Joint Nonnegative Matrix Factorization for Community Structures Detection in Signed Networks

Chao Yan, Hui-Min Cheng (co-first author),
Yi-Zi Ning, Xin Liu and Zhong-Yuan Zhang*
zhyuanzh@gmail.com

Central University of Finance and Economics

Introduction

Community structures detection in signed networks is very important for understanding not only the topology structures of signed networks, but also the functions of them, such as information diffusion, epidemic spreading, etc. In this paper, we develop a joint nonnegative matrix factorization model to detect community structures. The model has deep relationships with stochastic block model and probabilistic latent semantic indexing. In addition, we propose a modified partition density to evaluate the quality of community structures. We use it to determine the appropriate number of communities. The effectiveness of our approach is demonstrated based on both synthetic and real-world networks.

Methods Description

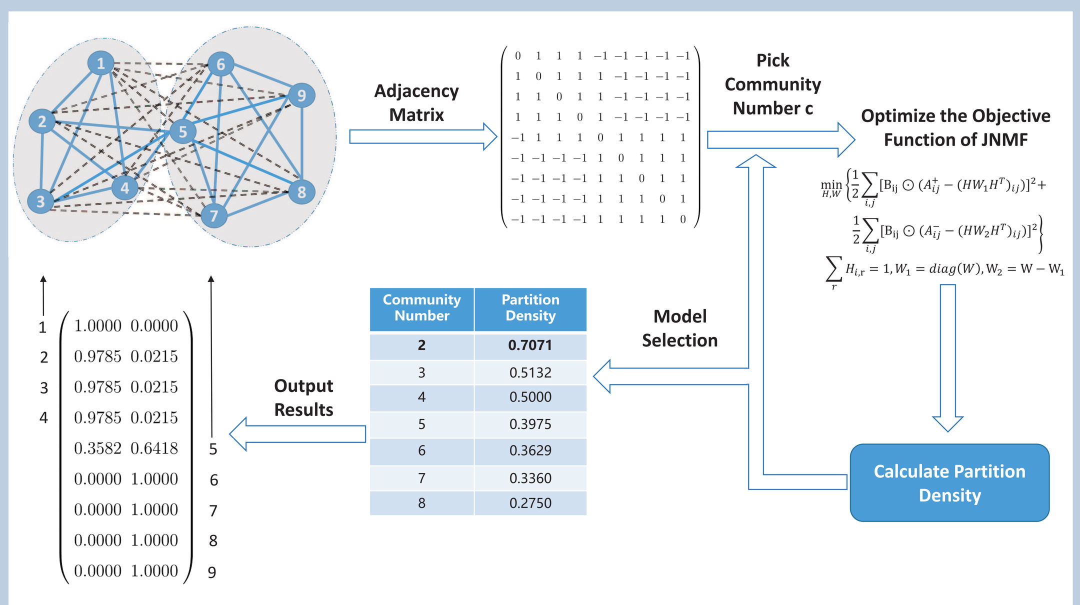

Figure 1: The flowchart of the proposed method. Solid edges in the network are positive, and dashed ones are negative.

Theorem 1 *The objective function is equivalent to the maximum likelihood of stochastic block model.*

Theorem 2 *The objective function is non increasing using the proposed algorithms.*

Conclusions and Future Work

We presented joint nonnegative matrix factorization method to detect community structures in signed networks, and also gave a revised partition density to infer the community number in networks automatically. The experiments conducted on both synthetic and real-world networks showed the effectiveness of the proposed method. In summary, the method is parameter-free, easy to implement.

Interesting problems for future work include generalization of the proposed method to weighted networks and directed networks, and applying to link prediction and recommendation systems in signed networks.

References

[1] Michelle Girvan and Mark EJ Newman. *Proceedings of the national academy of sciences*, 99(12):7821–7826, 2002.

[2] Daniel D Lee and H Sebastian Seung. *Nature*, 401(6755):788, 1999.

Acknowledgements

This work is supported by the Program for Innovation Research in Central University of Finance and Economics, Young Elite Teacher Project of Central University of Finance and Economics, and Disciplinary funding of Central University of Finance and Economics.

