之前EM算法不足

在我之前所采取的EM算法中，图 G 作为模型隐变量参与优化，而我们假设该隐变量由若干参数通过一定模型产生，该过程称为图结构建模；受当时参考文献影响，我采用SBM模型，该模型认为

- 一个图中的节点属于一个簇
- 同簇节点相互关联概率一定且相等
- 不同簇节点之间概率只与双方节点所在的簇有关

通过数学语言描述为：
$xG->P(G|\Omega)=\prod_{i<j}\Omega_{c_ic_j}^{G_{ij}}(1-\Omega_{c_ic_j})^{(1-G_{ij})}$
文献中认为c_i已知，因此给定参数Ω，G的后验分布$P(G|\Omega)$也已知；

这之后我们构建图参数对观测值的后验概率分布 $P(G,\Omega,\alpha,\beta|O)=\frac{P(O|G,\alpha,\beta)P(G|\Omega)P(\Omega)P(\alpha)P(\beta)}{P(O)}$ 认为参数的先验均为几何分布，且$P(G|\Omega)$已知，我们只需要通过EM算法计算$P(O|G,\alpha,\beta)$便可得出该后验分布。

但是，在本项目中，节点所在簇的信息未知，EM算法原理上讲行不通，但是之前版本算法通过假设**每个节点都属于一个簇**来规避该问题，使得在给定参数Ω 的情况下也能计算隐变量G；但是该算法的复杂度也直线上升为$O(n^2)$，相当笨重。这也使我产生了使用HMM解决较高计算复杂度问题的意图，但是中期后我继续查资料的时候没有找到使用该方法计估计邻接矩阵的方法，因此我也就将重点放在了SBM模型上

解决办法

目前关于图神经网络的研究中尚无对该问题的继续改进，因此我将文献检索范围扩大至离散数学与图论，终于发现如下论文

> Mariadassou, Mahendra, Stéphane Robin, and Corinne Vacher. 2010. “Uncovering Latent Structure in Valued Graphs: A Variational Approach.” *The Annals of Applied Statistics* 4 (2): 715–42.

该文献介绍了目前在SBM模型领域常见的解决算法；EM算法的不足是$P(G|\Omega)$无法计算。但是我们可以通过其他方式对该函数进行采样与模拟，也因此派生出两种解决办法

- 基于变分推理的算法
- 基于马尔可夫蒙特卡洛的算法

之前生信课上我已经用过MCMC算法计算系统发生树，因此这次我想试试变分推理法