leftarrow$ POP(m); // Next edge type
      $n \leftarrow l[1]$ //Current node
      $E \leftarrow$ GETEDGES(n, me) // Get edges of type me
      if $me =$ user-item **then**
        $\langle n, j, v \rangle \leftarrow$ WSAMPLE(E) //weighted
      else
        $\langle n, j, v \rangle \leftarrow$ USAMPLE(E) //uniform
      PUSH(j, l); //Add node j to path
      MPGENERATE(l,m)
    else
      return l

list with a single user as the start node and returns a single random walk guided by the meta-path. The functions *USample* and *WSample*, which is omitted for reasons of space, each returns an edge from the list. In the case of *USample* all edges have equal probability. In the case of *WSample*, the edge is chosen with probability proportional to e^w.

3. EXPERIMENTS AND RESULTS

In order to test the meta-path algorithm and its impact on recommendation model we build a movie recommendation using multi-relational matrix factorization (DMF) [4, 8] by incorporating additional relations built from extended meta-paths. For this paper, we used randomly selected a 33% subset of the MovieLens 1M dataset [1]. We generated four meta-path relations starting from the user($user \rightarrow movie \rightarrow actor, user \rightarrow movie \rightarrow director, user \rightarrow movie \rightarrow actor \rightarrow movie, user \rightarrow movie \rightarrow director \rightarrow movie$). The models were optimized using BPR as the optimization criterion (BPR-opt), as described in [4]. Figure 1 shows the

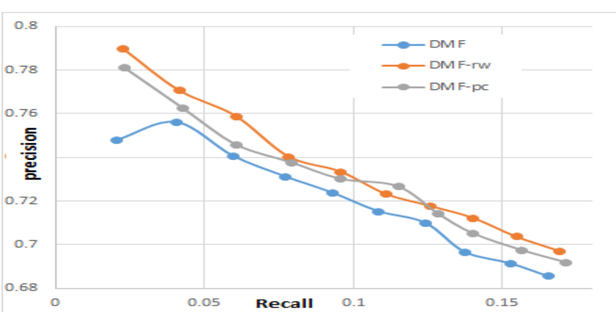

Figure 1: Recall vs. precision for MovieLens dataset

results for recall and precision on recommendation lists of length one through ten for the three recommendation models. *DMF-rw* ,the model using extended paths through random walk, outperforms other generated models. *DMF-pc* represents the model using meta-paths generated in a traditional way, finally *DMF* shows the result for models using just direct link of the network.

As anticipated, random sampling for meta-path generation is much faster than generating the full meta-path relations. On our test machine, the random walk method takes less than 5% of the time required by the baseline technique to generate *UMA* and *UMAM* meta-paths. Since meta-path generation is a major portion of the overall learning time for this system, the random sampling technique would be strongly preferred even if its accuracy were not better.

[1] http://grouplens.org/datasets/movielens/

4. CONCLUSION

We proposed a weighted sampling method to generate meta-paths in weighted heterogeneous network. The results show multi-relational matrix factorization recommendation using those meta-paths can be enhanced comparing to previous method. Additionally, random walk based meta-path generation is more efficient while not sacrificing accuracy. As our future work, we will be exploring other multi-relational algorithms and hybrid model using weighted sampling and extend the application of weighted sampling to other types of weighted edges.

Acknowledgments

This work was supported in part by the National Science Foundation under Grant No. IIS-1423368 (Multi-dimensional Recommendation in Complex Heterogeneous Networks).

5. REFERENCES

[1] R. Burke and F. Vahedian. Social web recommendation using metapaths. In *RSWeb@RecSys*, 2013.
[2] R. D. Burke, F. Vahedian, and B. Mobasher. Hybrid recommendation in heterogeneous networks. In *UMAP 2014*, pages 49–60, 2014.
[3] F. Christoffel, B. Paudel, C. Newell, and A. Bernstein. Blockbusters and wallflowers: Accurate, diverse, and scalable recommendations with random walks. In *Proceedings of the 9th ACM Conference on Recommender Systems*, RecSys '15, pages 163–170, New York, NY, USA, 2015. ACM.
[4] L. R. Drumond, E. Diaz-Aviles, L. Schmidt-Thieme, and W. Nejdl. Optimizing multi-relational factorization models for multiple target relations. In *CIKM 2014*, pages 191–200, New York, NY, USA, 2014. ACM.
[5] F. Vahedian. Weighted hybrid recommendation for heterogeneous networks. In *RecSys '14*, pages 429–432, 2014.
[6] F. Vahedian and R. D. Burke. Predicting component utilities for linear-weighted hybrid recommendation. In *RSWeb 2014*, 2014.
[7] F. Vahedian, R. D. Burke, and B. Mobasher. Network-based extension of multi-relational factorization models. In *Poster Proceedings of the 9th ACM Conference on Recommender Systems, RecSys 2015, Vienna, Austria, September 16, 2015.*, 2015.
[8] F. Vahedian, R. D. Burke, and B. Mobasher. Meta-path selection for extended multi-relational matrix factorization. In *Proceedings of the Twenty-Ninth International Florida Artificial Intelligence Research Society Conference, FLAIRS 2016, Key Largo, Florida, May 16-18, 2016.*, pages 566–571, 2016.
[9] X. Yu, X. Ren, Y. Sun, Q. Gu, B. Sturt, U. Khandelwal, B. Norick, and J. Han. Personalized entity recommendation: A heterogeneous information network approach. In *WSDM 2014*, pages 283–292, 2014.
[10] X. Yu, X. Ren, Y. Sun, B. Sturt, U. Khandelwal, Q. Gu, B. Norick, and J. Han. Recommendation in heterogeneous information networks with implicit user feedback. In *Proceedings of the 7th ACM Conference on Recommender Systems*, RecSys '13, pages 347–350, New York, NY, USA, 2013. ACM.