Experimental Results on Synthetic Networks

From Fig. 2 and 3, one can observe that: (1) The results of the methods decrease when P_{in} is decreasing or μ is increasing. (2) The proposed method is less sensitive to noise. (3) The proposed method is the best, and the standard deviations are low. For example, when $\mu = 0.6$, the NMI of our method (96.28%) is 73% higher than that of the FEC (23.44%).

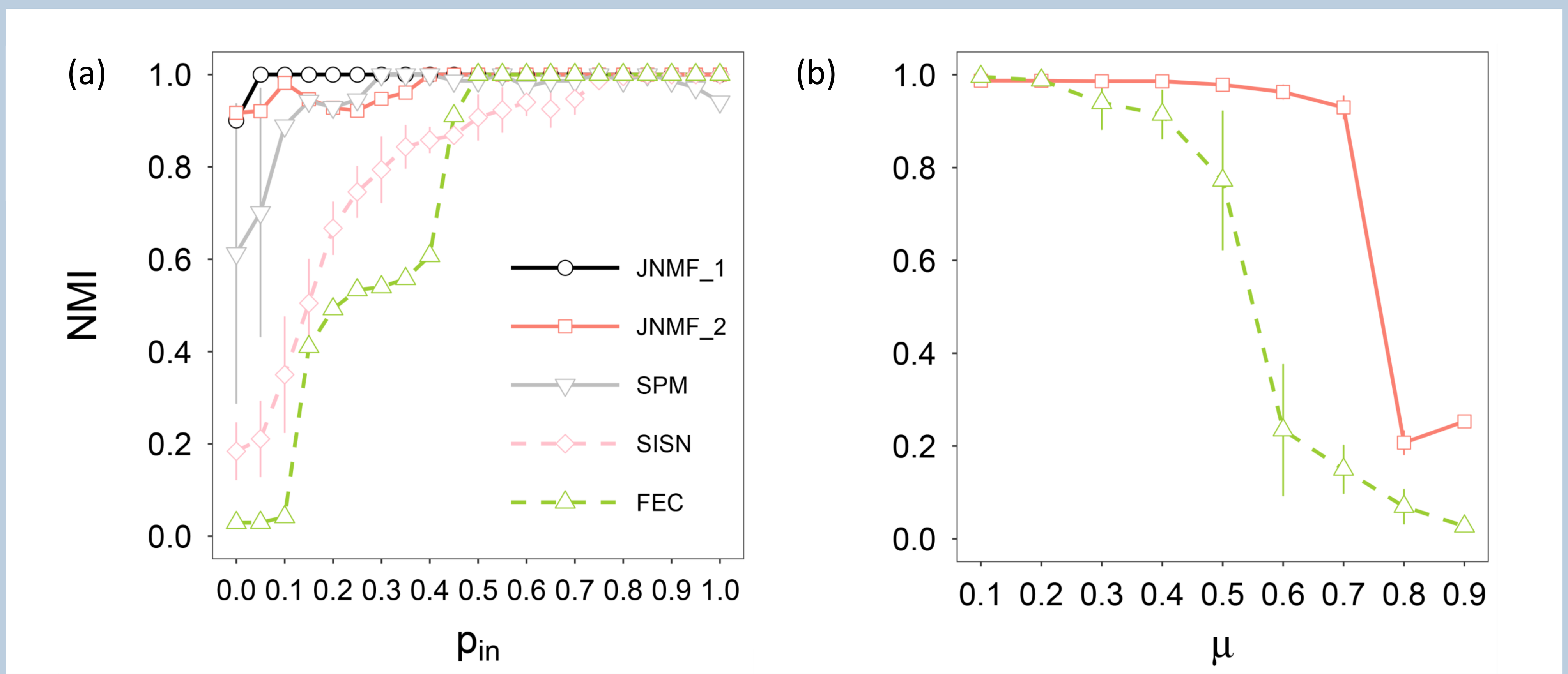

Figure 2: Community detection performance of different methods on SG networks (Left) and SLFR networks (Right).

