INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY

HYDERABAD

BRICS – Efficient Techniques for Estimating the Farness-Centrality in Parallel

Sai Charan Regunta, Sai Harsh Tondomker and Kishore Kothapalli


INTERNATIONAL INSTITUTE OF
INFORMATION TECHNOLOGY
HYDERABAD

Why Graphs?

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Farness Centrality

For a node v in graph G(u,v), it is the sum

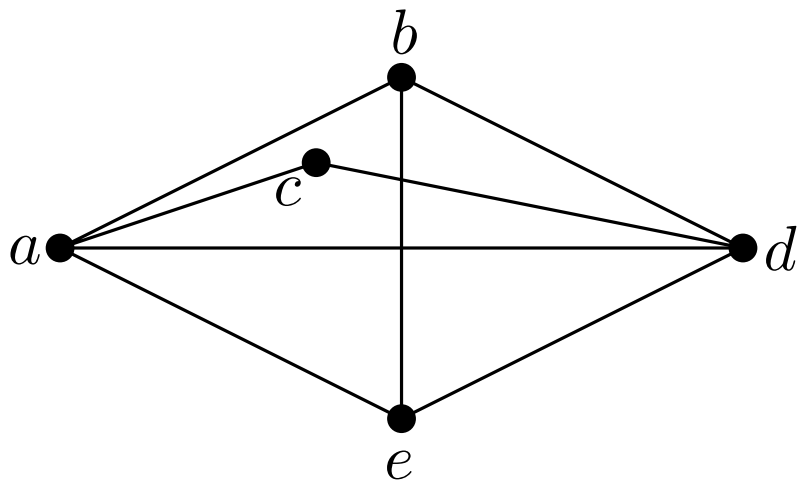
of the shortest distances from node v to all other nodes in the graph.

$$farness[u] = \sum_{v \in V} d(u, v)$$

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Farness Centrality

Example

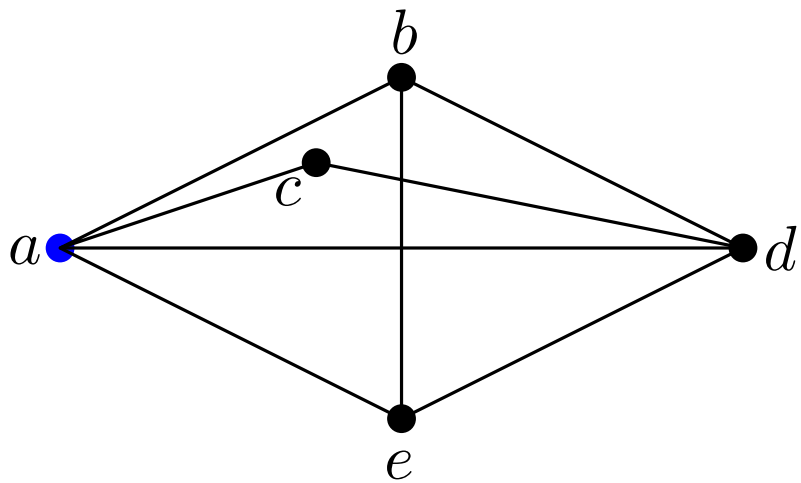


	a	b	c	d	e	FC
a						
b						
c						
d						
e						

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Farness Centrality

Example

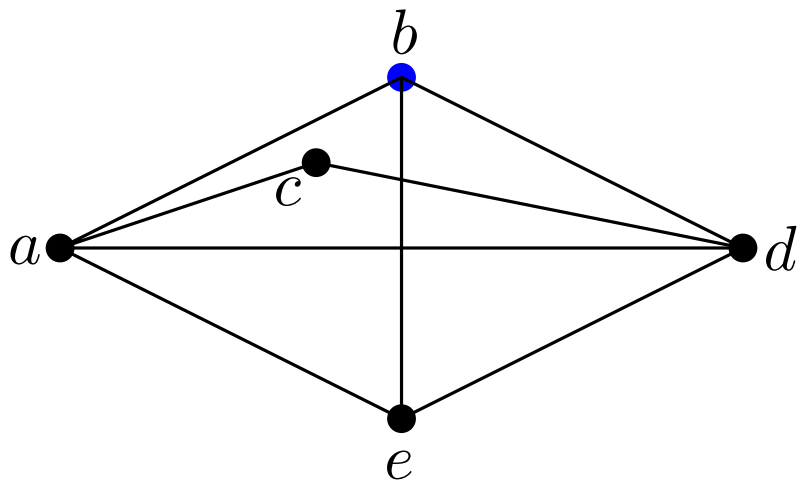


	a	b	c	d	e	FC
a	0	1	1	1	1	**4**
b						
c						
d						
e						

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Farness Centrality

Example

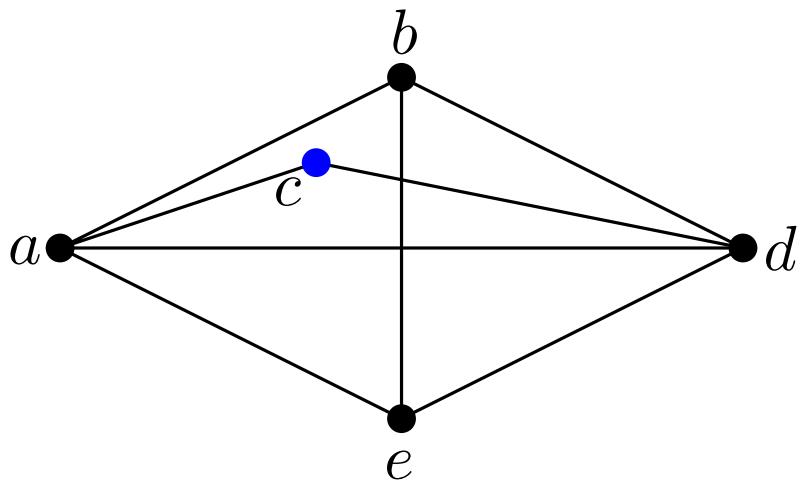


	a	b	c	d	e	FC
a	0	1	1	1	1	**4**
b	1	0	2	1	1	**5**
c						
d						
e						

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Farness Centrality

Example

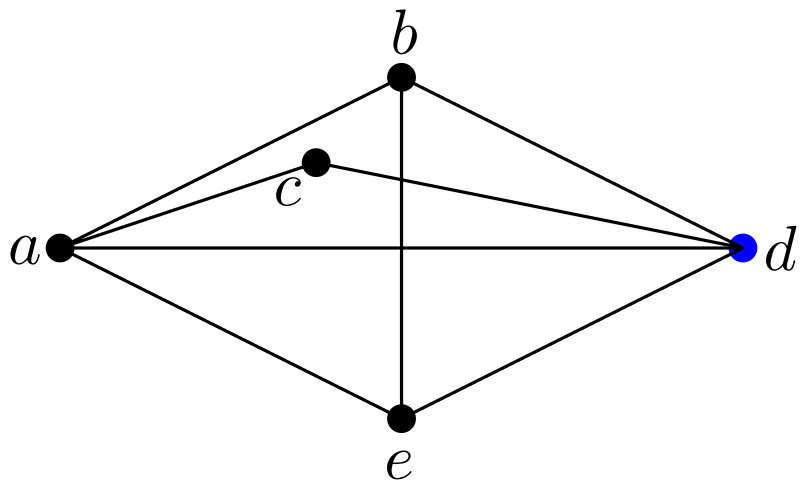


	a	b	c	d	e	FC
a	0	1	1	1	1	**4**
b	1	0	2	1	1	**5**
c	1	2	0	1	2	**6**
d						
e						

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Farness Centrality

Example

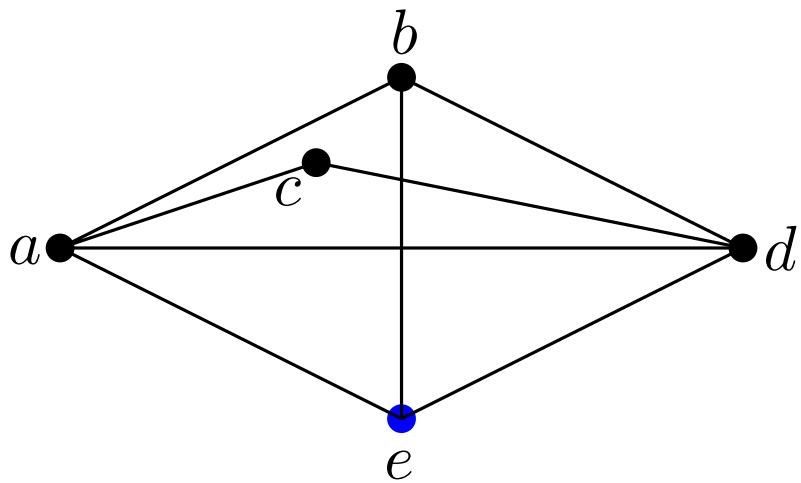


	a	b	c	d	e	FC
a	0	1	1	1	1	**4**
b	1	0	2	1	1	**5**
c	1	2	0	1	2	**6**
d	1	1	1	0	1	**4**
e						

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Farness Centrality

Example

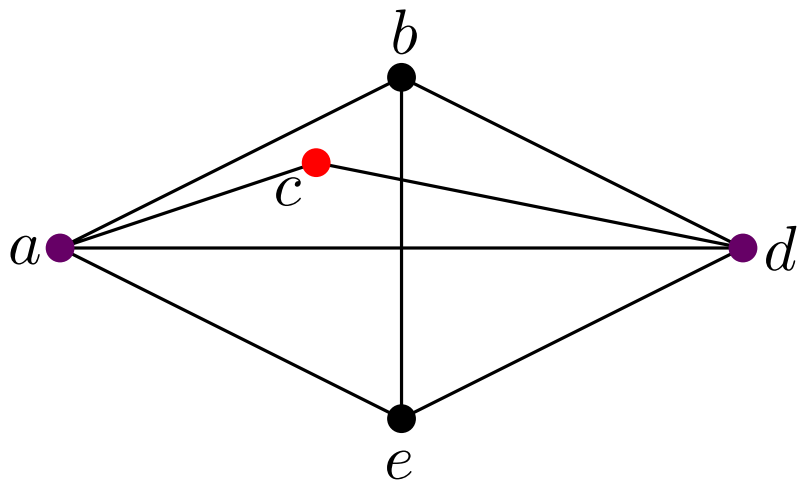


	a	b	c	d	e	FC
a	0	1	1	1	1	**4**
b	1	0	2	1	1	**5**
c	1	2	0	1	2	**6**
d	1	1	1	0	1	**4**
e	1	1	2	1	0	**5**

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Farness Centrality

Example



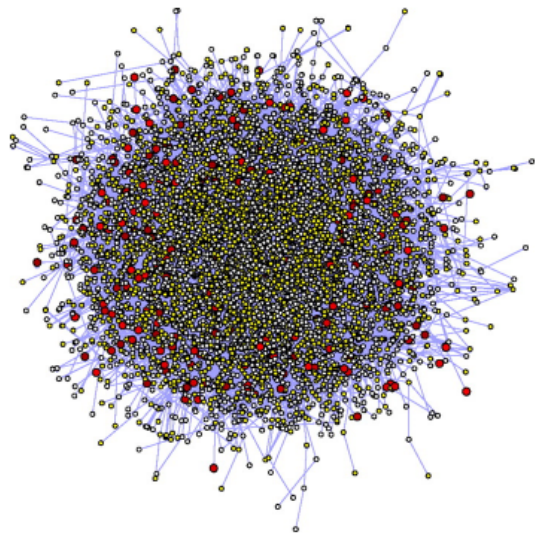
	a	b	c	d	e	FC
a	0	1	1	1	1	**4**
b	1	0	2	1	1	**5**
c	1	2	0	1	2	**6**
d	1	1	1	0	1	**4**
e	1	1	2	1	0	**5**

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Why Approximate FC?

