Appendix of the KDD 2020 paper "Local Community Detection in Multiple Networks"

APPENDIX

The proof of Theorem 1

In Theorem 1, we discuss the weak-convergence of the modified transition matrix $\mathcal{P}_i^{(t)}$ in general multiple networks. Because after each iteration, $\mathcal{P}_i^{(t)}$ needs to be column-normalized to keep the stochastic property, we let $\hat{\mathcal{P}}_i^{(t)}$ represent the column-normalized one. Then Theorem 1 is for the residual $\Delta(t+1) = \| \hat{\mathcal{P}}_i^{(t+1)} - \hat{\mathcal{P}}_i^{(t)} \|_\infty$. In the proof, we will utilize some results in the proof of Theorem 2.

Proof.

$$\begin{aligned}\Delta(t+1) &= \| \hat{\mathcal{P}}_i^{(t+1)} - \hat{\mathcal{P}}_i^{(t)} \|_\infty \\ &= \max\{|\hat{\mathcal{P}}_i^{(t+1)}(x,y) - \hat{\mathcal{P}}_i^{(t)}(x,y)|\} \\ &= \max\{|\frac{\mathcal{P}_i^{(t+1)}(x,y)}{\sum_z \mathcal{P}_i^{(t+1)}(z,y)} - \frac{\mathcal{P}_i^{(t)}(x,y)}{\sum_z \mathcal{P}_i^{(t)}(z,y)}|\}\end{aligned}$$

Based on Eq. (2), we have for all t, $\frac{1}{K} \leq \sum_z \mathcal{P}_i^{(t)}(z,y) \leq 1$.

We denote $\min\{\mathcal{P}_i^{(t+1)}(x,y), \mathcal{P}_i^{(t)}(x,y)\}$ as p and $\min\{\sum_z \mathcal{P}_i^{(t+1)}(z,y), \sum_z \mathcal{P}_i^{(t)}(z,y)\}$ as m. From the proof of Theorem 2 , we know that

$$|\mathcal{P}_i^{(t+1)}(x,y) - \mathcal{P}_i^{(t)}(x,y)| \leq \lambda^t K$$

and

$$|\sum_z \mathcal{P}_i^{(t+1)}(z,y) - \sum_z \mathcal{P}_i^{(t)}(z,y)| \leq \lambda^t K |V_i|$$

Then, we have

$$\begin{aligned}&|\frac{\mathcal{P}_i^{(t+1)}(x,y)}{\sum_z \mathcal{P}_i^{(t+1)}(z,y)} - \frac{\mathcal{P}_i^{(t)}(x,y)}{\sum_z \mathcal{P}_i^{(t)}(z,y)}| \\ \leq& \frac{p + \lambda^t K}{m} - \frac{p}{m + \lambda^t K |V_i|} \\ \leq& \frac{m\lambda^t K + |V_i|(\lambda^t K)^2 + pK|V_i|\lambda^t}{m^2} \\ =& \lambda^t \frac{mK + K^2|V_i|\lambda^t + pK|V_i|}{m^2}\end{aligned}$$

When $t > \lceil -\log_\lambda(K^2|V_i|) \rceil$, we have $\lambda^t K^2 |V_i| \leq 1$. Thus,

$$\Delta(t+1) \leq \lambda^t \frac{mK + K^2|V_i|\lambda^t + pK|V_i|}{m^2} \leq \lambda^t K^2(|V_i| + 2)$$

Then, it's derived that $\Delta(t+1) \leq \epsilon$ when $t > \lceil \log_\lambda \frac{\epsilon}{K^2(|V_i|+2)} \rceil$. □

The proof of Theorem 2

Theorem 2 discusses the weak-convergence of the modified transition matrix $\mathcal{P}_i^{(t)}$ in the special multiplex networks which have the same node set in all networks. In the multiplex networks, cross-transition matrices are just $\mathbf{I}$, so the stochastic property of $\mathcal{P}_i^{(t)}$ can be naturally guaranteed without the column-normalization of $\mathcal{P}_i^{(t)}$. We then define $\Delta(t+1) = \| \mathcal{P}_i^{(t+1)} - \mathcal{P}_i^{(t)} \|_\infty$.

Proof. Based on Eq. (3) and Eq. (2), we have

$$\begin{aligned}\Delta(t+1) =& \| \mathcal{P}_i^{(t+1)} - \mathcal{P}_i^{(t)} \|_\infty \\ =& \| \sum_{j=0}^{K} [\hat{\mathbf{W}}^{(t+1)}(i,j) - \hat{\mathbf{W}}^{(t)}(i,j)]\mathbf{P}_j \|_\infty \\ =& \| \sum_{j \in L_i} [\hat{\mathbf{W}}^{(t+1)}(i,j) - \hat{\mathbf{W}}^{(t)}(i,j)]\mathbf{P}_j \\ &+ \sum_{j \in \bar{L}_i} [\hat{\mathbf{W}}^{(t+1)}(i,j) - \hat{\mathbf{W}}^{(t)}(i,j)]\mathbf{P}_j \|_\infty\end{aligned}$$

where $L_i = \{j | \hat{\mathbf{W}}^{(t+1)}(i,j) >= \hat{\mathbf{W}}^{(t)}(i,j)\}$, and $\bar{L}_i = \{1, 2..., K\} - L_i$. Since all entries in $\mathbf{P}_j$ are non-negative, for all j, all entries in the first part is non-negative and all entries in the second part is non-positive. Thus, we have

$$\begin{aligned}\Delta(t+1) = \max\{&\| \sum_{j \in L_i} [\hat{\mathbf{W}}^{(t+1)}(i,j) - \hat{\mathbf{W}}^{(t)}(i,j)]\mathbf{P}_j \|_\infty, \\ &\| \sum_{j \in \bar{L}_i} [\hat{\mathbf{W}}^{(t+1)}(i,j) - \hat{\mathbf{W}}^{(t)}(i,j)]\mathbf{P}_j \|_\infty\}\end{aligned}$$

