

Network Supervised Topic Modeling: General Idea

[Mei et al. 08]

- Probabilistic topic modeling as optimization: maximize likelihood

$$\Lambda^* = \arg \max_{\Lambda} p(\text{TextData} | \Lambda)$$

- Main idea: network imposes constraints on model parameters Λ
 - The text at two adjacent nodes of the network tends to cover similar topics
 - Topic distributions are smoothed over adjacent nodes
 - Add network-induced regularizers to the likelihood objective function

Any generative model

Any network

$$\Lambda^* = \arg \max_{\Lambda} f(p(\text{TextData} | \Lambda), r(\Lambda, \text{Network}))$$

Any way to combine

Any regularizer