- Time complexity:
  - FC: $O(nm)$
  - Approx FC: $O(km)$

10M nodes graph



  - FC: 2 - 3 days
  - Approx FC: 3 - 5 Hrs

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Related Work

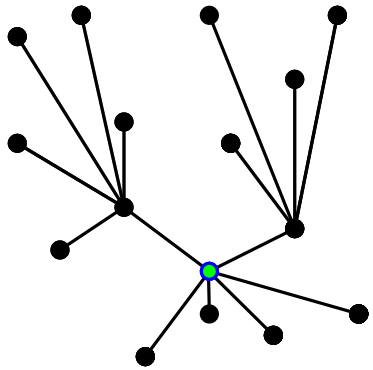
- Eppstein et al: Random-K, first randomized Approximation of Centrality.(1999)

- Cohen et al: Hybrid method of sampling and pivoting to estimate the closeness centrality.

- Sariyuce et al: BADIOS framework for computing exact betweenness centrality.

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

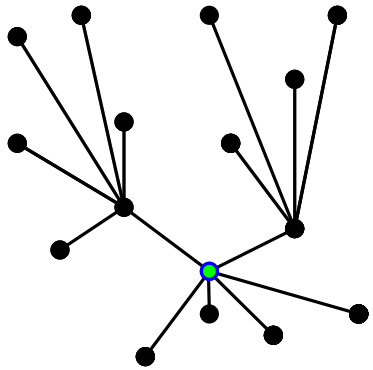
Approaches



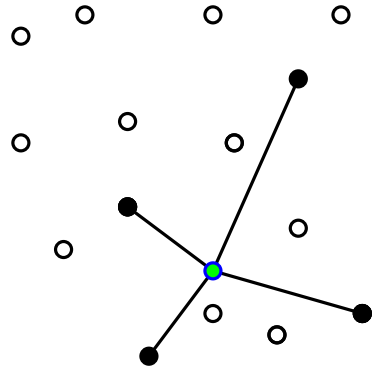
Exact FC

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Approaches



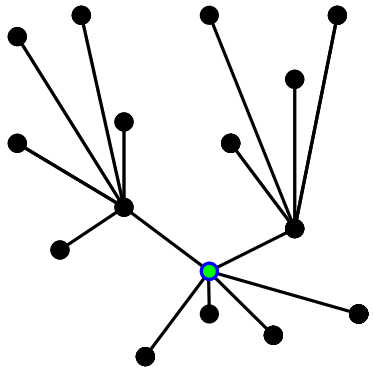
Exact FC



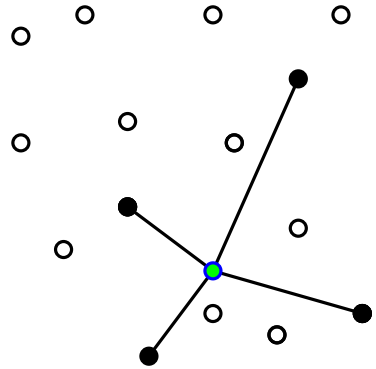
Approx FC(Random-K)

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

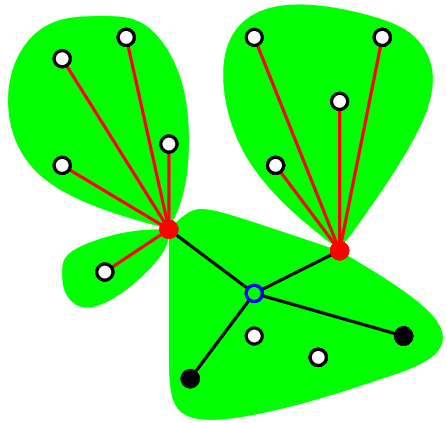
Approaches



Exact FC



Approx FC(Random-K)



Our Approach(BRICS)

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Our Techniques

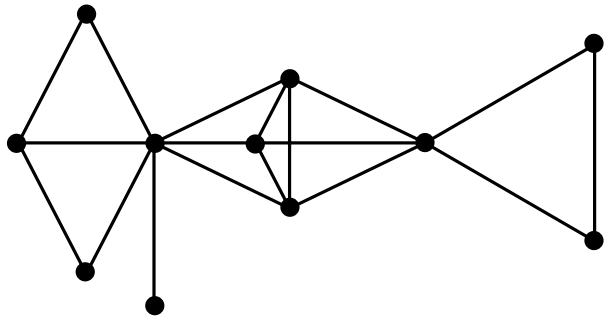
BRICS-FrameWork

- BCT
- Redundant Nodes 3D & 4D
- Identical Nodes
- Chaining

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

BCC based approach

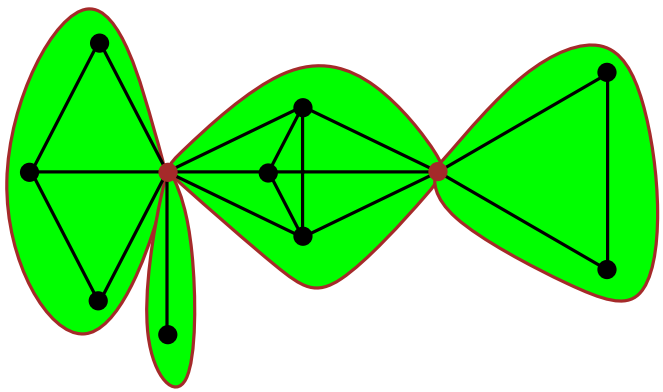
Graph G



INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

BCC based approach

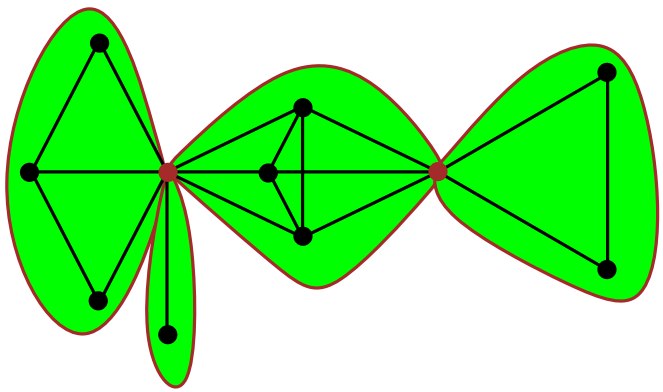
BCC(G)



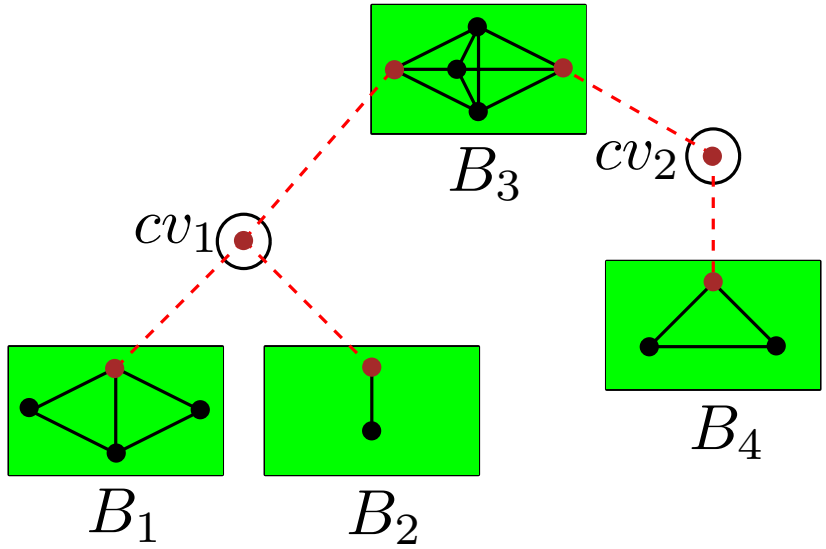
INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

BCC based approach

BCC(G)



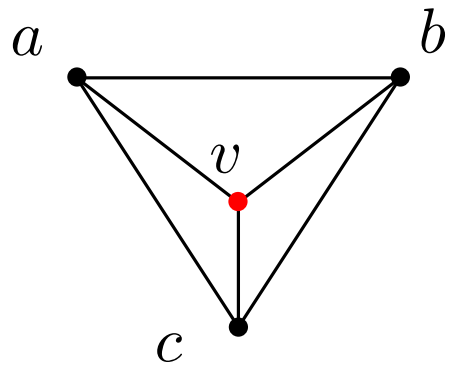
BCT(G)

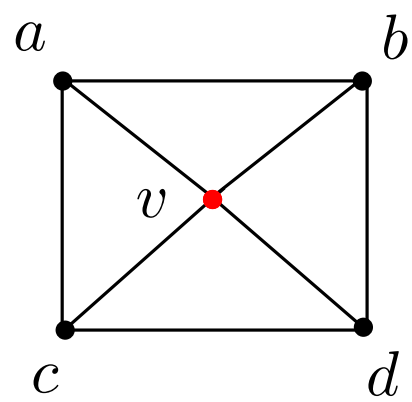


INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Our Techniques

Redundant Nodes



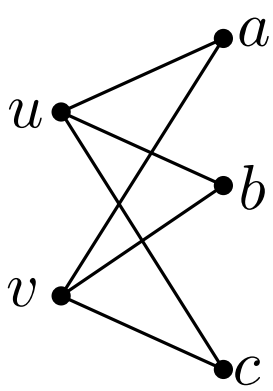


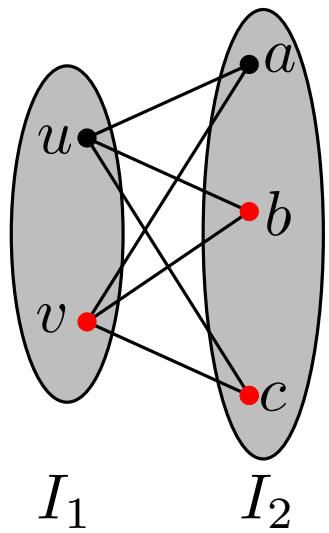
No shortest path pass through v, unless v is source/destination in R3, similarly in R4.

