Efficient Construction of 2K+ Graphs

Minas Gjoka†, Balint Tillman‡, Athina Markopoulou†, Rasmus Pagh ‡
† University of California, Irvine, ‡ ITU Copenhagen

Introduction

Goal: Generate simple graphs that have exactly the following properties:

- A target Joint Degree Matrix (JDM):

$$\mathrm{JDM}(k,l) = \sum_{v \in V_k} \sum_{w \in V_l} 1_{\{\{v,w\} \in E\}}$$

- A target degree-dependent average clustering coefficient:

$$\bar{c}(k) = \frac{1}{|V_k|} \sum_{v \in V_k} \frac{T_v}{k \cdot (k-1)/2}$$

(T_v: all triangles using node v; $k \cdot (k-1)/2$: all possible triangles; V_k: all nodes of degree k)

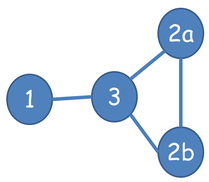
Example

k/l	1	2	3
1			1
2		2	2
3	1	2	

$JDM^{\odot}(k,l)$

k	$\bar{c}$(k)
2	1
3	0.33

$\bar{c}(k)$



Prior Work

- Configuration Model for JDM [Mahadevan '06]
  - graphs may not be simple; low clustering
- Balance Degree Invariant [Stanton '12]
  - constraints during construction; low clustering

Contributions

- New **exact JDM (2K)** construction algorithm
  - provably produces a simple graph
  - less constrained construction
- New **heuristic 2K+** construction algorithm
  - add clustering, while preserving JDM
  - orders of magnitude faster than prior approaches

Exact JDM construction algorithm

JDM (or 2K) Construction

Input: $JDM^{\odot}(k,l)$

for $(k,l) \in JDM^{\odot}(k,l)$
  while $JDM(k,l) < JDM^{\odot}(k,l)$
    Pick disconnected nodes $v \in V_k$ and $w \in V_l$
    if v does not have free stubs
      neighbor switch for v (preserve 2K + free a stub)
    if w does not have free stubs
      neighbor switch for w (preserve 2K + free a stub)
    add edge between (v, w)
    $JDM(k,l)$ ++

