Social Recommendation System Using Network Embedding & Temporal Information

*Saranya, *Sowmya A S, *Mohammed Shebin K K, **Anuraj Mohan
*UG Scholar, **Assistant Professor, Department of Computer Science and Engineering
NSS College of Engineering Palakkad

Social Recommendation System (SRS)

Recommendation systems which target the social media domain are termed as Social Recommendation Systems. User interests are captured in order to recommend top-K friends or items, or both.

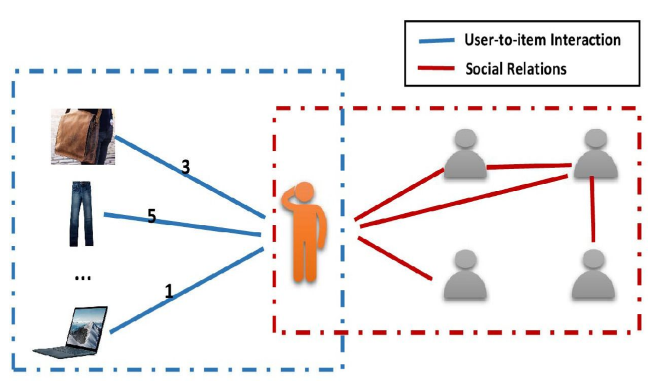

Fig. 1 Pictorial representation to demonstrate how SRS uses social information for recommendation

Heterogeneous User-Item Network (HUI)

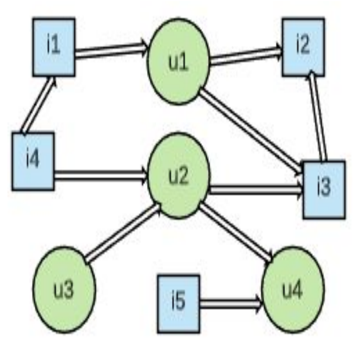

Fig. 2 An HUI Network

An information network with two types of nodes -
1. Users
2. Items

and three types of edges -
1. User-User
2. User-Item
3. Item-Item

Objective

To create a model - **Social Recommendation System** -

- That uses the temporal semantic effects
- Social relationships & user behavior sequential patterns of a heterogeneous information network
- To recommend top-K friends and items