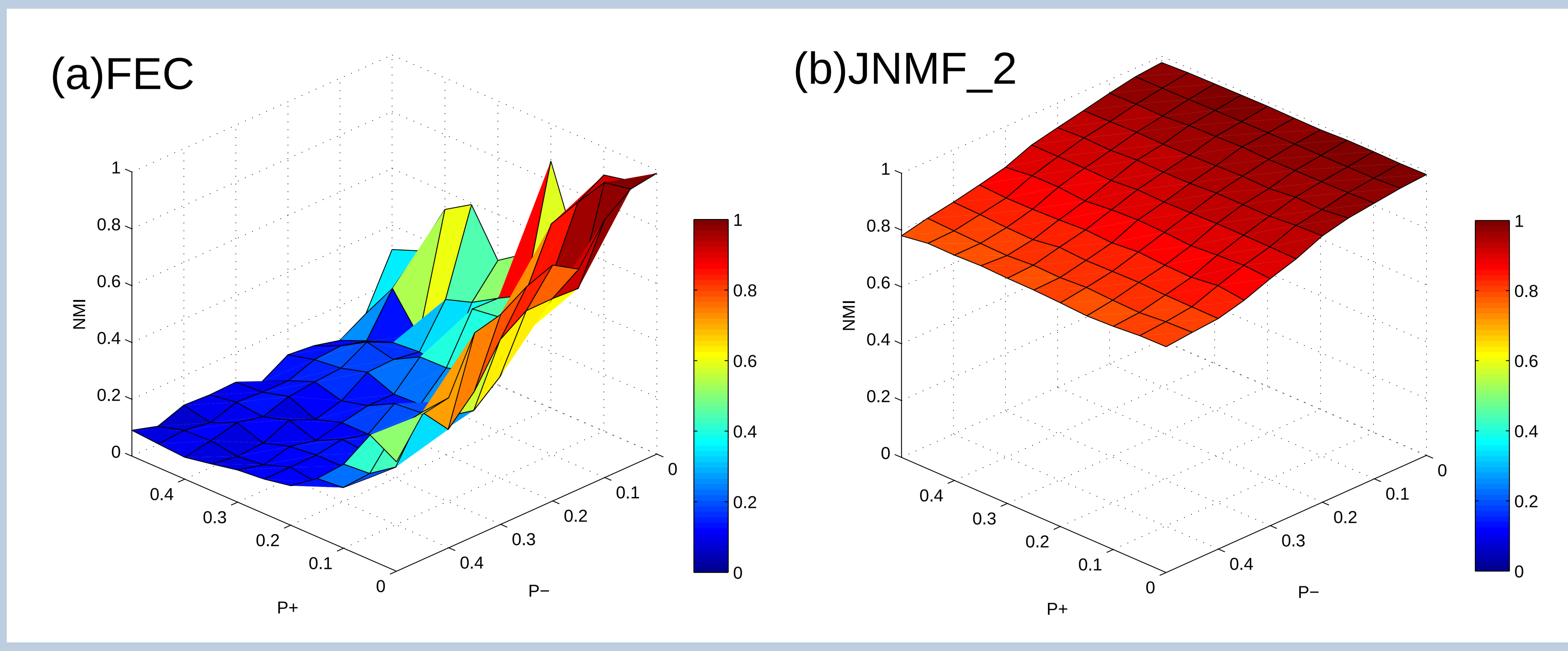

Figure 3: Community detection performance of FEC and our proposed method on SLFR.

Experimental Results on Real-World Networks

From Fig. 4 and Table 1, one can observe that: (1) The inferred numbers of communities are reasonable, and the detected communities are practical, and easy to explain. (2) The time complexity of our method is $O(cn^2)$, making it faster and more suitable for large-scale computation.