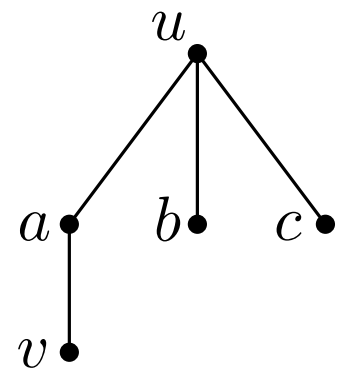
INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Our Techniques

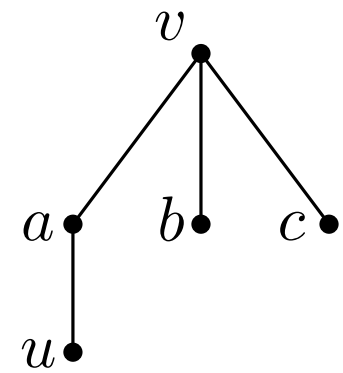
Identical Nodes







BFS(u)

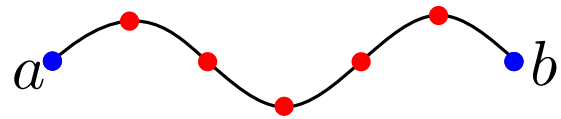


BFS(v)

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Our Techniques

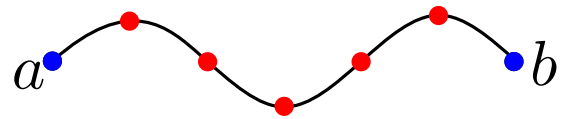
Chain Nodes

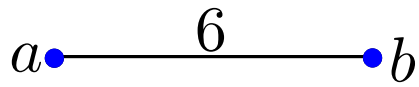


INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Our Techniques

Chain Nodes

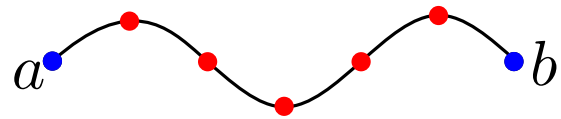




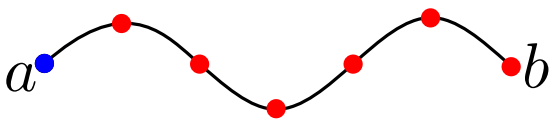
INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Our Techniques

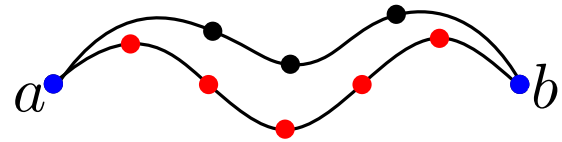
Chain Nodes



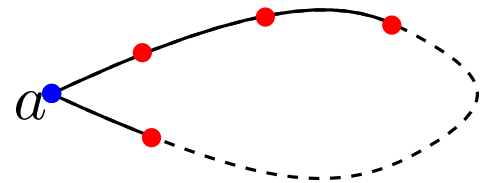
Type-1



Type-2



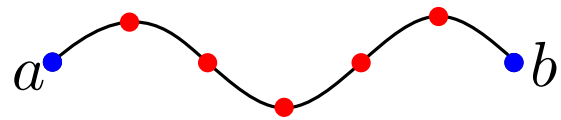
Type-3



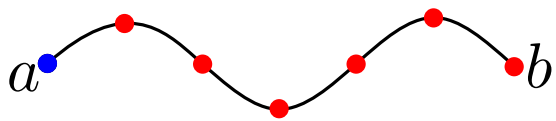
INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Our Techniques

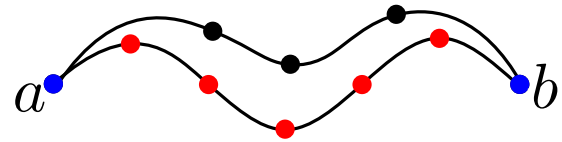
Chain Nodes



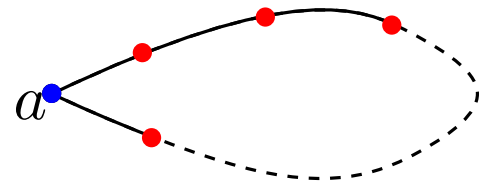
Type-1



Type-2



Type-3

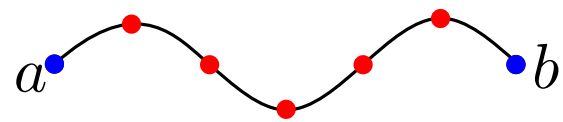


All the redundant nodes can be removed directly from Graph G.

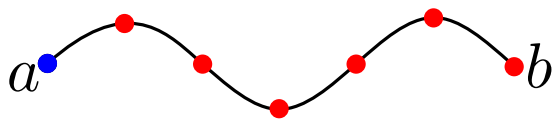
INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Our Techniques

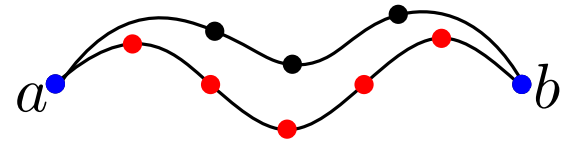
Chain Nodes



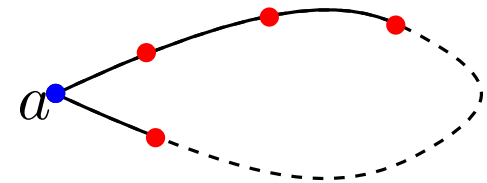
Type-1



Type-2



Type-3



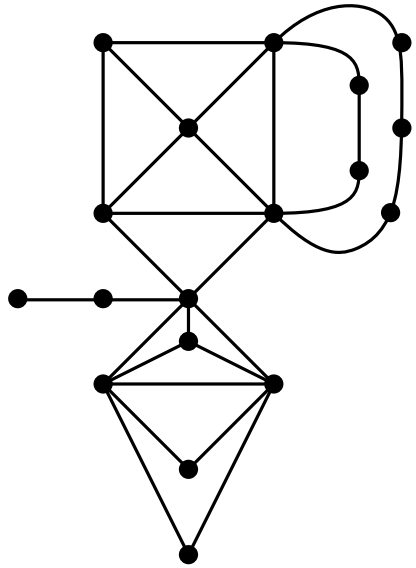
All the redundant nodes can be removed directly from Graph G.

Upto 60% of total chains are of Type-2 and 80% are redundant nodes.

