Fast Centrality-Driven Diffusion in Dynamic Networks

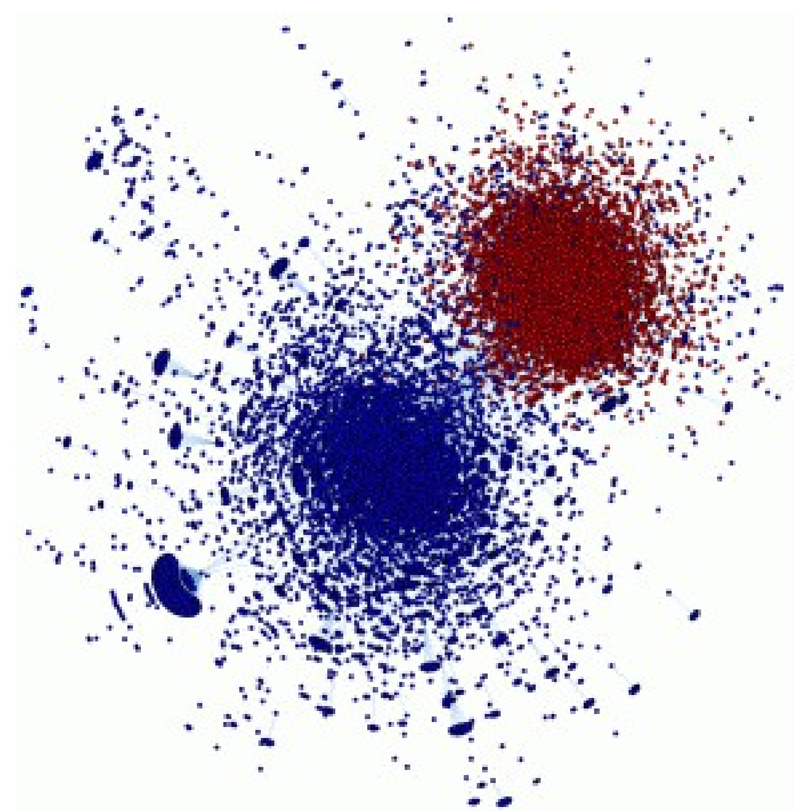
Abraão Guimarães (UFJF, Brazil)
Alex B. Vieira (UFJF, Brazil)
Ana Paula Couto (UFMG, Brazil)
Artur Ziviani (LNCC, Brazil)

(Simplex – WWW/2013)

WDN 2013 – Buenos Aires, Argentina, November 2013

Motivation

- Diffusion processes in complex networks
  - data search
  - data routing
  - information spreading
  - ...



How to speed up the diffusion process?

Centrality-driven diffusion

biased forwarding isn't new,
but how different centralities may help?

Centrality

- **Centrality** notion reflects the relative importance of nodes (links) in the network
- Examples of node centrality
  - degree centrality
  - betweenness centrality
    - relative number of shortest paths passing through each node
  - closeness centrality
    - inverse of the sum of distances between each node and all other nodes in the network

Network model

- Static view
  - only one graph: $G = (V, E)$
- Dynamic view
  - set of graphs: $G_t = \{G_1, G_2, \ldots, G_n\}$
  - each graph $G_t = (V_t, E_t)$ is a *snapshot* from the network model during δ time units

Network model

- Set of nodes remains unchanged
- Network dynamics is given by link changes

Diffusion process

Node **u** needs to send a message to node **v**

- **Flooding**: at each step, all direct neighbors receive the message
- **Random walk**: at each step, only one randomly selected neighbor receives the message
- **Centrality walk**: at each step, select the highest centrality neighbor to receive the message