Since for all i, score vector $\mathbf{x}_i^{(t)}$ is non-negative, we have $\cos(\mathbf{x}_i^{(t)}, \mathbf{x}_k^{(t)}) \geq 0$. Thus, $\sum_k \mathbf{W}^{(t+1)}(i,k) \geq \sum_k \mathbf{W}^{(t)}(i,k)$. For the first part, we have

$$\begin{aligned}&\| \sum_{j \in L_i} [\hat{\mathbf{W}}^{(t+1)}(i,j) - \hat{\mathbf{W}}^{(t)}(i,j)]\mathbf{P}_j \|_\infty \\ =& \| \sum_{j \in L_i} [\frac{\mathbf{W}^{(t+1)}(i,j)}{\sum_k \mathbf{W}^{(t+1)}(i,k)} - \frac{\mathbf{W}^{(t)}(i,j)}{\sum_k \mathbf{W}^{(t)}(i,k)}]\mathbf{P}_j \|_\infty \\ \leq& \| \sum_{j \in L_i} [\frac{\mathbf{W}^{(t)}(i,j) + \lambda^t}{\sum_k \mathbf{W}^{(t)}(i,k)} - \frac{\mathbf{W}^{(t)}(i,j)}{\sum_k \mathbf{W}^{(t)}(i,k)}]\mathbf{P}_j \|_\infty \\ \leq& \| \sum_{j \in L_i} \frac{\lambda^t}{\sum_k \mathbf{W}^{(t)}(i,k)} \mathbf{P}_j \|_\infty \\ \leq& \| \sum_{j \in L_i} \lambda^t \mathbf{P}_j \|_\infty \\ \leq& \lambda^t K\end{aligned}$$

Similarly, we can prove the second part has the same bound.

$$\begin{aligned}&\| \sum_{j \in \bar{L}_i} [\hat{\mathbf{W}}^{(t+1)}(i,j) - \hat{\mathbf{W}}^{(t)}(i,j)]\mathbf{P}_j \|_\infty \\ =& \| \sum_{j \in \bar{L}_i} [\frac{\mathbf{W}^{(t)}(i,j)}{\sum_k \mathbf{W}^{(t)}(i,k)} - \frac{\mathbf{W}^{(t+1)}(i,j)}{\sum_k \mathbf{W}^{(t+1)}(i,k)}]\mathbf{P}_j \|_\infty \\ =& \| \sum_{j \in \bar{L}_i} [\frac{\mathbf{W}^{(t)}(i,j)}{\sum_k \mathbf{W}^{(t)}(i,k)} \\ &- \frac{\mathbf{W}^{(t)}(i,j) + \lambda^{(t+1)}\cos(\mathbf{x}_i^{(t+1)}, \mathbf{x}_j^{(t+1)})}{\sum_k \mathbf{W}^{(t)}(i,k) + \lambda^{(t+1)}\cos(\mathbf{x}_i^{(t+1)}, \mathbf{x}_k^{(t+1)})}]\mathbf{P}_j \|_\infty \\ \leq& \lambda^t | \frac{\sum_k \cos(\mathbf{x}_i^{(t+1)}, \mathbf{x}_k^{(t+1)}) - \sum_{j \in \bar{L}_i} \cos(\mathbf{x}_i^{(t+1)}, \mathbf{x}_j^{(t+1)})}{\sum_k \mathbf{W}^{(t)}(i,k)} | \\ \leq& \lambda^t K\end{aligned}$$

Thus, we have

$$\Delta(t+1) \leq \lambda^t K$$

Then, we know $\Delta(t+1) \leq \epsilon$ when $t > \lceil \log_\lambda \frac{\epsilon}{K} \rceil$. □

The proof of Theorem 4

Proof. According to the proof of Theorem 1, when $T_e > \lceil -\log_\lambda(K^2|V_i|)\rceil$, we have

$$\begin{aligned}
&\| \mathcal{P}_i^{(\infty)} - \mathcal{P}_i^{(T_e)} \|_\infty \\
\leq &\| \sum_{t=T_e}^{\infty} \| \mathcal{P}_i^{(t+1)} - \mathcal{P}_i^{(t)} \|_\infty \\
\leq &\sum_{t=T_e}^{\infty} \lambda^t K^2(|V_i| + 2) \\
= &\frac{\lambda^{T_e} K^2(|V_i| + 2)}{1 - \lambda}
\end{aligned}$$

So we can select $T_e = \lceil \log_\lambda \frac{\epsilon(1-\lambda)}{K^2(|V_i|+2)} \rceil$ such that when $t > T_e$, $\| \mathcal{P}_i^{(\infty)} - \mathcal{P}_i^{(t)} \|_\infty < \epsilon$.

□