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Putting Together

Procedure

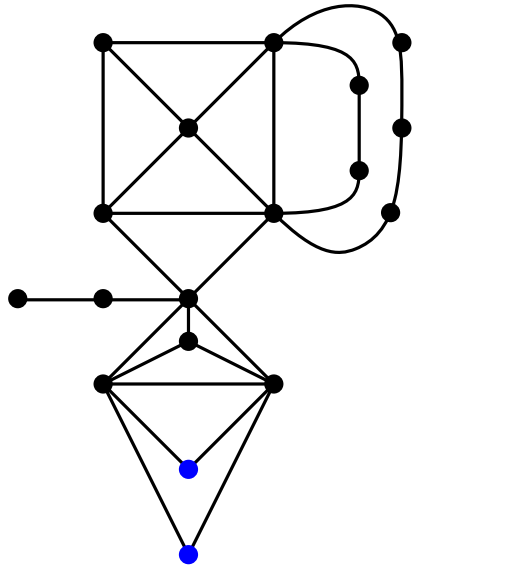


G

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Putting Together

Procedure

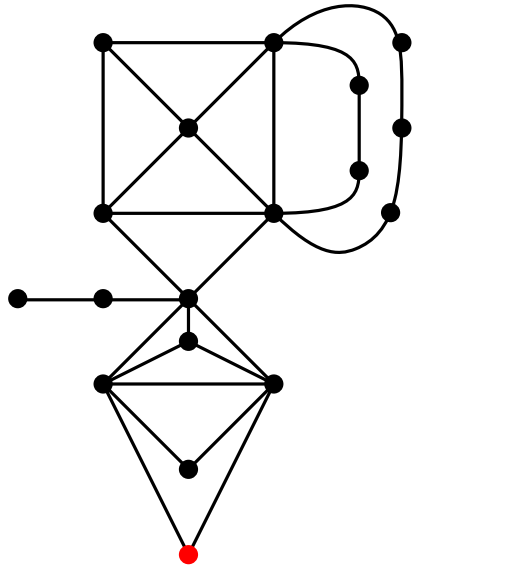


G $\longrightarrow$

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Putting Together

Procedure

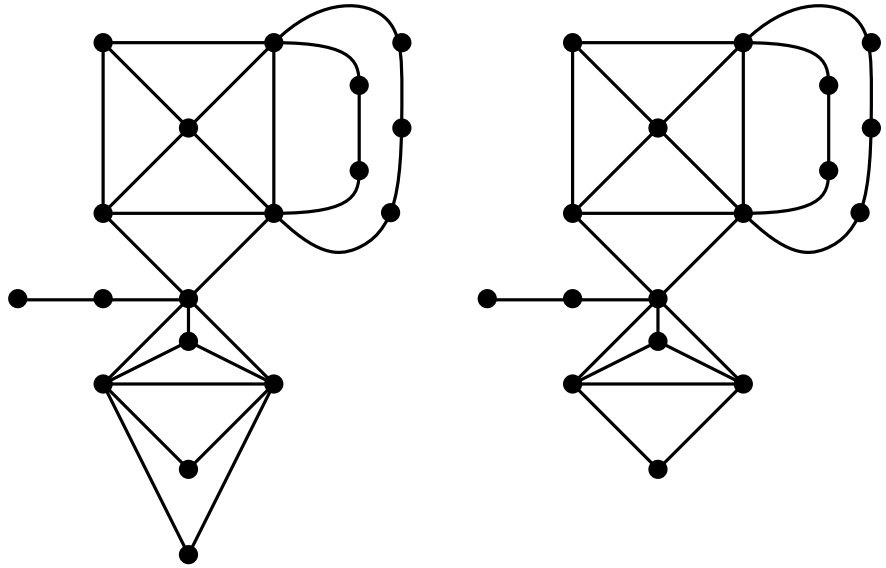


$G \xrightarrow{-I}$

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Putting Together

Procedure

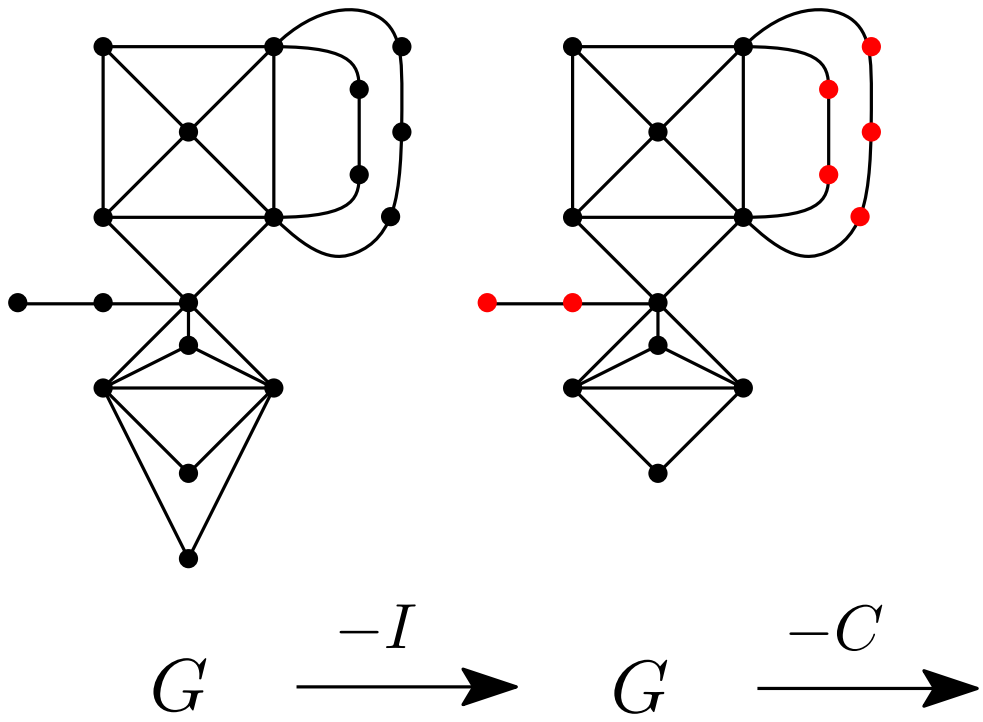


$$G \xrightarrow{-I} G$$

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Putting Together