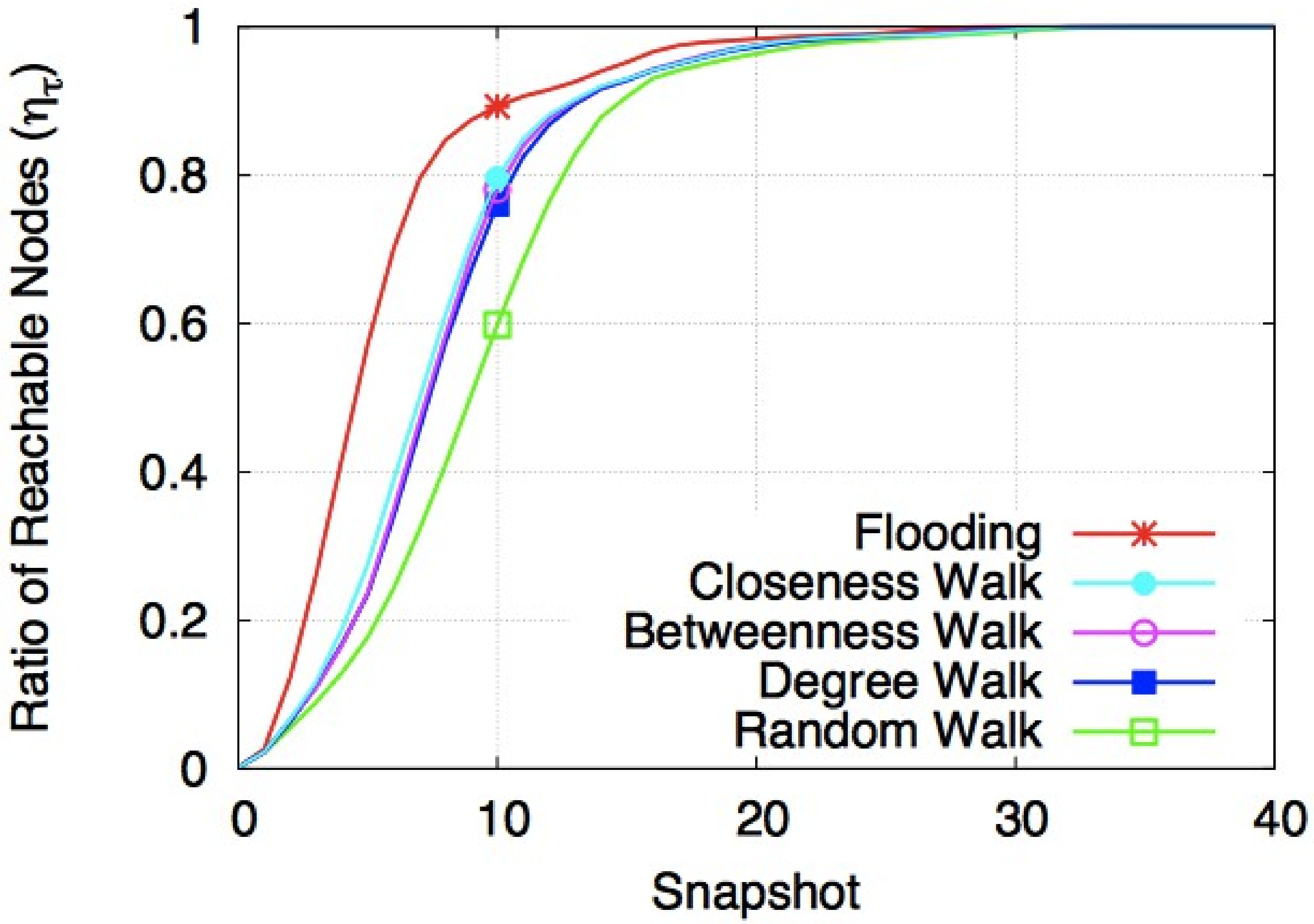
Datasets

- **Infocom**
  - two-day contacts between iMotes devices
  - ~ 70 students and researchers in Infocom 2006 conference
  - snapshots: δ = {1,15} minutes
- **SopCast**
  - P2P live video streaming system
  - one-hour trace with 334 nodes
  - snapshot: δ = 1 second

Static view - SopCast

- Cover time is very fast
  - to reach 100% of nodes:
    - flooding: two snapshots
    - centrality walks: three snapshots
    - random walk: four snapshots

