University of Konstanz
Algorithmics Group
V. Amati / J. Lerner / D. Schoch

Network Modeling
Winter Term 2013/2014

Assignments $\mathcal{N}^{\underline{o}}$ 2

released: 06.11.2013 **due:** 13.11.2013, 12:00

Task 1: Planted Partition Model — 5 points

(1) We take a planted partition model to explain a social network of observed friendship relationships among a group of boys and girls. More precisely, we assume a partition of the actor set into girls and boys $V = V_1 \uplus V_2$, and restrict the intra group parameters to be equal, i.e. $p_{11} = p_{22}$. Consequently, we have only two parameters: one to control for intra group density (p_1) and one to control for inter group density(p_2). Calculate the maximum likelihood estimate of these two parameters, given an observed network G_{obs}.
Explain what these parameters might tell us about *social selection.*

(2) Provide a modification of the efficient $G(n, p)$ algorithm such that we are able to draw loop-free, undirected graphs from the general planted partition model (arbitrary k and p_{ij}). *(It is sufficient to give a textual description. An Implementation is considered in Task 3)*

Task 2: Model with given expected degree sequence — 5 points

Let $d_1, d_2, \ldots, d_n$ be a sequence of real numbers in the intervall $[0, n-1]$ with the following property:

$$d_i^2 \leq \sum_{j=1}^{n} d_j \text{ for all } i \in \{1, n\}.$$

Define a random graph model on the set of undirected graphs *with loops* in which the expected degree of vertex i is equal to d_i. *Hint: It is easiest to do this with a fully independent model.*

Task 3: R: facebook graph and planted partition model 10 points

preparatory steps:

- Download the edgelist and node attributes of the facebook graph from the lecture homepage.
- Create a network object with the edgelist and add the "dorm" attribute from the node attribute table to the vertices.
- Delete all vertices with dorm attribute equal to zero.
- Calculate all p_{ij} values for all pairs of dorms such that the expected number of edges in each dorm and between each pair of dorms is equal to the observed numbers in the facebook graph. *(Hint: Task 1 (1))*
- Extend your $G(n, p)$ implementation, such that you can sample from a planted partition model.

Use the parameters of n and p_{ij} to create 1000 network samples with your planted partition model. Count the number of samples, where more triangles are observed than in the facebook graph. What do you observe this time? Further, calculate the mean number of triangles you observed in your samples. Compare it to the expected number of triangles in the $G(n, p)$ model and explain the difference.

Send your R-Script to `david.schoch@uni-konstanz.de`
A code example with additional hints can be found on the lecture homepage.