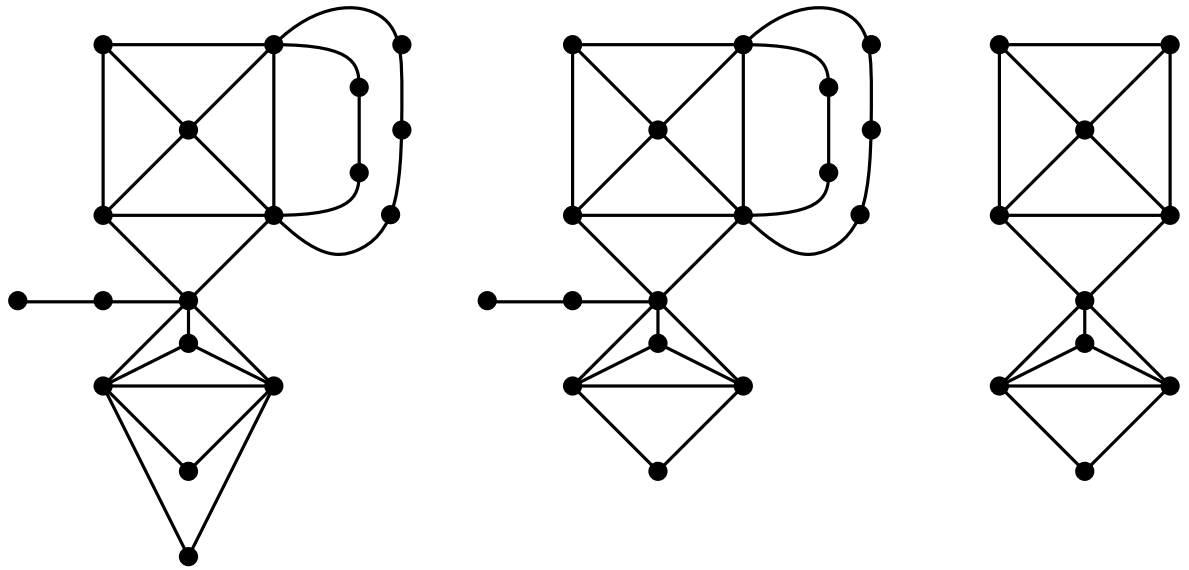
Procedure



INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Putting Together

Procedure

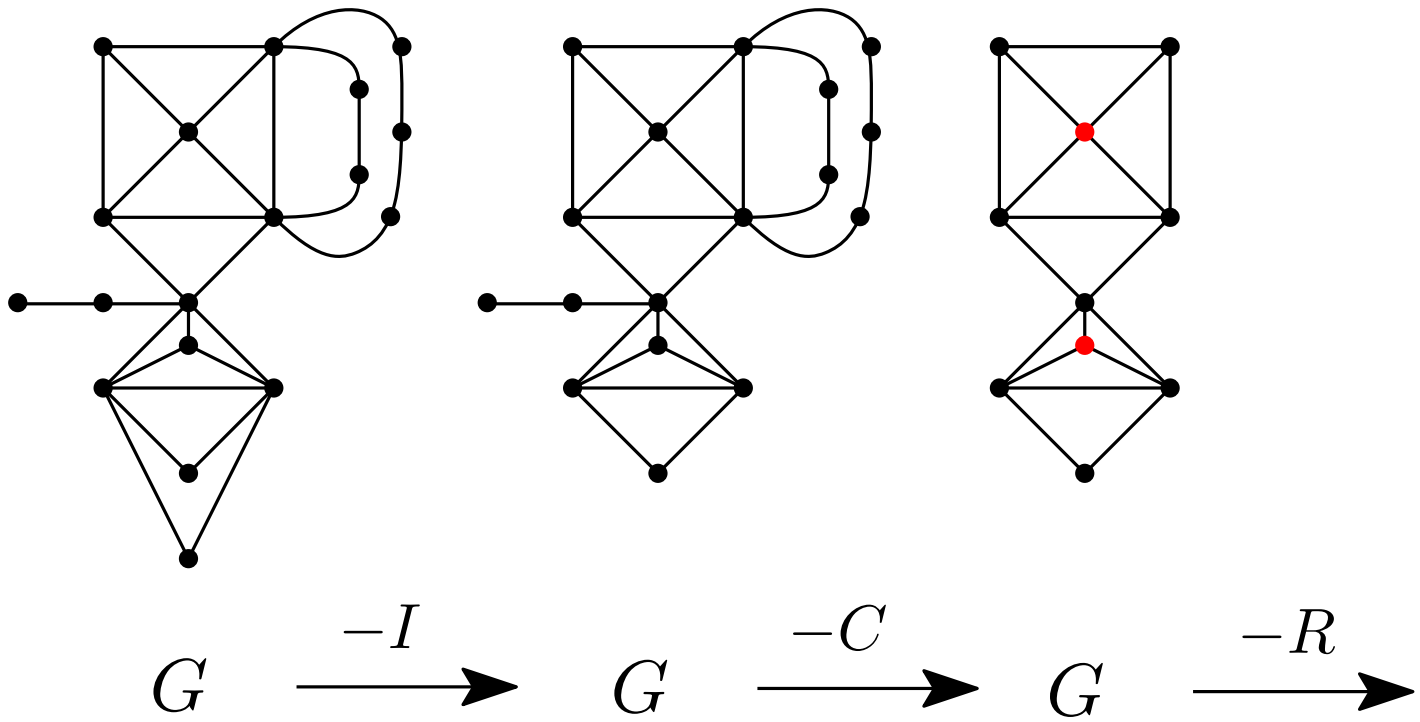


$$G \xrightarrow{-I} G \xrightarrow{-C} G$$

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Putting Together

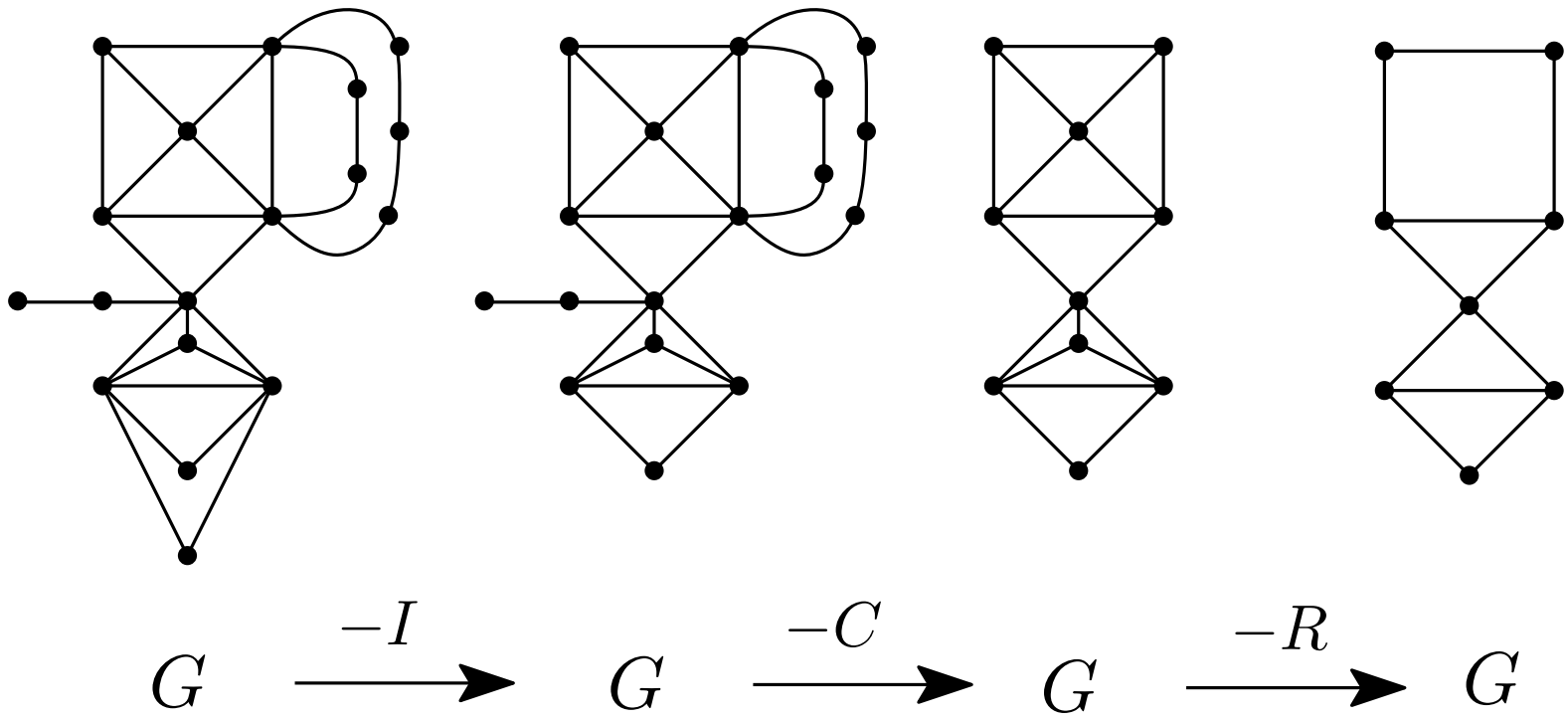
Procedure



INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Putting Together

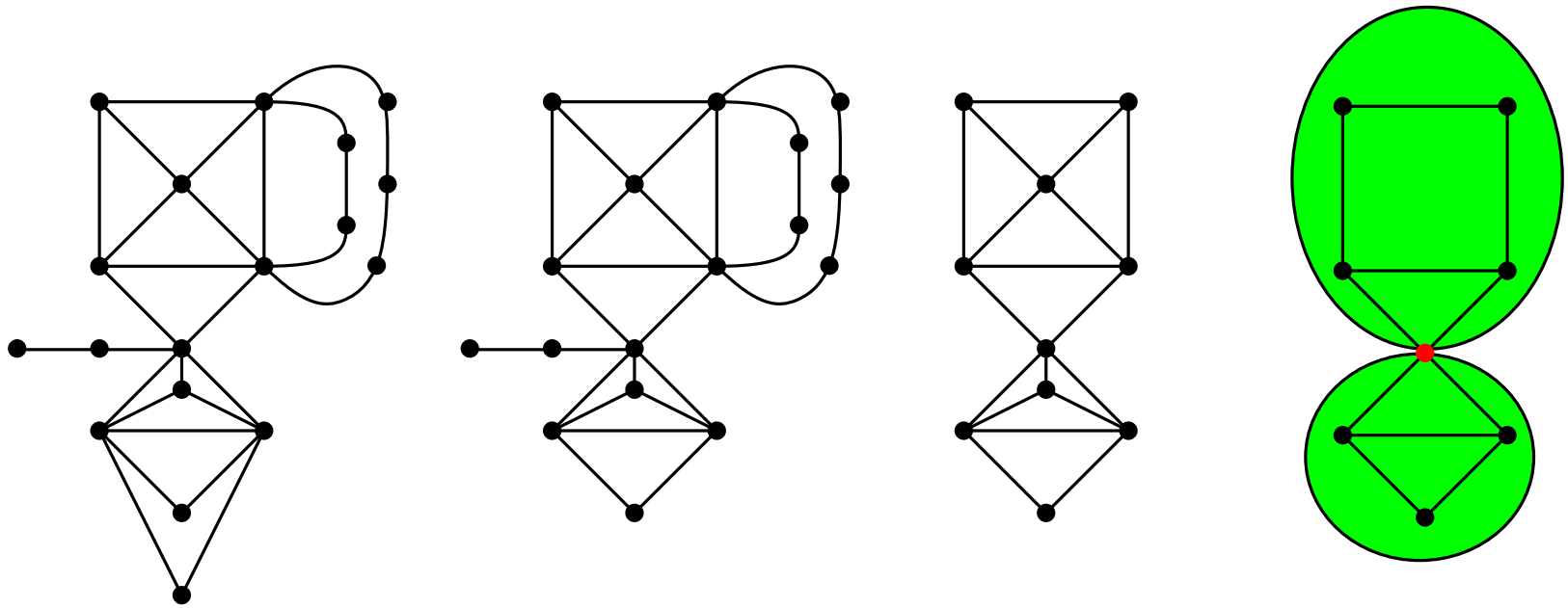
Procedure



INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Putting Together

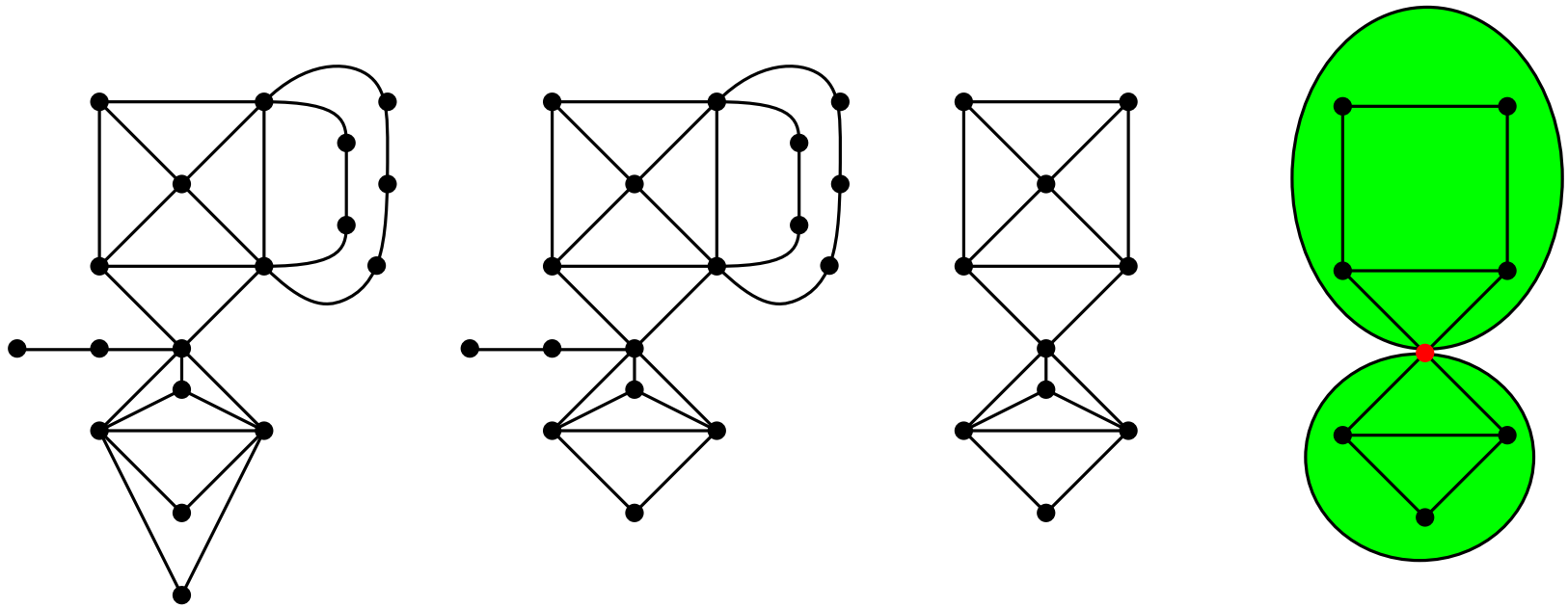
Procedure



INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Putting Together

