Graph-based recommendations for big data systems

Lev Denisov, Ferdiansyah Dolot and Anna Turu Pi – Facultat d'Informàtica de Barcelona, Universitat Politècnica de Catalunya

Objectives

Research state of the art in graph-based recommendation systems deployed in large scale production environments.

Identify whether graph-based recommenders overcome limitations of traditional recommenders such as lack of scalability and lack of real-time recommendations.

Production systems

Random walk

- Pixie
- Who To Follow
- RealGraph
- GraphJet

Online motif detection

- MagicRecs

Graph Convolutional Network

- PinSAGE

Motivation

Traditional recommendation systems basic structure is an utility matrix and it entails scalability and sparsity problems, as well as cold-start problem.

Sparsity as each user only rates a small subset of the items available, the utility matrix used will be sparse, leading to unreliable results.

Scalability In a matrix-based system, the growth of users and items is reflected exponentially in the growth of the matrix, hence exponential growth of computations.

Cold-start problem The initial lack of information generates unreliable recommendations.

Real-time recommendations Traditional methods need to be reprocessed from scratch. Would be expensive and entail a high latency.

Although graph-based systems are already in use, the market has not adapted yet to Big Data requirements; the most common system is the collaborative filtering based on matrix decomposition and it encounters those limitations.

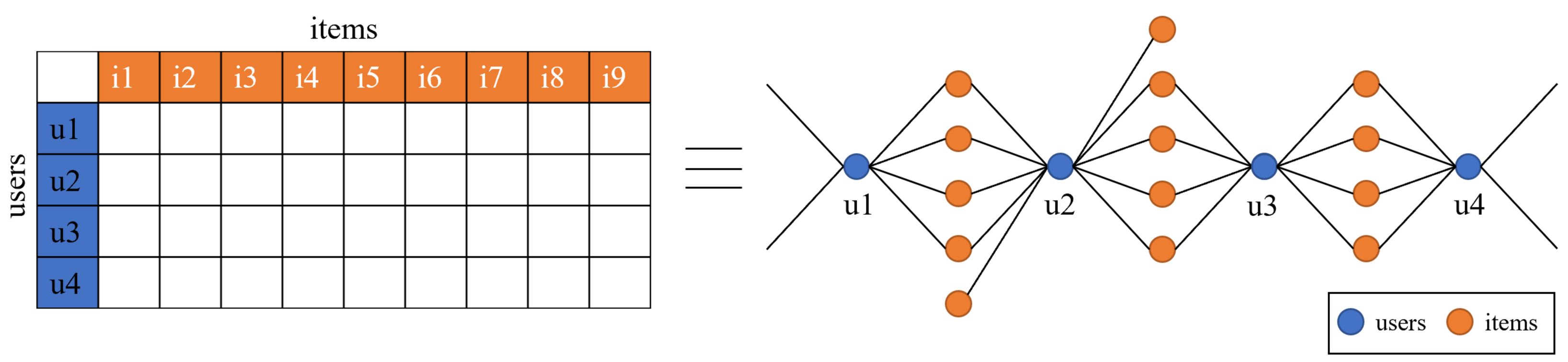
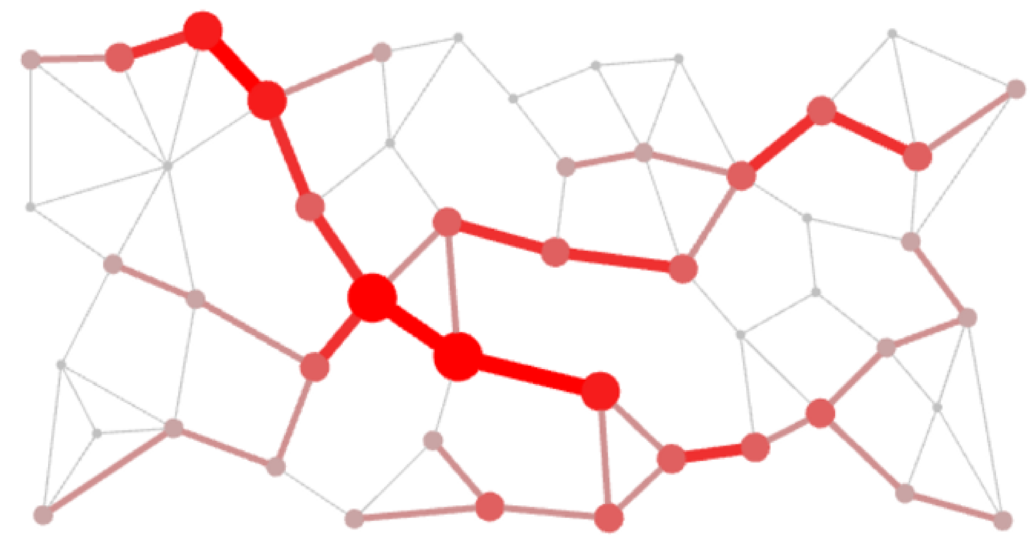