Output: Simple graph with exactly $JDM^{\odot}(k,l)$

Example

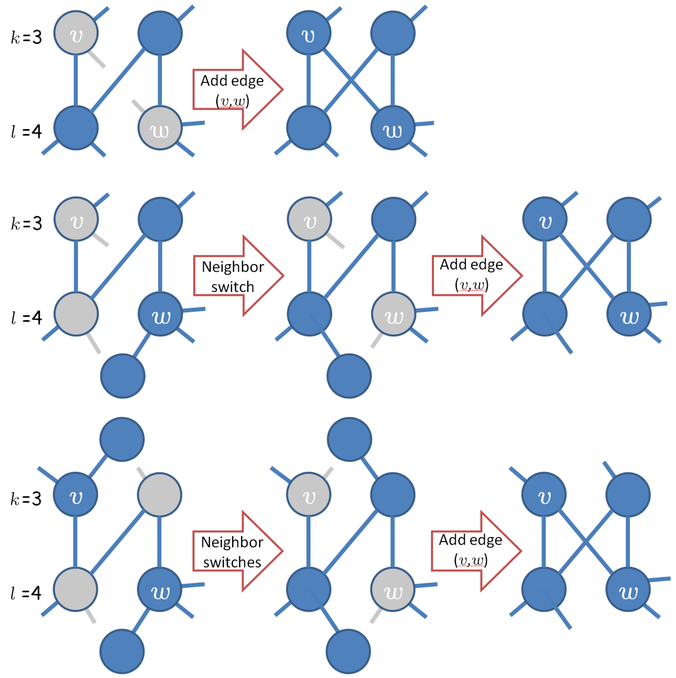
$JDM(3,4) < JDM^{\odot}(3,4)$

Used stub
Free stub

Case 1: $\mathfrak{u}$,$\mathfrak{w}$ both have free stubs

Case 2: $\mathfrak{u}$ has free stubs, but $\mathfrak{w}$ does not

Case 3: neither $\mathfrak{u}$ nor $\mathfrak{w}$ have free stubs



Summary

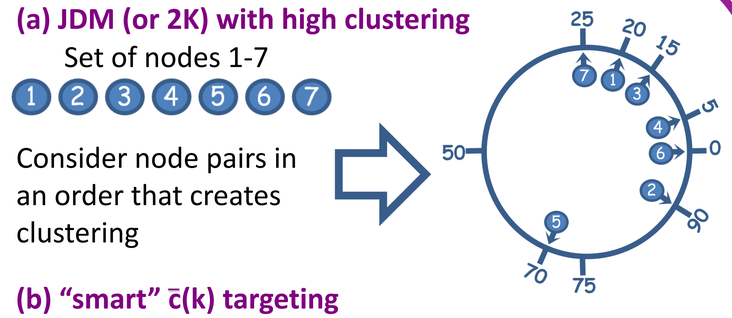
- Running time O(|E|k_{max})
  - at each iteration we add one edge
  - a neighbor switch for a node in degree group k is O(k)
- Flexibility during construction
  - any order of degree pairs (k,l) or node pairs ($\mathfrak{v}$,$\mathfrak{w}$)
  - can start with empty or partially built graph

Exact JDM + approximate c̄(k)

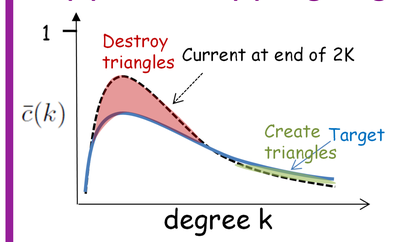
Construction

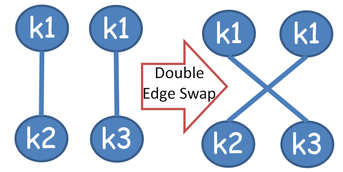
(a) JDM (or 2K) with high clustering

Set of nodes 1-7



(b) "smart" c̄(k) targeting





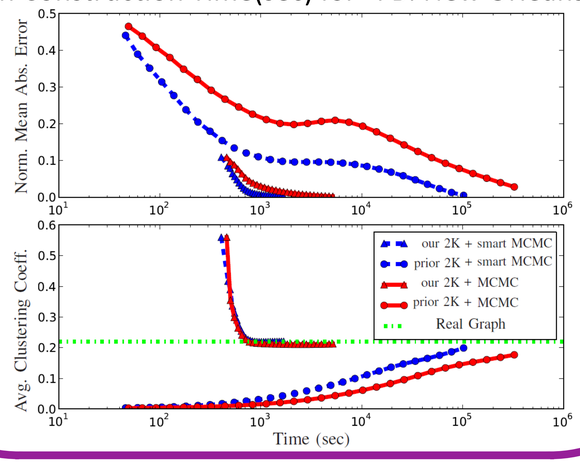
Edge selection for double edge swaps

- triangle creation: low number of shared partners
- triangle destruction: random edges

Simulation Results

Dataset	$\|V\|$	$\|E\|$	$\bar{c}$	our 2K + smart MCMC	prior 2K + MCMC
FB: UCSD	14 948	443 221	0.227	568 s	177 533 s
FB: Harvard	15 126	824 617	0.212	1 182 s	387 506 s
FB: New Orl.	63 392	816 884	0.222	1 463 s	381 397 s
soc-Epinions	75 877	405 737	0.138	888 s	8 958 s
email-Enron	36 692	183 831	0.497	4 279 s	196 202 s
CAIDA AS	26 475	53 377	0.208	121 s	168 s

Graph Construction Time(sec) for "FB: New Orleans"



This work was supported by NSF Award 1028394 and by AFOSR Award FA9550-09-0643.