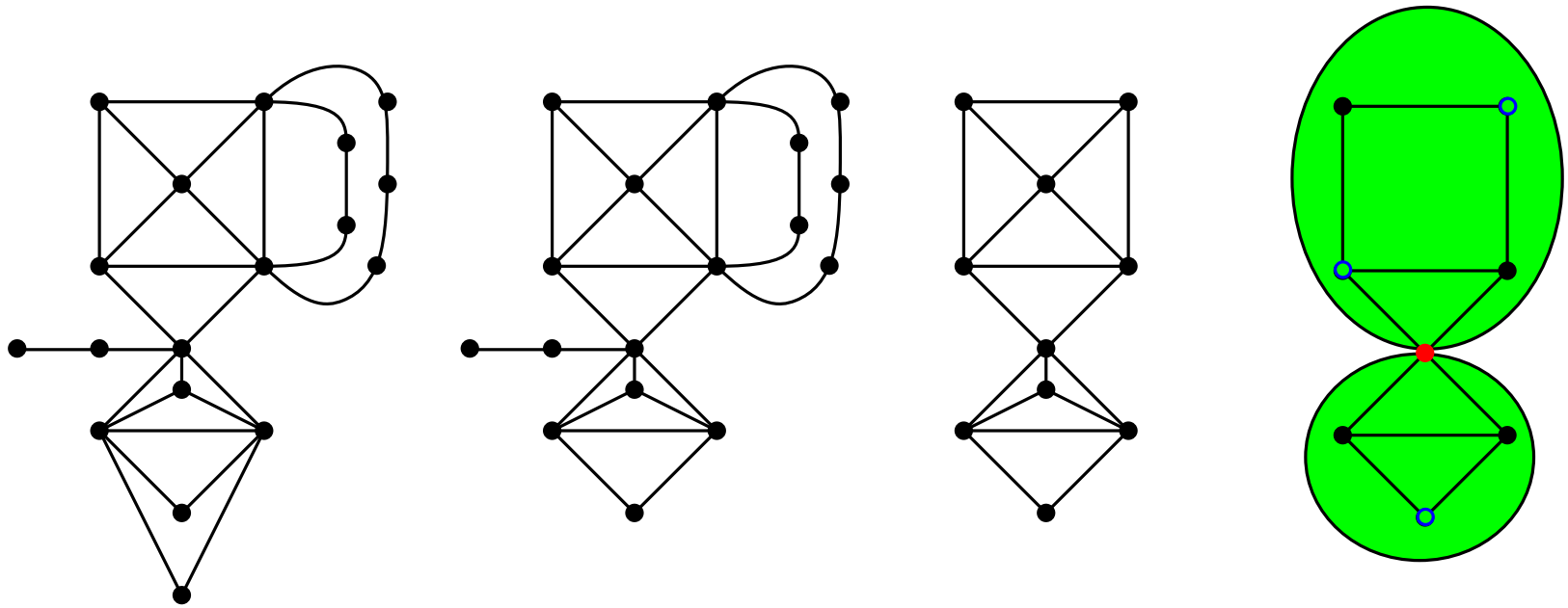
Procedure



INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Putting Together

Procedure

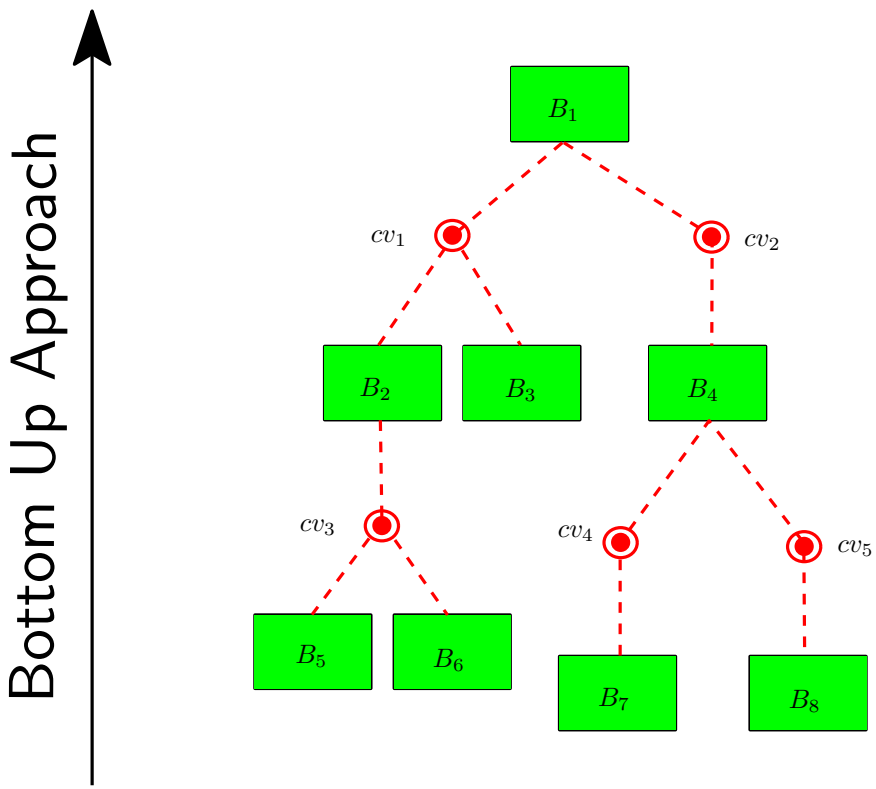


$$G \xrightarrow{-I} G \xrightarrow{-C} G \xrightarrow{-R} G$$

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Putting Together

Procedure

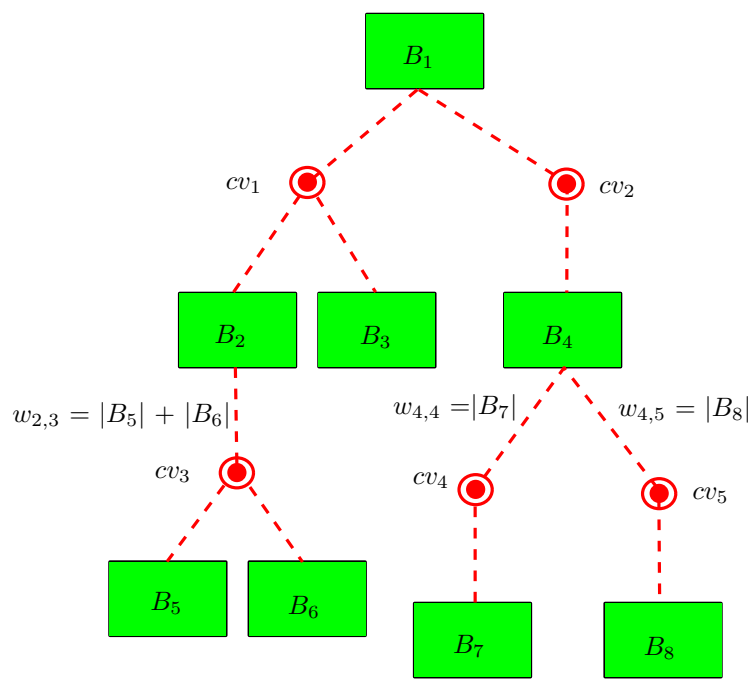


INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Putting Together

Procedure

Bottom Up Approach



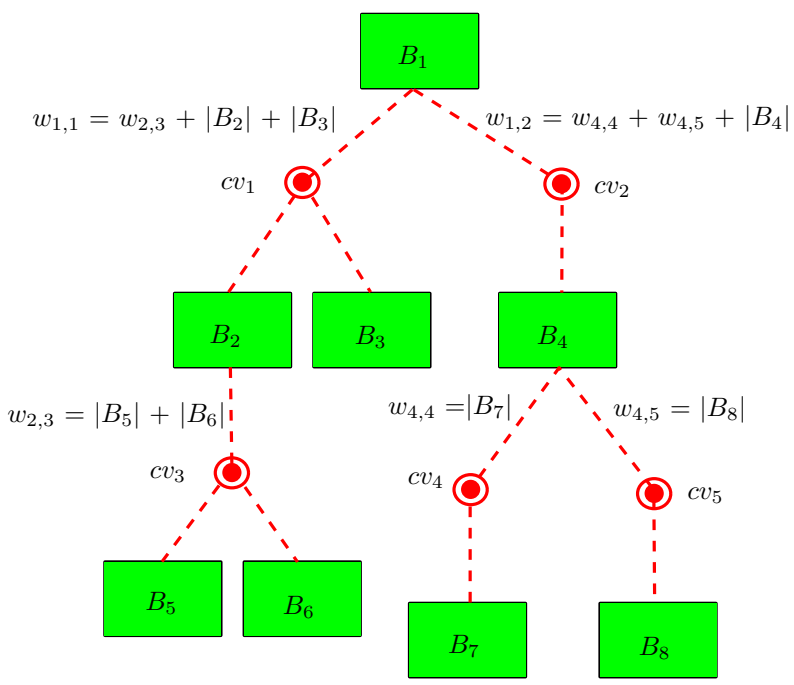
Top Down Approach

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Putting Together

Procedure

Bottom Up Approach

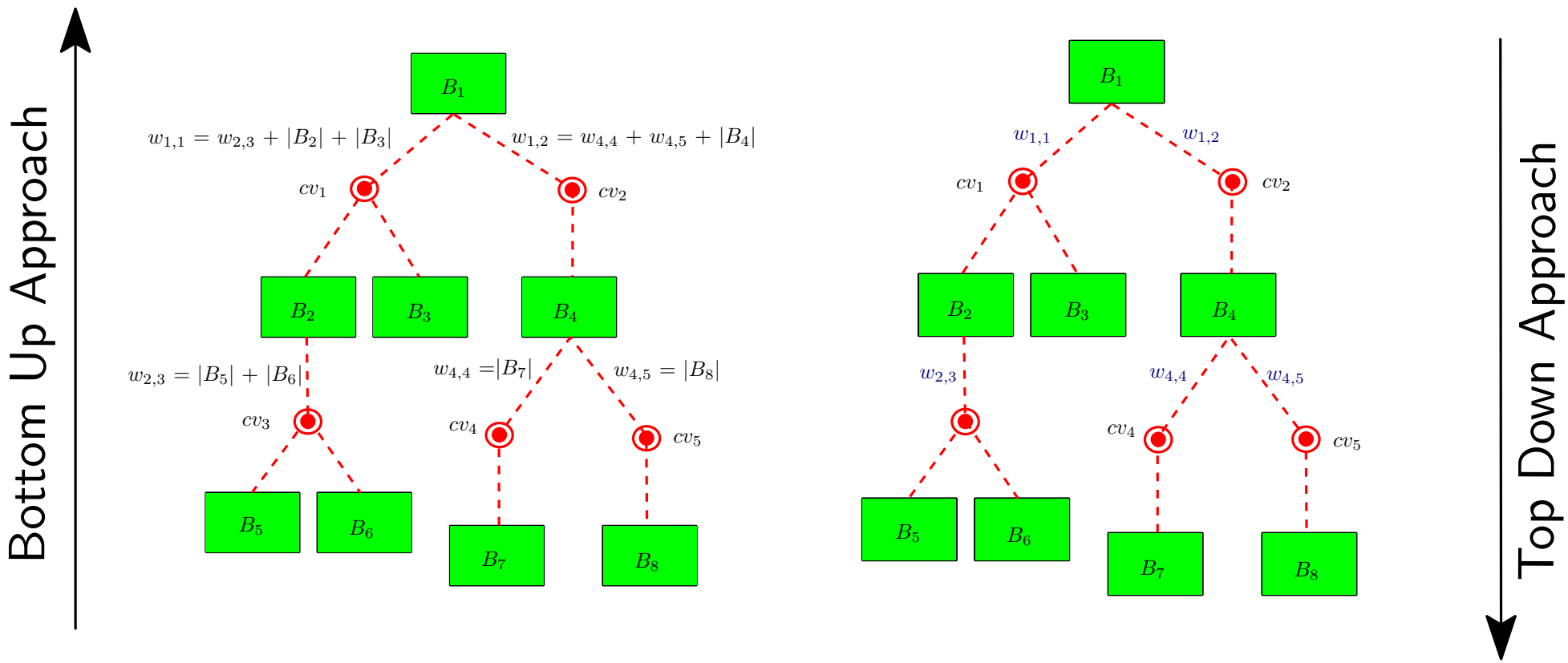