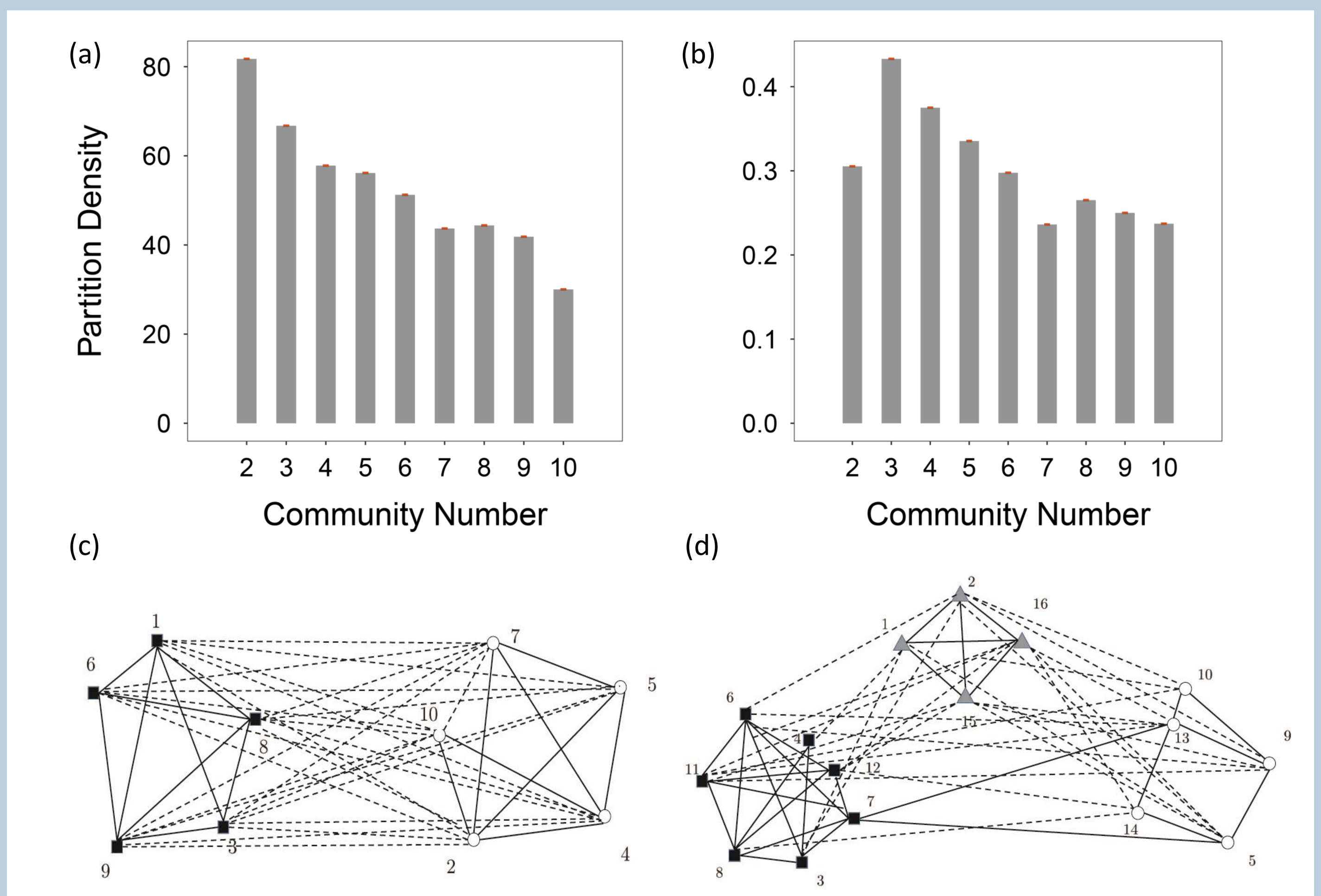

Figure 4: Community detection results of our proposed method on real world networks.

Table 1: Number of communities estimated by different methods. The last row is the time complexity.

	JNMF_2	*SISN*	*FEC*
Slovene parliamentary party	2 ± 0	8.8 ± 0.42	2 ± 0
Gahuku-Gama subtribes	3 ± 0	3 ± 0	4 ± 0
U.S.supreme court	2 ± 0	2 ± 0	2 ± 0
Time complexity	$O(cn^2)$	$O(n^4)$	$O(n^3)$