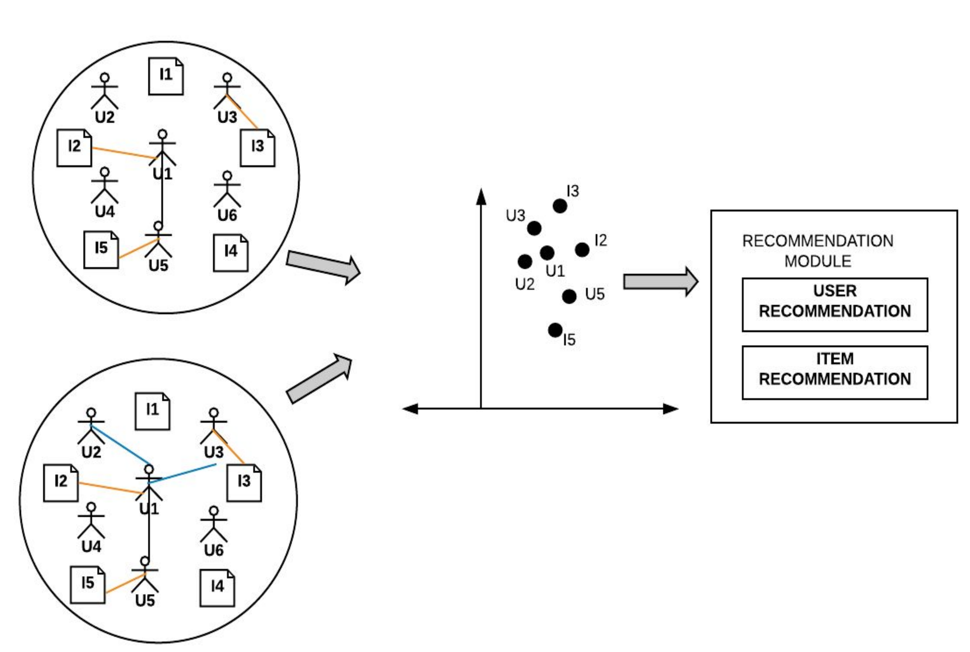

Fig. 3 Flowchart of Graph Based Embedding With Temporal Informations

The updations of links in dataset is captured by -

- Dividing the HUI network into different snapshots
- Generating meaningful metapaths that captures the semantic relationships in the network, corresponding to each snapshot
- Generating node representations / embeddings

Problem Statement

To design a system that can capture both the network property and its evolution, show the importance of network embedding and temporal information for improving the performance of recommendation systems to predict top-K friends and items.

Methodology

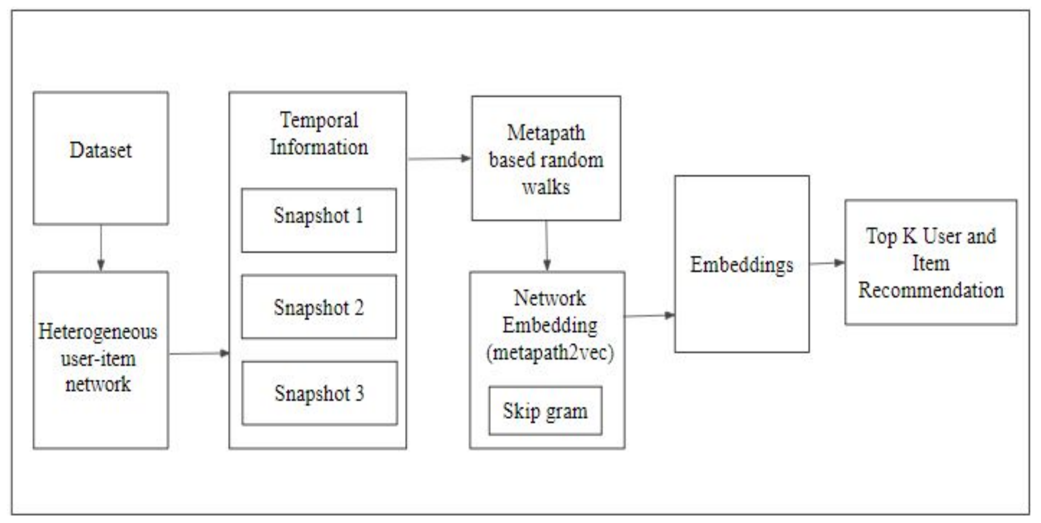

Fig.4 Proposed System

The various steps involved in the process are -

- Construction of HUI
- Metapath generation
- Metapath2vec
- Top-K friend & item recommendation

A. Construction of HUI -
A HUI network shown as in Fig.2 is constructed

B. Metapath generation -
Metapath-based random walks are introduced to generate paths that are capable of capturing the semantic and structural correlations between different types of nodes

C. Metapath2vec distinguishes the context of nodes conditioned on their types while constructing its neighborhood function $N_t(v)$ by maximizing the probability of having the heterogeneous context