Dynamics impact - SopCast



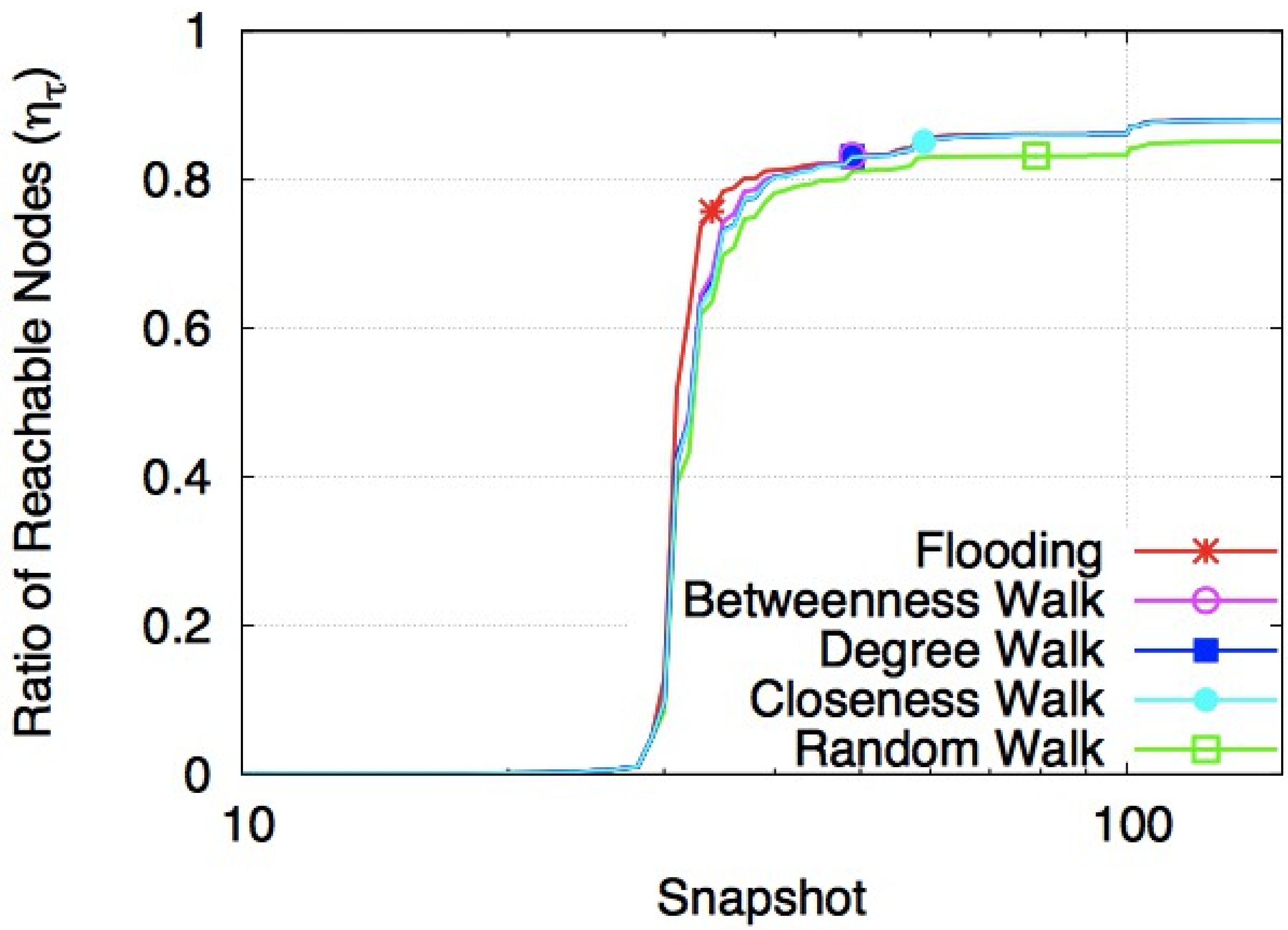
Centrality-driven walks accelerate the diffusion

