Community Detection Algorithms II

Leonid E. Zhukov

National Research University Higher School of Economics

19.03.2012

Community Detection Algorithms

Randomized Min-Cut (Karger, 1993)

Algorithm: Randomized Min-Cut

Data: graph $G(V, E)$

Result: min-cut

repeat

  choose and edge at random;

  contract its endpoints;

until *two vertices remain* ;

Community Detection Algorithms

Let $\min(cut) = k$, then every node degree $k_i \geq k$, edges in graph $m \geq \frac{nk}{2}$
Let E_i event that on i-th step selected edge is *not* in min cut.

$$\begin{aligned} P(E_1) &\geq 1 - \frac{k}{m} = 1 - \frac{2}{n} \\ P(E_2|E_1) &\geq 1 - \frac{2}{n-1} \\ P(E_i|E_1 \cup E_2 \cup ...E_{i-1}) &\geq 1 - \frac{2}{n-i+1} \end{aligned}$$

$$P(E_1 \cup ..E_{n-2}) \geq \prod_{i-1}^{n-2} \frac{n-i-1}{n-i+1} = \frac{2}{n(n-1)}$$

probability of *not* finding min cut after $n^2/2$ independent runs:

$$P(error) \leq \left(1 - \frac{2}{n(n-1)}\right)^{n^2/2} \sim \frac{1}{e}$$

Community Detection Algorithms

Multilevel algorithms(Karypis, 2006)

Algorithm: Multilevel graph partitioning

Data: graph $G(V, E)$

Result: cut

coarsen graph ;

partition;

refine graph ;

---

coarsening is done by randomized maximal matching
partition on coarse graph can be done by spectral, Kernighan-Lin, etc methods