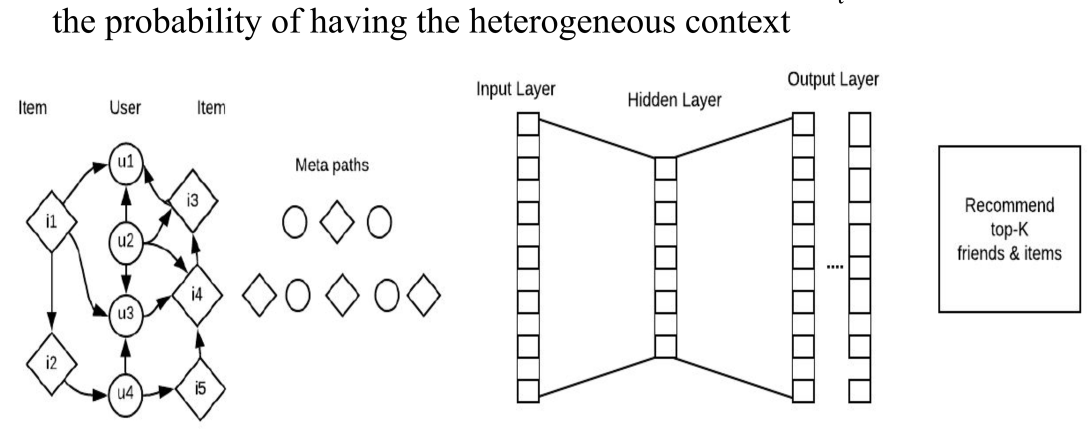

Fig.5 Skip-gram in metapath2vec

D. Given a target user $u_i \in U$ with the query time t, for each user node u_j who has not been connected with u_i, we compute its ranking score, and then select the k ones with the highest ranking scores as recommendations. Top-K friends and items for a particular user is found using cosine similarity

$$\text{Similarity} = \bar{u}_i.\bar{u}_j/\|\bar{u}_i\|.\|\bar{u}_j\|$$

(similarity between users u_i and u_j)

Then, they are sorted and top-K records are extracted. Recommendation is done in an effective manner as potential friends and item information is included using metapath based random walks

Dataset & Tools Used

- Datasets - Last.Fm (song level - tags and similar songs) & MovieLens (describes 5-star rating and tagging activity from MovieLens)
- Timestamp information - used for dividing the dataset into three snapshots

The system was implemented using Python language. Following libraries were used -

- NumPy
- TensorFlow
- Scikit-learn

Datasets	#Users	#Items	#User-user Links	#User-item Links	#Item-item Links
Last.fm	125	205	335	595	425
MovieLens	306	50	600	535	100

Table 1. Statistics of dataset

Results & Analysis

1. Comparison with DeepWalk

Method	Last.fm			MovieLens		
	Recall@3	Recall@5	Recall@10	Recall@3	Recall@5	Recall@10
DeepWalk	0.56	0.60	0.66	0.66	0.71	0.72
Metapath2vec	0.73	0.80	0.84	0.73	0.76	0.82

Table 2. Top-K Friends Recommendation Effectiveness

2. Comparison with & without inclusion of temporal information

Dataset	Timestamp Based	Without Temporal Information
Last.fm	0.83	0.75
MovieLens	0.81	0.72

Table 3. Recall@10 (Friend Recommendation) for Time based comparison

3. Parameter Sensitivity

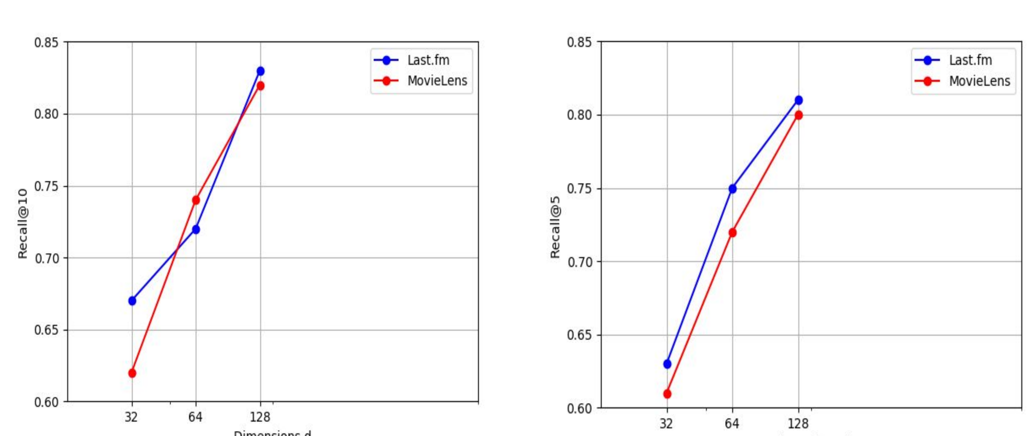

Fig. 6 Recall@10 vs Dimension (Friends Recommendation & Item Recommendation)