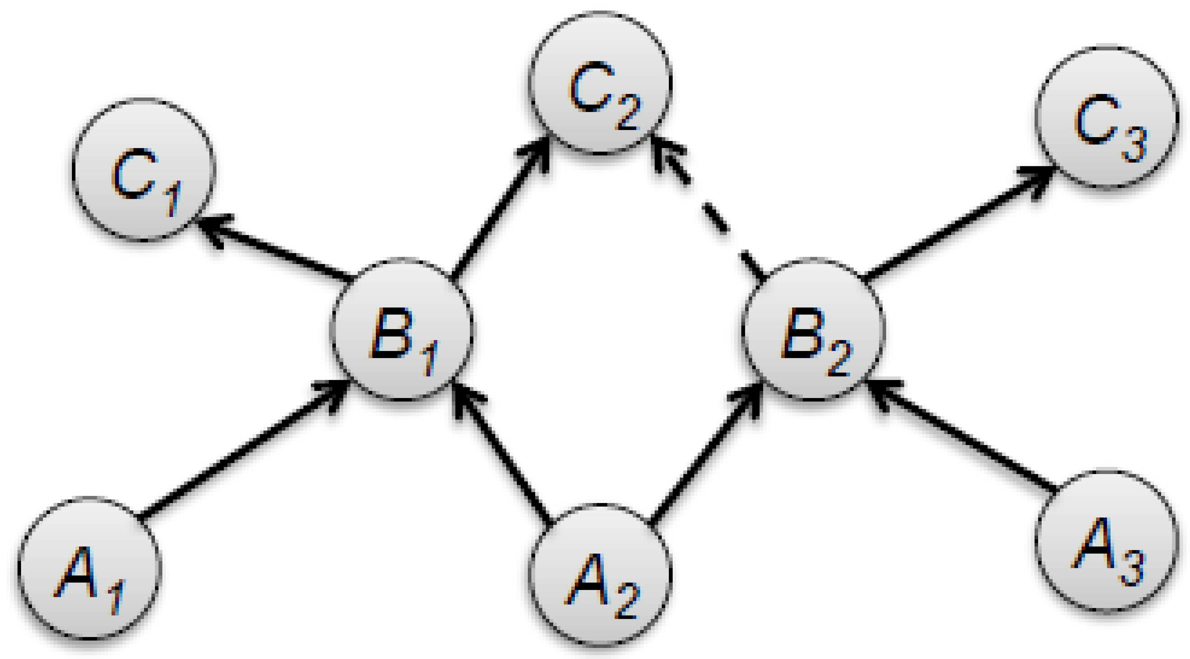
Figure 1: Comparison of a traditional recommendation system basic structure, utility matrix, (left) to a graph based one (right)

