



Push-Pull Gossiping for Information Sharing in Peer-to-Peer Communities

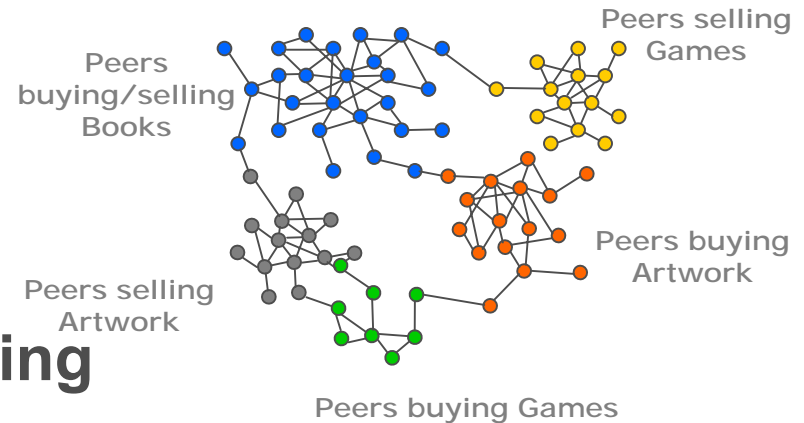
Mujtaba Khambatti, Kyung Ryu, and Partha Dasgupta

Arizona State University



Abstract + Outline

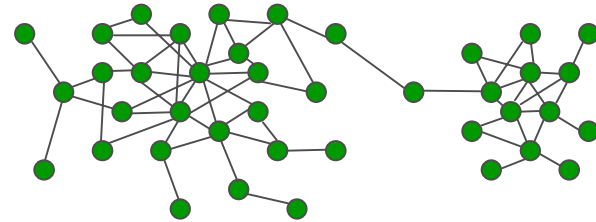
- Introduce notion of
 - communities
- Distributed Discovery
 - identify special peers
- Novel Push-Pull Gossiping
 - information dissemination
- Evaluation
 - push to few peers achieves information availability



Communities in P2P Systems

■ Peer-to-Peer systems

- distributed systems
- computing elements called peers
- comparable roles, ...
- communicate, share ...