Impact of system knowledge

system knowledge: proportional to the δ snapshot size

System knowledge - Infocom

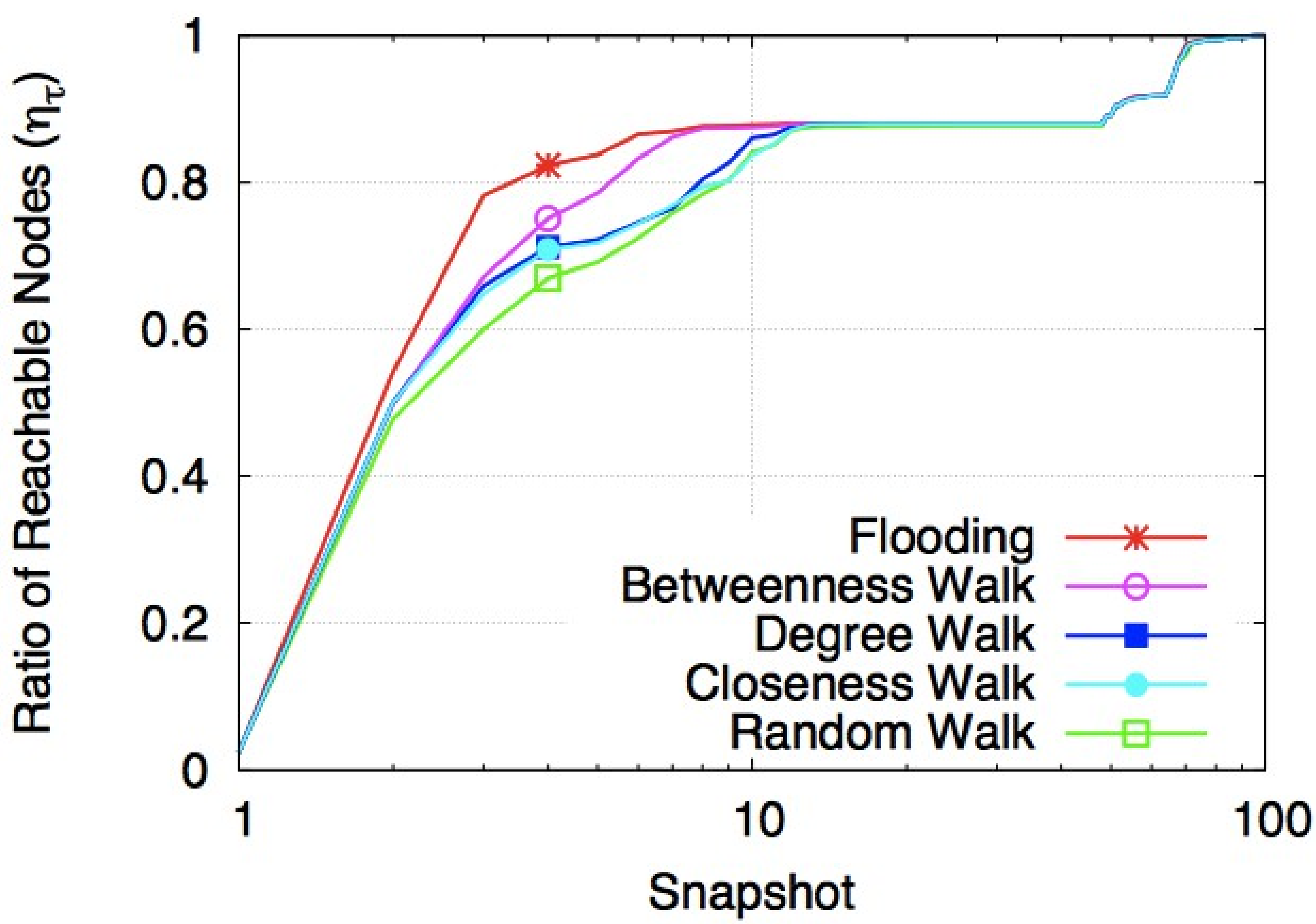
δ = 1 minute



- Only 85% of nodes reached
- Different centrality walks behave almost equally

