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

Network Modeling
Winter Term 2013/2014

Assignments $\mathcal{N}^{\underline{o}}$ 6

released: 04.12.2013 **due:** 10.12.2013,12:00h

Task 1: Hammersley Clifford Theorem 6 points

Let $\mathcal{G}$ the set of undirected, loopless graphs with n vertices and let $c\colon V \to \{A, B\}$ divide the set of vertices $V = \{1, \ldots, N\}$ into two disjoint subsets, $V = A \uplus B$.

Consider the class of random graph models $\mathcal{K}_c = \{(\mathcal{G}, P)\}$ containing all models which fulfill $P(G) > 0$ for all graphs in $\mathcal{G}$ and in which the following independence assumption holds.

> For all pairs of dyads d_1, d_2 it holds that d_1 and d_2 are conditionally independent, unless both of the following properties hold:
>
> - d_1 and d_2 are incident
> - all nodes incident to d_1 and d_2 belong to the same subset. More precisely, if $d_1 = \{u, v\}$ and $d_2 = \{x, y\}$, then
>
> $$c(u) = c(v) = c(x) = c(y) \ .$$

(a) Which random graph models in $\mathcal{K}_c$ are *Markov random graphs*?

(b) Provide an ERGM formula for the probability function of a general random graph model in the class $\mathcal{K}_c$.

(c) Let $V = \{1, 2, 3\} \uplus \{4\}$. Draw the *dependence graph* of $(\mathcal{G}, P) \in \mathcal{K}_c$.

Task 2: Hammersley Clifford Theorem **4 points**

We define a class of (*anti-Markov*) random graph models satisfying

(1) the probability of every graph is positive and

(2) incident dyads $\{i, j\}$ and $\{j, k\}$ are conditionally independent, given the rest of the graph.

Consequently, for every set of four pairwise different vertices $\{i, j, u, v\}$ the dyads $\{i, j\}$ and $\{u, v\}$ might be conditionally dependent, given the rest of the graph.

(a) Describe the cliques of the resulting dependence graph in words.

(b) Provide an ERGM formula for the probability function of a general *homogeneous* anti-Markov graph model.

Task 3: R: Finding a suitable network **5 extra points**

In this Task you should find a network g, with the following properties:

(1) The number of nodes is 100.

(2) The network should be partitioned into 20 groups with 5 vertices each.

(3) This partition should be used as a categorical vertex attribute called "group".

(4) The number of edges and triangles within each group should be higher than between groups.

So we are basically looking for a network of a planted partition model. But here comes the tricky part: the network should fulfill the following properties when fitted.

(1) *ergm(g ∼ edges+triangles)* should be degenerated.

(2) *ergm(g ∼ edges+gwesp(1,fixed=TRUE))* should not be degenerated.

(3) *ergm(g ∼ edges+triangles("group")+nodematch("group"))* should have a lower BIC than *ergm(g ∼ edges+nodematch("group"))*.

The conclusion should be that although g "looks" like a planted partition model, it is better fitted when the local triangles are also included.