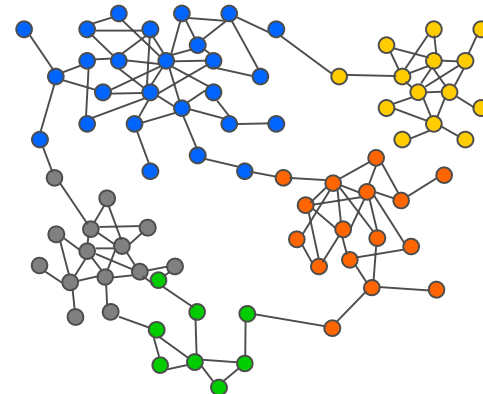
■ Peers have interests

- implicitly / explicitly obtained
- reflect the activities / interests



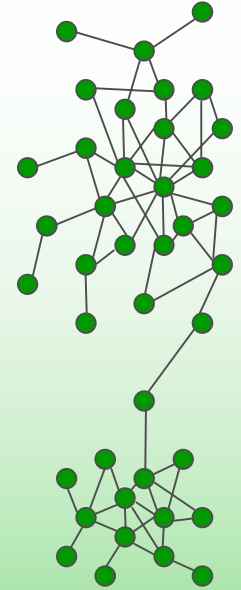
■ Communities of peers

- formed implicitly based on interests
- collection of peers sharing interests

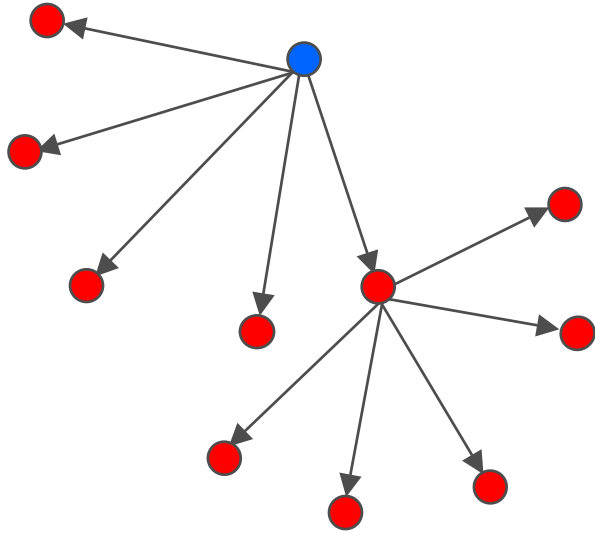


Forming a P2P Network

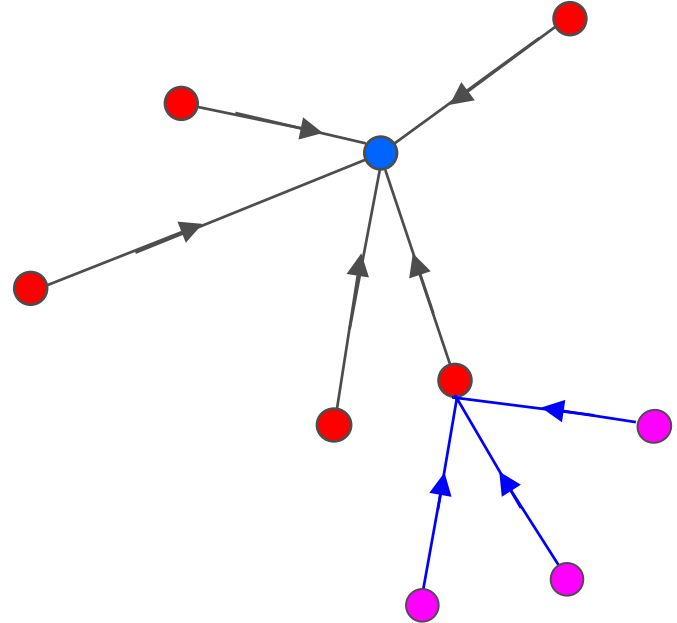
- P2P Network as a graph (V,E)
 - V = peers
 - E = links
- Links are overlay connections
 - similar to friendships / http links
- Links can be created using:
 - a special peer chosen by the domain
 - a known peer (friend)
 - a well-known peer
- Links give weights to Interest Attributes
 - outlink weight
 - inlink weight



Weight Calculation for each Interest



Outlink Weight



Inlink Weight

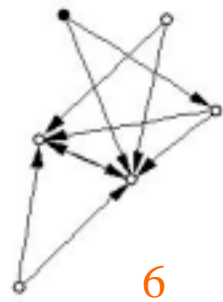
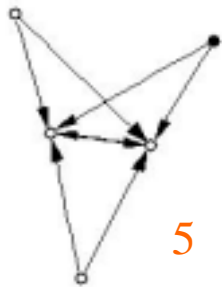
Rules for joining peers