Top Down Approach

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Putting Together

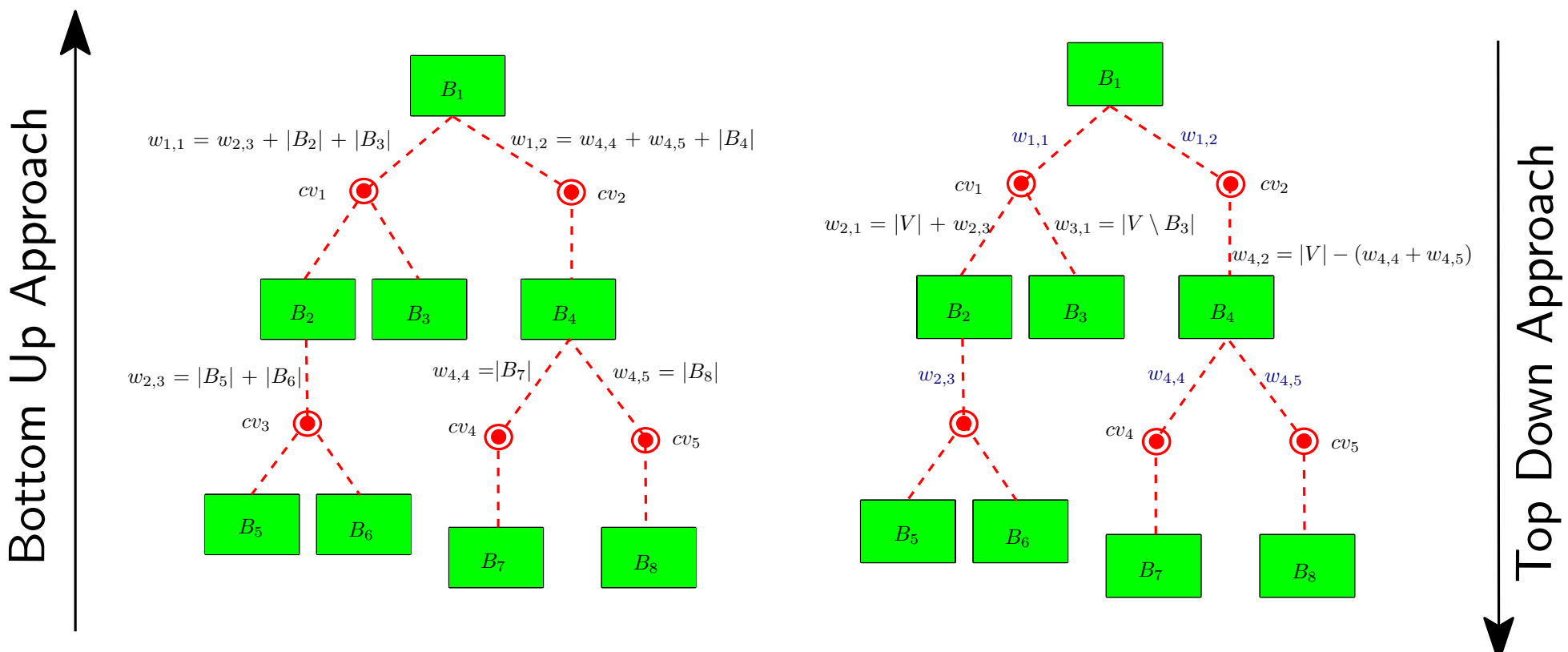
Procedure



INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Putting Together

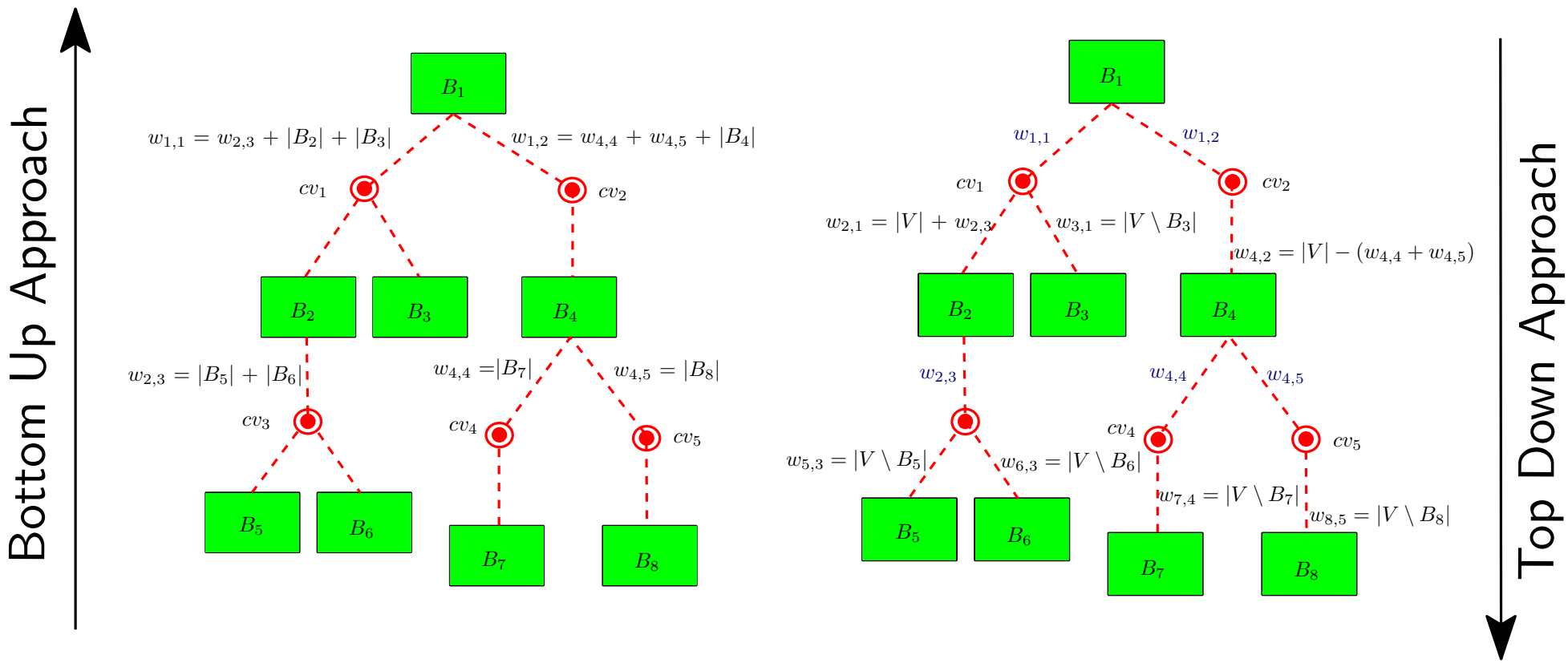
Procedure



INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Putting Together

Procedure

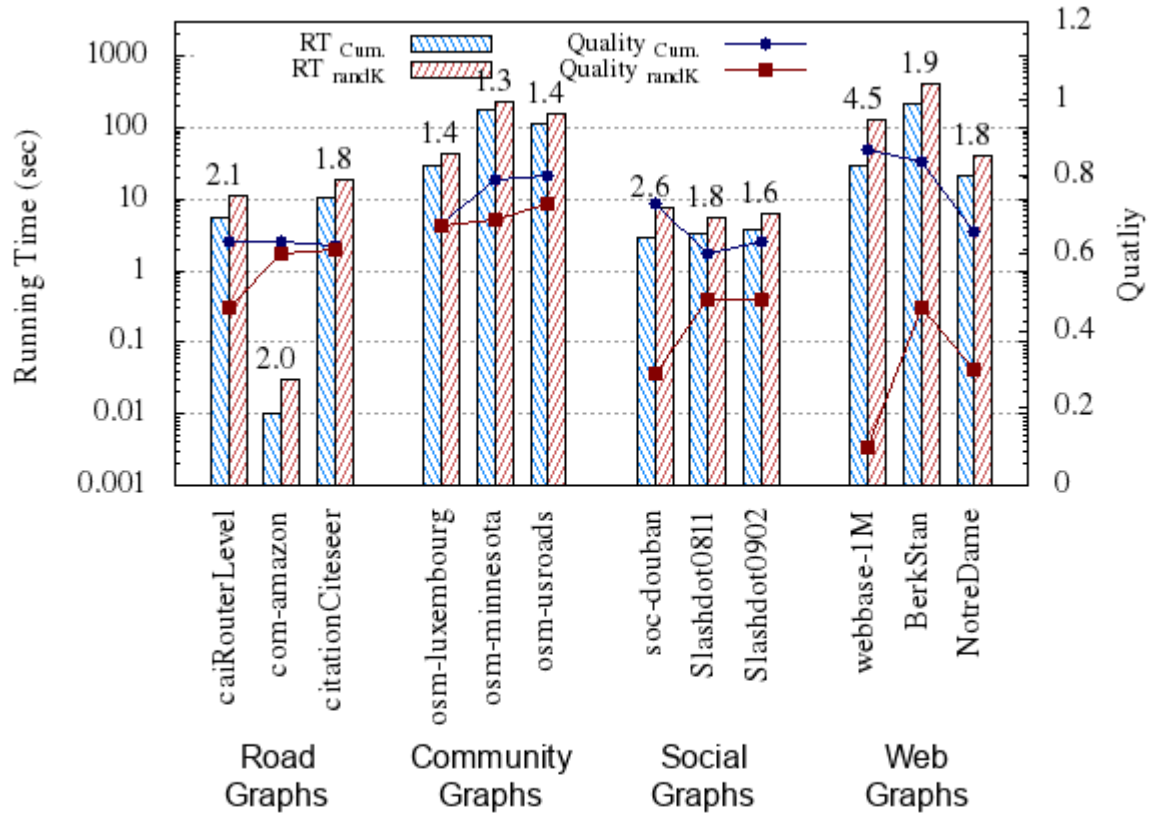


INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Results

Speed up Vs Quality

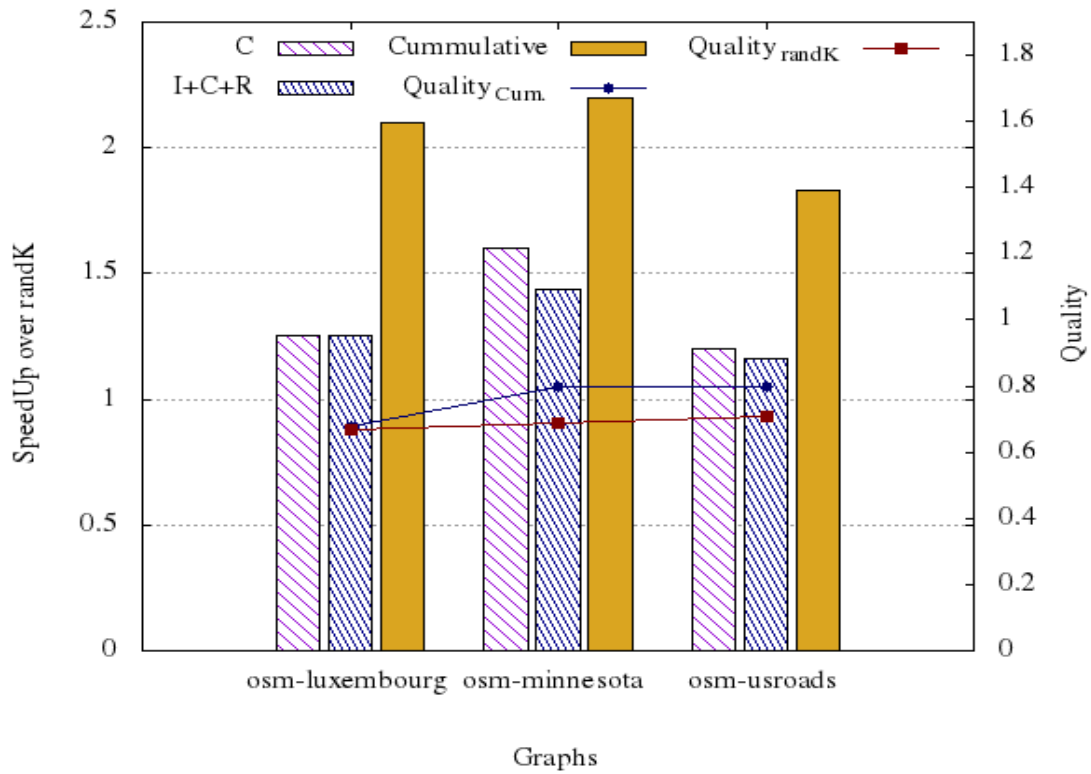
$$\text{Quality} = \frac{\sum_{v \in V} AR(v)}{n} \text{ where } AR(v) = \frac{farness\_estimated(v)}{farness\_actual(v)}$$



INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Results

Profiling for Road Networks



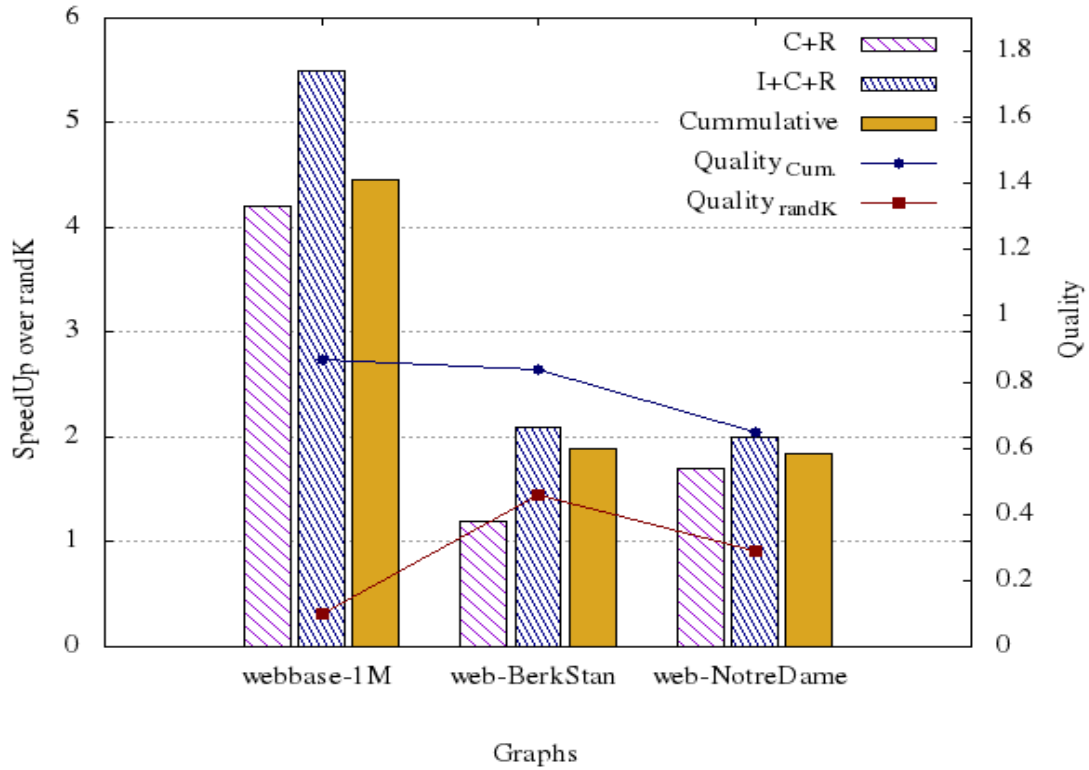
I & R $<$ 1% C: 70 - 85% and Largest BiCC 90%

Distribution of BiCC: Skewed, Slightly better quality

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Results

Profiling for Web graphs



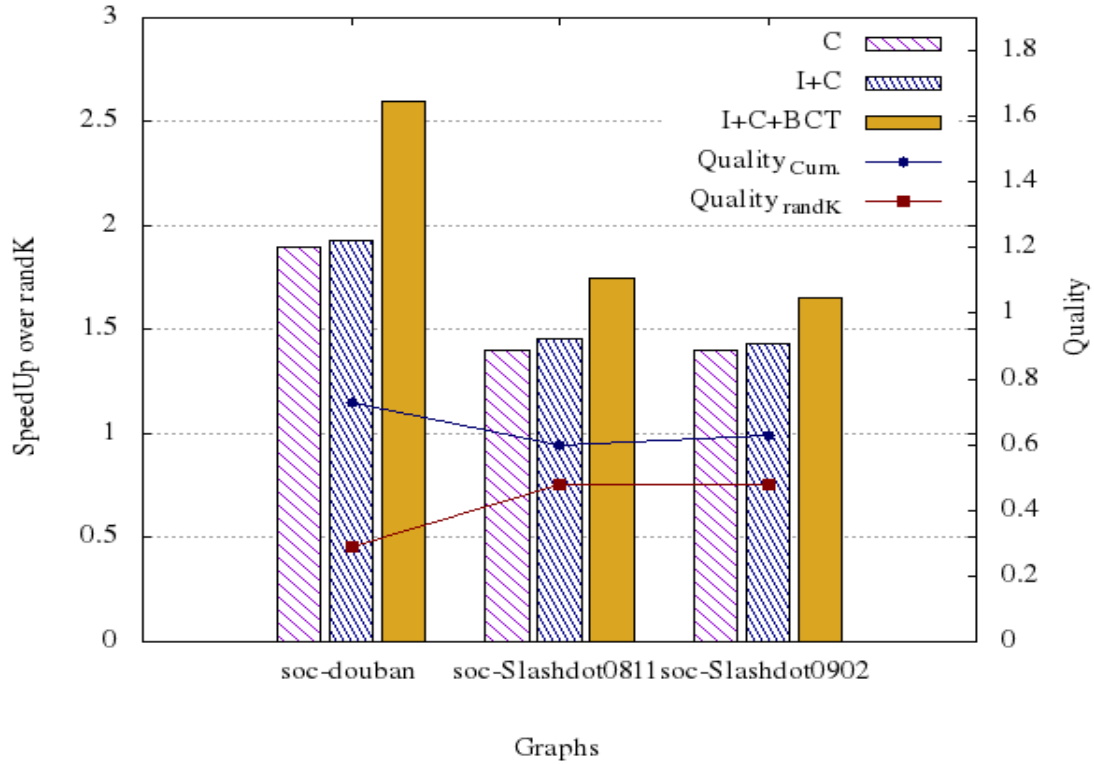
I: 44 %, C: 54%, R: 2.4 % and Large BiCC 27.3 % Nodes

Distribution of BiCC: nearly balanced, better quality

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Results

Profiling for Social graphs



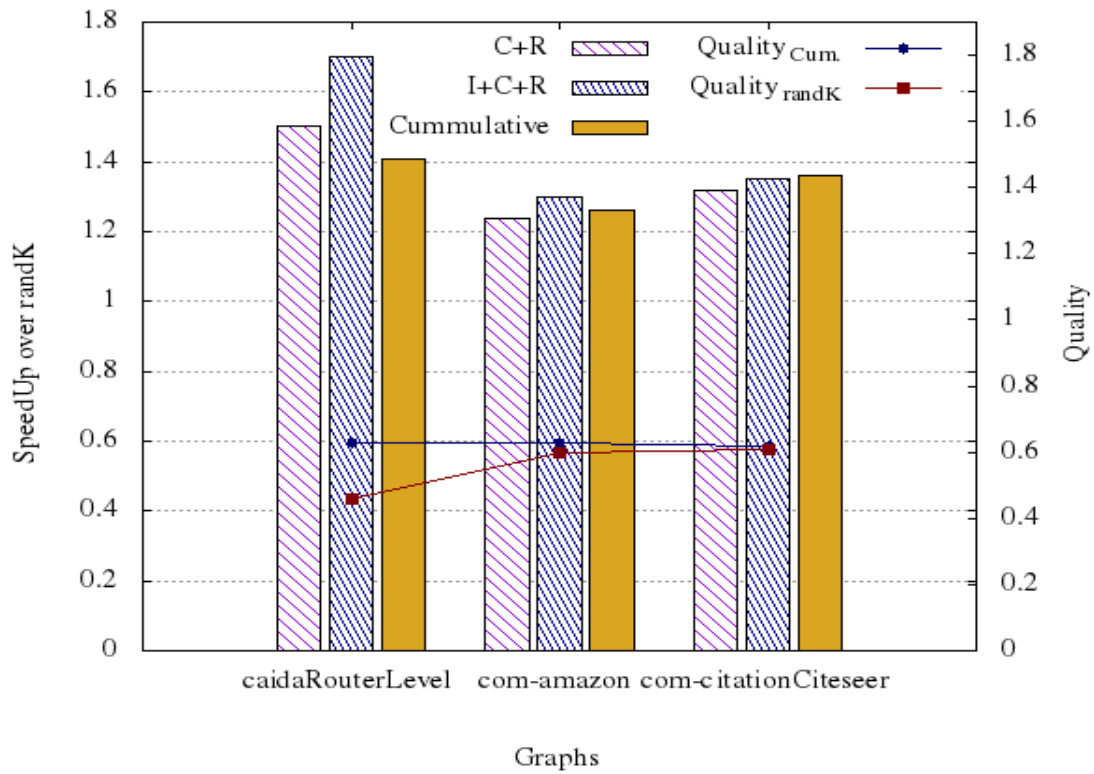
I & C $\leq$ 40 % but R $<$ 1% and Large BiCC: 72 %

Distribution of BiCC: Skewed, Slightly better quality

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Results

Profiling for Community graphs



I, C & R $\leq$ 40 % and Large BiCC: 80 %

Distribution of BiCC: Skewed, Slightly better quality

Thank You

Questions???