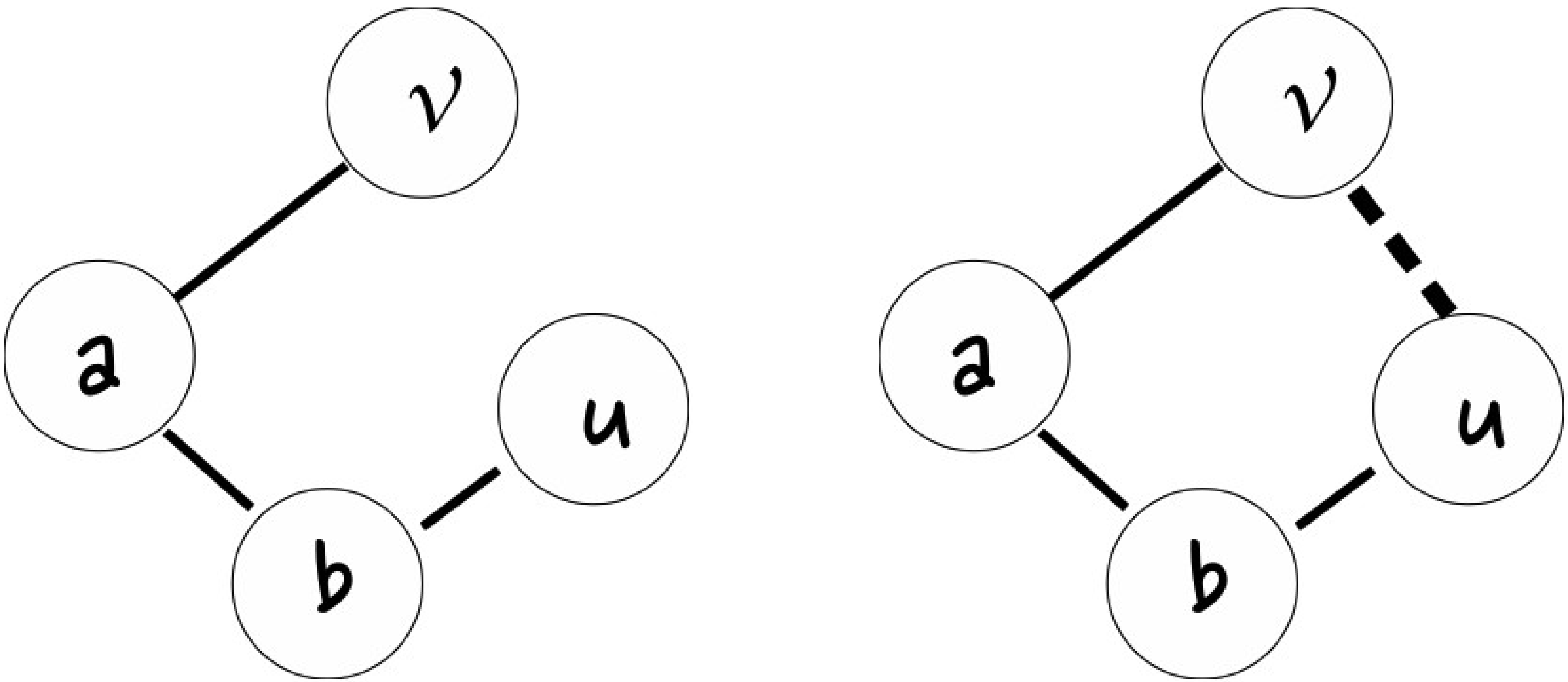
System knowledge - Infocom

δ = 15 minute

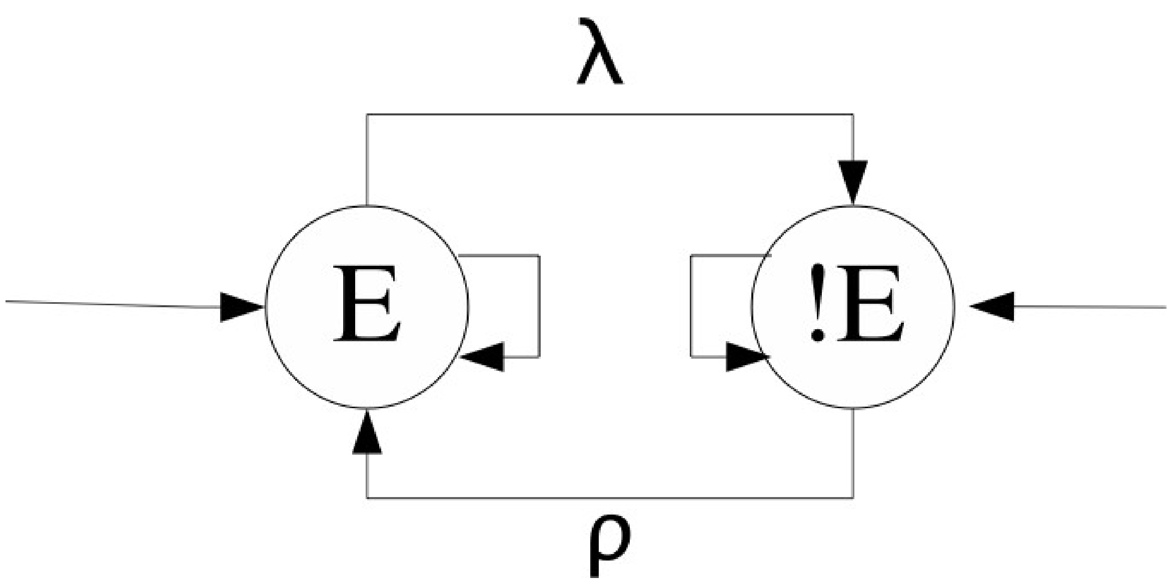


- 100% of nodes reached
- faster diffusion
- betweenness walk approximates better flooding
- most popular participants are better identified

Link prediction