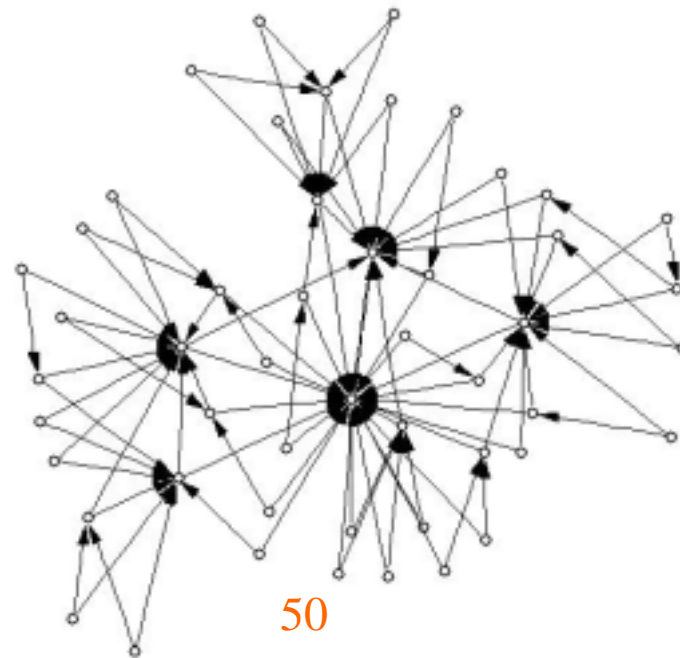
+ $N+1$ fully connected peers

+ $X \xrightarrow{p(k)} A, A \in \text{list of existing peers}$

+ X connects with N most popular neighbors of A

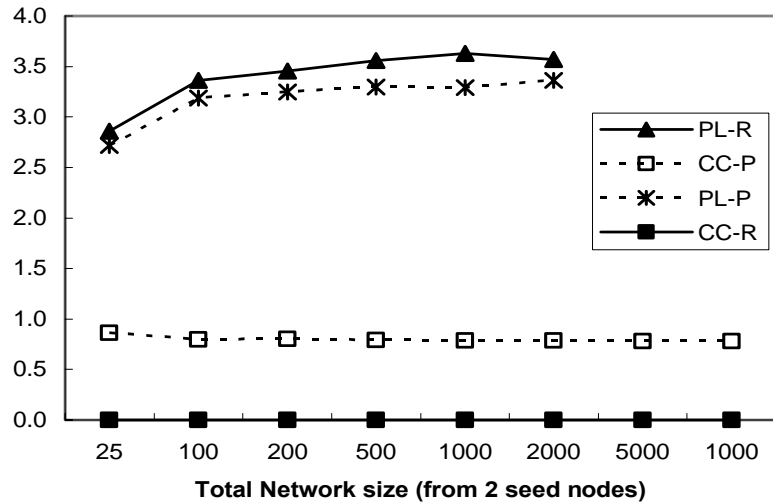


10



50

Evaluation of Rules



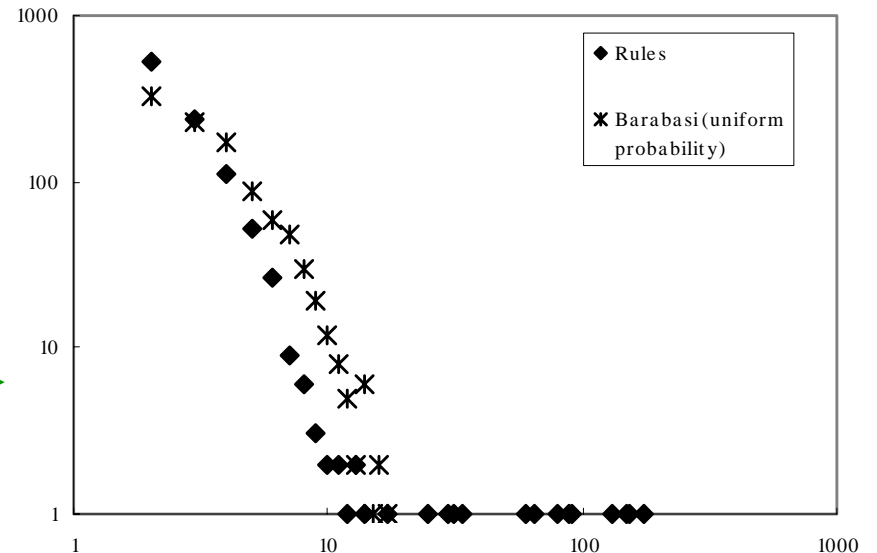
■ Small-World Property

- short characteristic path length
- high clustering coefficient



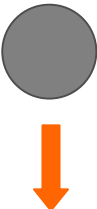
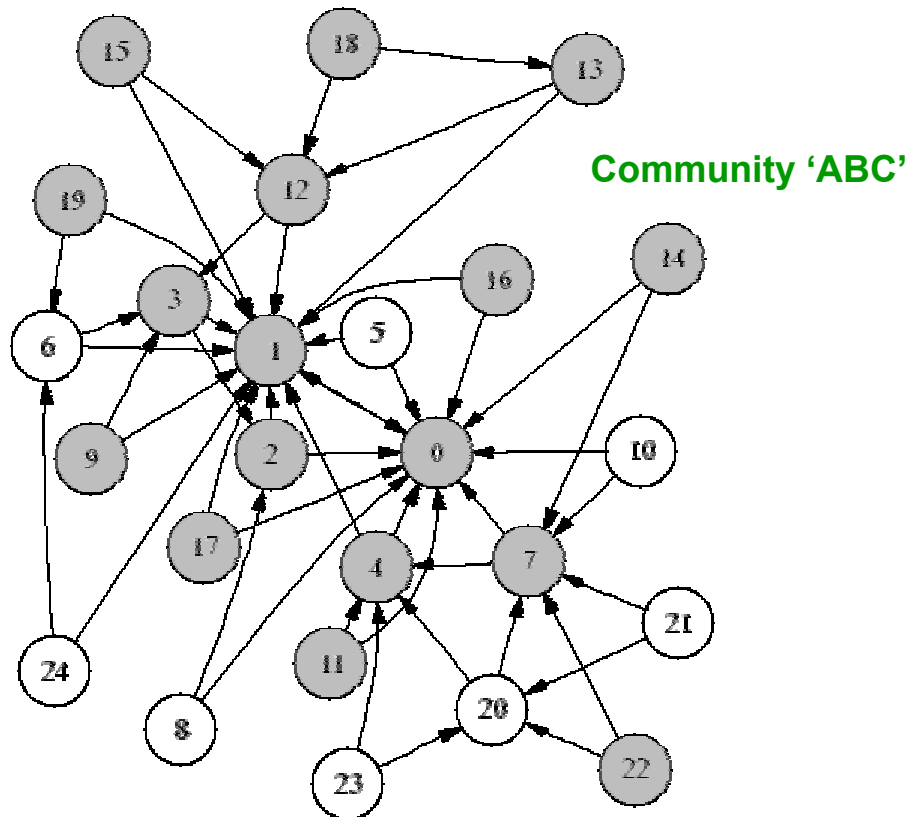
■ Scale-Free Property

- power-law exponent: -1.13
- network size, 1000 nodes



Information Dissemination

- After Formation & Discovery [**earlier work**]
 - p2p network organized into communities, link weights computed



Interest	OulinkW	InlinkW
G	0.4	0.25
F	0.8	0.12
T	0.6	0.45
D	0.23	0.32



Information Dissemination

- **Challenge:**
 - disseminate information within a community of peers
- **Related Approaches:**
 - flooding
 - depth-limited flooding
 - **CANNOT** be compared to rumor spreading
 - ✦ N is unknown
 - ✦ each peer only knows a subset of N
- **Our Solution**
 - distributed discovery to gather data on members of a community
 - push-pull gossiping phase

Distributed Discovery: Overview

- Low overhead, simple, protocol for community C
 - terminates easily, although N is not known
- Differs from Lamport Distributed Snapshot [see paper why]
- Identifies two things:
 - number and identity of peer members of a community
 - list of special peers – called SEERS
- Seers are peers with high involvement (/) values
 - they know more peers within community C than most other peers

calculation explained in
paper & earlier work

Distributed Discovery: Nuts & Bolts

Vector ID	Peer ID	Peer Involvement
-----------	---------	------------------	-----	-----	-----

■ Initiation

- any peer can initiate by creating a vector
 - ✦ vector ID helps detect conflicts and discard vectors with lower IDs

■ On receipt of vector

- append information
- send to all neighbors claiming common interest of community C