Random walk



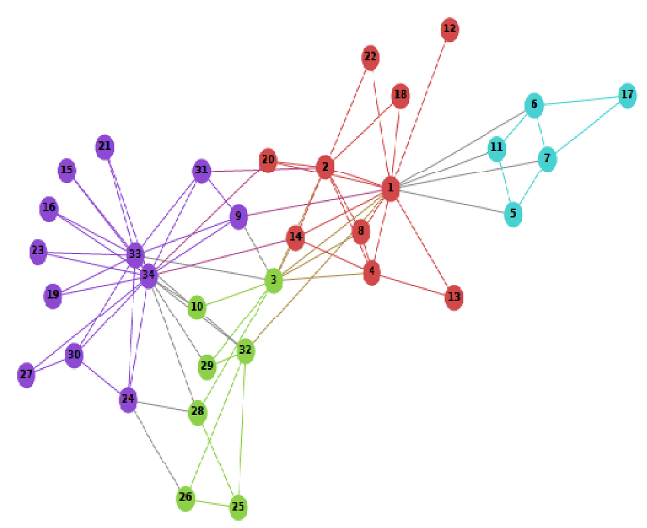
- Stochastic algorithm, efficient neighbor nodes sampling
- Ranks nodes by the number of visits
- Can be applied on the arbitrarily large graphs
- Approximate rating; precision depends on the number of steps
- Can be biased individually for every user
- Can adjust variability by changing walk length

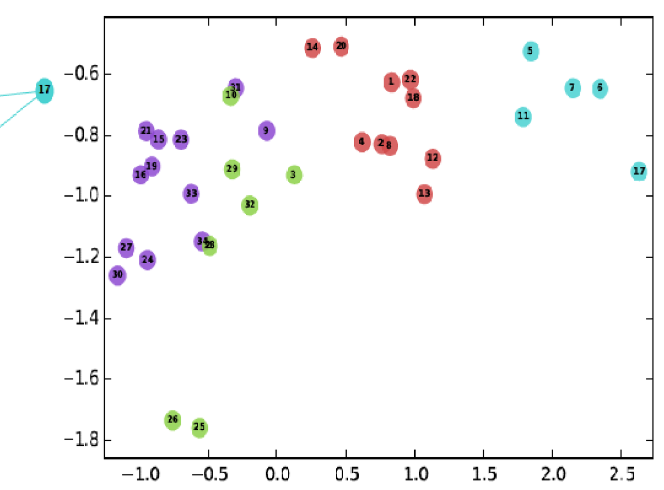
Online motif detection



- Detects motifs – meaningful structures in the graph
- Works with nodes neighborhood, does not depend on graph size
- Detects if a motif formed when new edge is added
- Can be efficiently implemented as intersection of adjacency lists
- If a motif detected push a recommendation

Graph Convolutional Network





- Hybrid algorithm
- Neural network generates embeddings of the graph structure
- Works with nodes neighborhood
- Training optimized with random walks
- Can be combined with content embeddings to improve quality
- Trained network efficiently computes embeddings for new items
- Fast recommendations with Locality Sensitive Hashing

Conclusion

Graph-based recommendations are a viable alternative to the traditional methods. The literature offers graph-based recommendation algorithms that are in lesser degree dependent on the data size allowing them to scale well. Graph-based methods can easily be applied in systems with millions of items providing real time high quality personalized recommendations which is not feasible with the traditional methods. Such systems are conceptually simple and robust. The next step in the state of the art is to provide ready solutions, as only custom solutions are available.

References

[1] Jesús Bobadilla et al. Recommender systems survey. *Knowl.-Based Syst.*, 2013.

[2] A. Sharma et al. Graphjet: Real-time content recommendations at twitter. 2016.

[3] C. Eksombatchai et al. Pixie: A system for recommending 3+ billion items to 200+ million ... 2018.

[4] R. Ying et al. Graph Convolutional Neural Networks for Web-Scale Recommender Systems. 2018.

Contact Information

L. Denisov
denisovlev@ya.ru

F. Dolot
ferdi.dolot@gmail.com

A. Turu Pi
annaturupi@gmail.com