Simple link prediction model

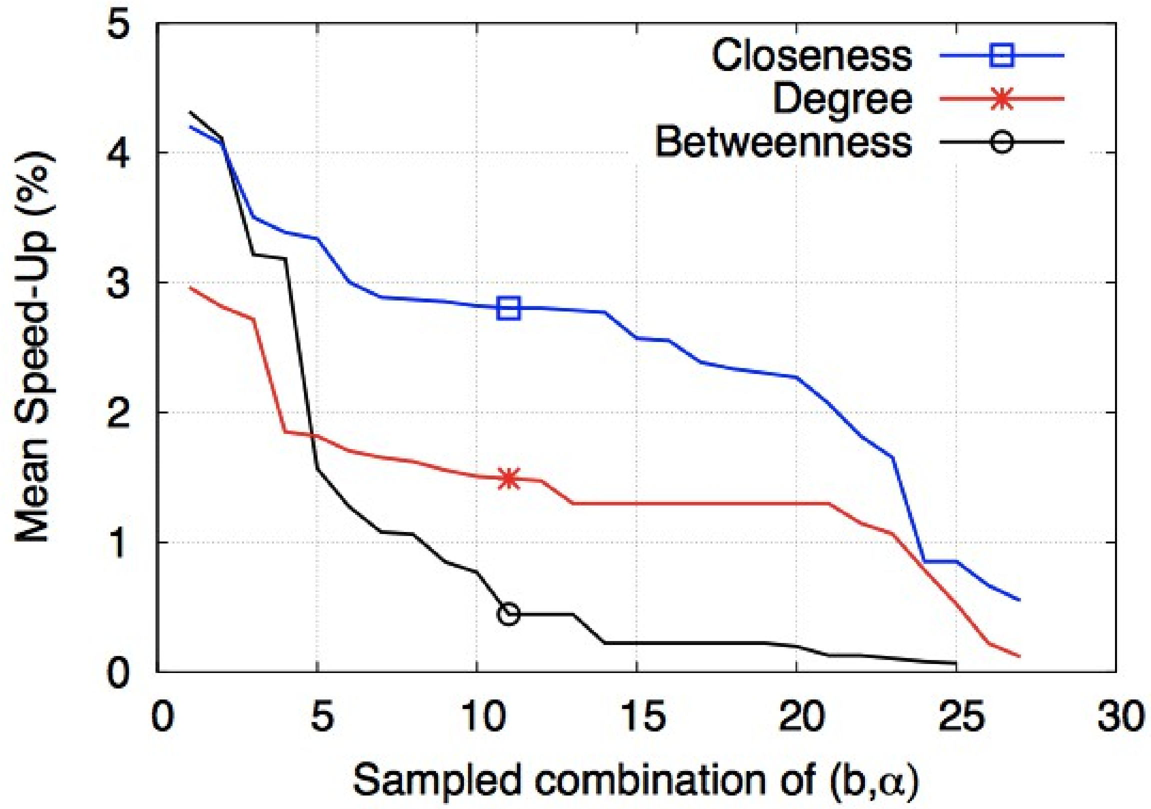


$$P(E_{v\to u}^{t+1}) = \frac{E_{v\to u}^{t} + E_{v\to u}^{t-1} * \alpha + E_{v\to u}^{t-2} * \alpha^2 + \ldots + E_{v\to u}^{t-b} * \alpha^b}{1 + \alpha + \alpha^2 + ... + \alpha^b}$$