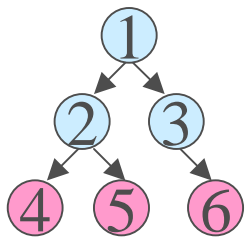
■ Termination detection

- if all neighbors have already received vector, send vector to initiator
- initiator checks frequency of vector arrivals

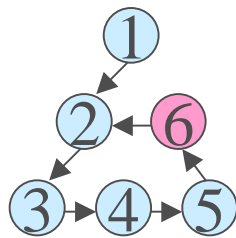
■ Discovery at the Initiator

- union of all vectors provides initiator with information about community members

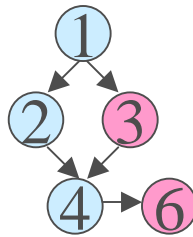
Explanation & Evaluation



Case 1



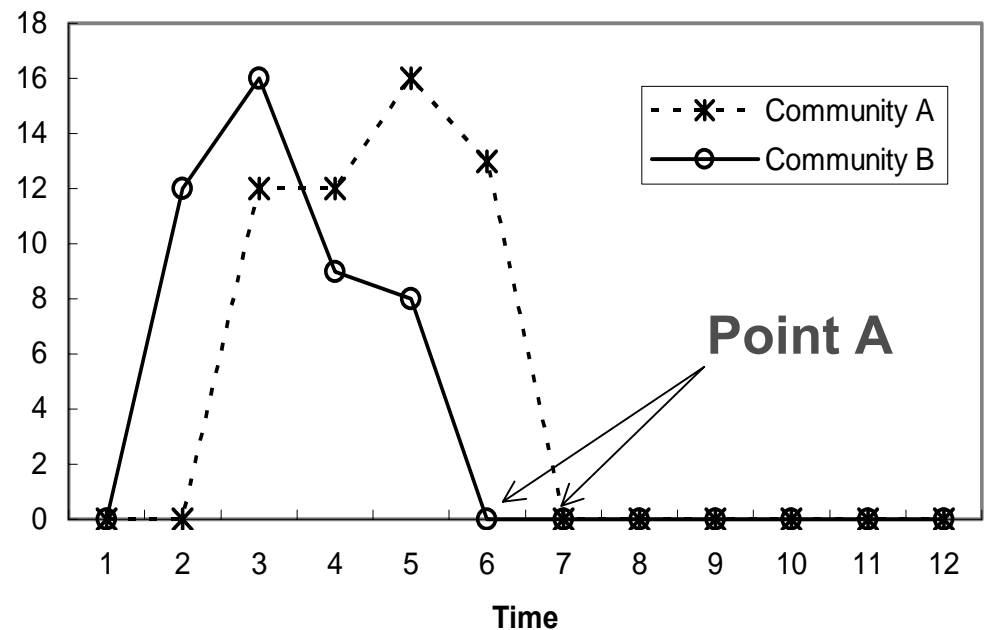
Case 2



Case 3

■ Frequency Graph

- vectors arriving
- network = 1000 nodes
- random delays & faults



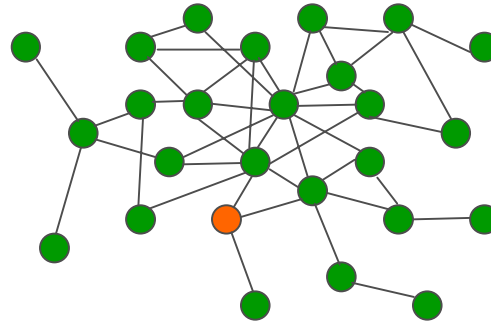
P2P Gossiping: Overview

- Undirected intra-community information dissemination
- Differs from Rumor Spreading:
 - N is not known
 - peers are not selected at random
- Initiator sorts list of member peers based on involvement
- Top 5%-10% are called Seers
- P2P Gossiping
 - information is PUSHED to seers
 - information is PULLED from seers (by other peers)