Inferences

- The proposed model can recommend top K friends & items more efficiently than the model using DeepWalk.
- By the inclusion of temporal information the model can easily capture the instant interests and the evolution of the network.
- Dimension d=128 is found to give better results compared to other dimensions

System Performance

- The performance is better when Last.fm dataset is used when compared to MovieLens dataset
- Top-10 friends and top-5 item recommendation gives better results (ie. precision value)

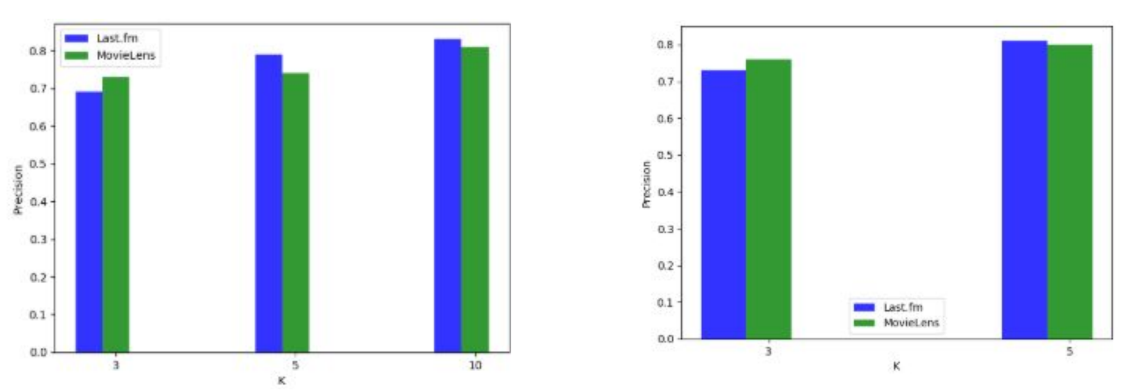

Fig. 7 Precision vs K (Top-K Friends & Item Recommendation)

Conclusion

- The system captures the temporal semantic effects, social relationships and user behavior sequential patterns
- For efficiently handling large-scale social media streams, temporal information is integrated to generate top-K recommendations. The recommendation process is based on the proximity of the related users and items while considering the freshness of the items
- The system is enhanced by using metapath based random walks and metapath2vec for network embedding, thus captures both the structural and semantic correlations of differently typed nodes efficiently

Future Works

- Incorporation of DGE to better capture temporal effects
- Investigation on how to apply deep learning methods to better fuse the embeddings of multiple metapaths
- Metapath2vec++ could be used, which enables the simultaneous modeling of structural and semantic correlations in heterogeneous network
- Threshold based algorithm (TA) can be used for faster and efficient top K friend and item recommendations

References

[1] Liu, P., Zhang, L. and Gulla, J.A., 2019. Real-time social recommendation based on graph embedding and temporal context. International Journal of Human-Computer Studies, 121, pp.58-72.
[2] Dong, Y., Chawla, N.V. and Swami, A., 2017, August. metapath2vec: Scalable representation learning for heterogeneous networks. In Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining (pp. 135-144). ACM.
[3] Shi, C., Hu, B., Zhao, W.X. and Philip, S.Y., 2018. Heterogeneous information network embedding for recommendation. IEEE Transactions on Knowledge and Data Engineering, 31(2), pp.357-370.
[4]Perozzi, B., Al-Rfou, R. and Skiena, S., 2014, August. Deepwalk: Online learning of social representations. In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 701-710).