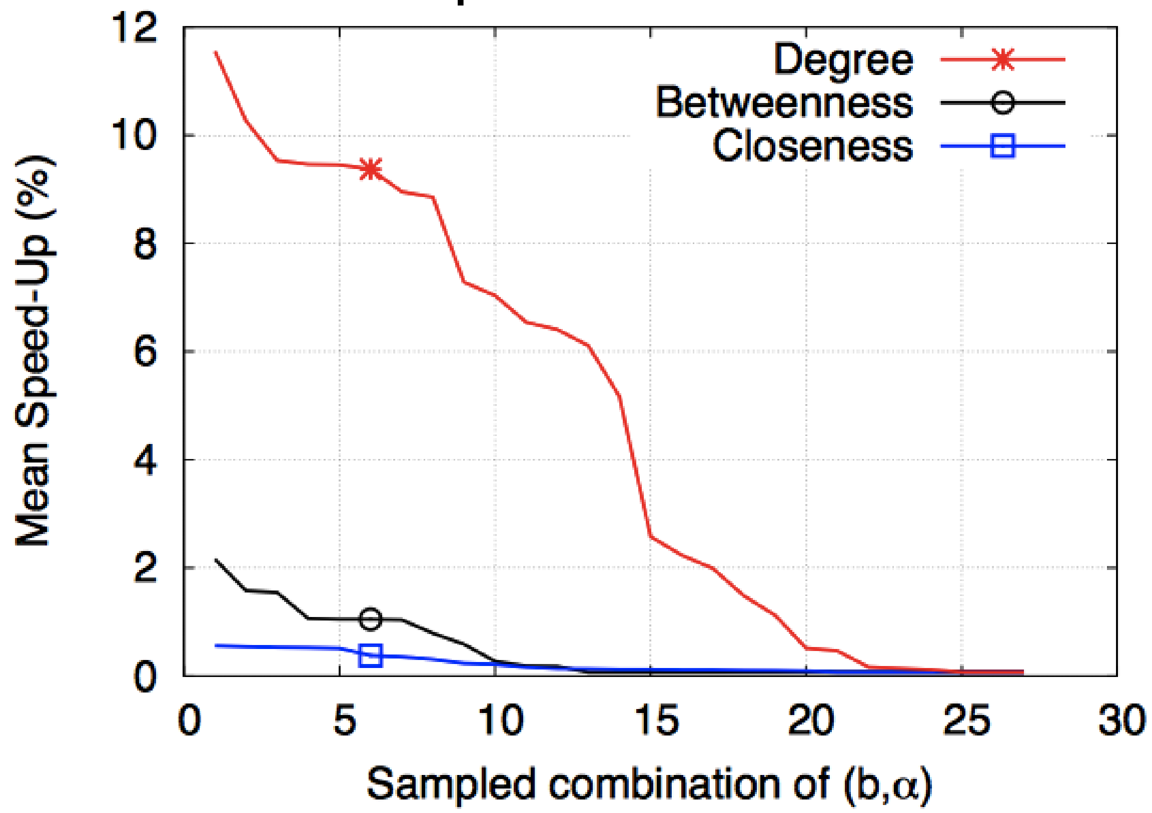
- moving window of last **b** snapshots
- each past snapshot is weighted by α ($0 < \alpha < 1$) that decreases exponentially

Experimental results

Infocom dataset



SopCast dataset



combinations of (b, α)

$5 \leq b \leq 50; \quad 0.1 \leq \alpha \leq 0.9$

Final remarks

- Centrality-driven selection of next-hop
  - accelerates diffusion process
  - levels off a trade-off between
    - flooding: low cover time and high message cost
    - random walk: large cover time and low message cost

Final remarks

- Preliminary results on link prediction
  - accelerates the diffusion process
- Future work: further investigate link prediction
  - heuristics to determine convenient combinations of (b, α) for particular networks
  - adaptive methods?

Thanks!
Obrigada!

Ana Paula Couto
Universidade Federal de Minas Gerais (UFMG)

ana.coutosilva@dcc.ufmg.br

Acknowledgements:

CAPES

CNPq
Conselho Nacional de Desenvolvimento
Científico e Tecnológico

FAPEMIG
Fundação de Amparo à Pesquisa do
Estado de Minas Gerais

FAPERJ
Fundação Carlos Chagas Filho de Amparo
à Pesquisa do Estado do Rio de Janeiro