P2P Gossiping: Nuts & Bolts

■ Scenario 1

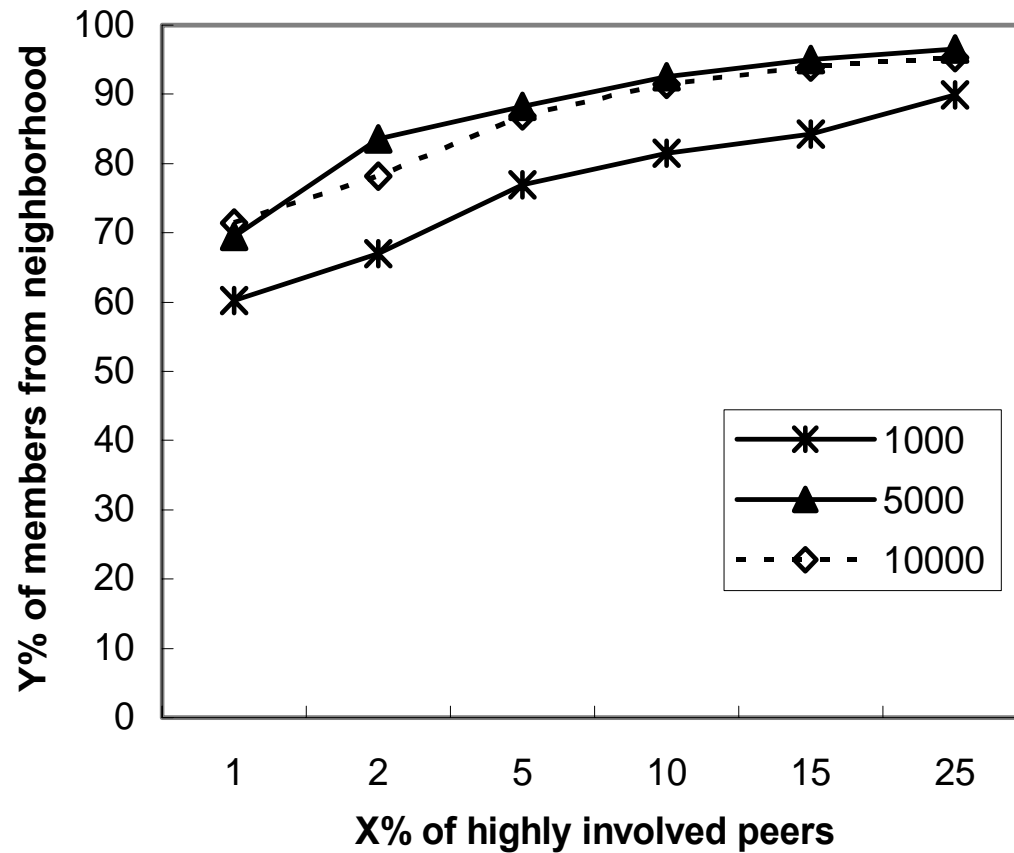
- initiator broadcasts list of seers to all community members
- any peer with gossip information sends it to seers (**PUSH**)
- peers can retrieve this information from seers (**PULL**)



■ Scenario 2

- any peer with gossip information sends it to initiator
- initiator sends information to list of seers (**PUSH**)
- peers can retrieve this information from seers (**PULL**)

Evaluation



Conclusions

- **P2P communities**
 - **natural, implicit organization of distributed systems**
- **P2P Network formed with Rules**
 - **guarantees small-world & scale-free properties**
- **Undirected intra-community communication**
 - **distributed discovery**
 - + **simple, low overhead protocol with trivial termination point**
 - **push-pull gossiping**
 - + **guarantees information availability, economy of messages**



More Information

- Current paper
- Efficient Discovery of Implicitly Formed Peer-to-Peer Communities
Int'l. Journal of Parallel and Distributed Systems and Networks
- Peer-to-Peer Communities: Formation and Discovery
14th IASTED Int'l. Conf. Parallel & Distributed Computing Systems
- <http://calypso.eas.asu.edu/>

