Search/Data Dissemination for Decentralized Online Social Networks

Anis Nasir

Online Social Networks

- Search (User/Groups)
- Data Dissemination
- Storage

- Security/Privacy

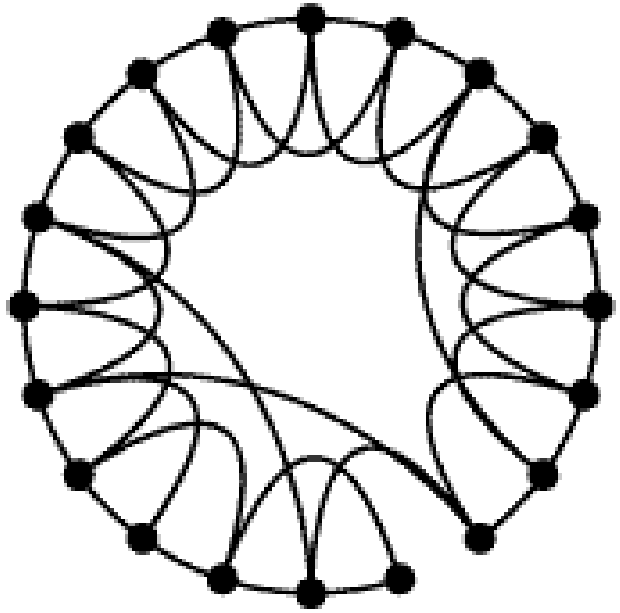
Online Social Networks

- **Search (User/Groups)**
- **Data Dissemination**
- Storage

- Security/Privacy

Nature of Social Networks

- Small world networks
  - High Clusterization
  - Small diameters

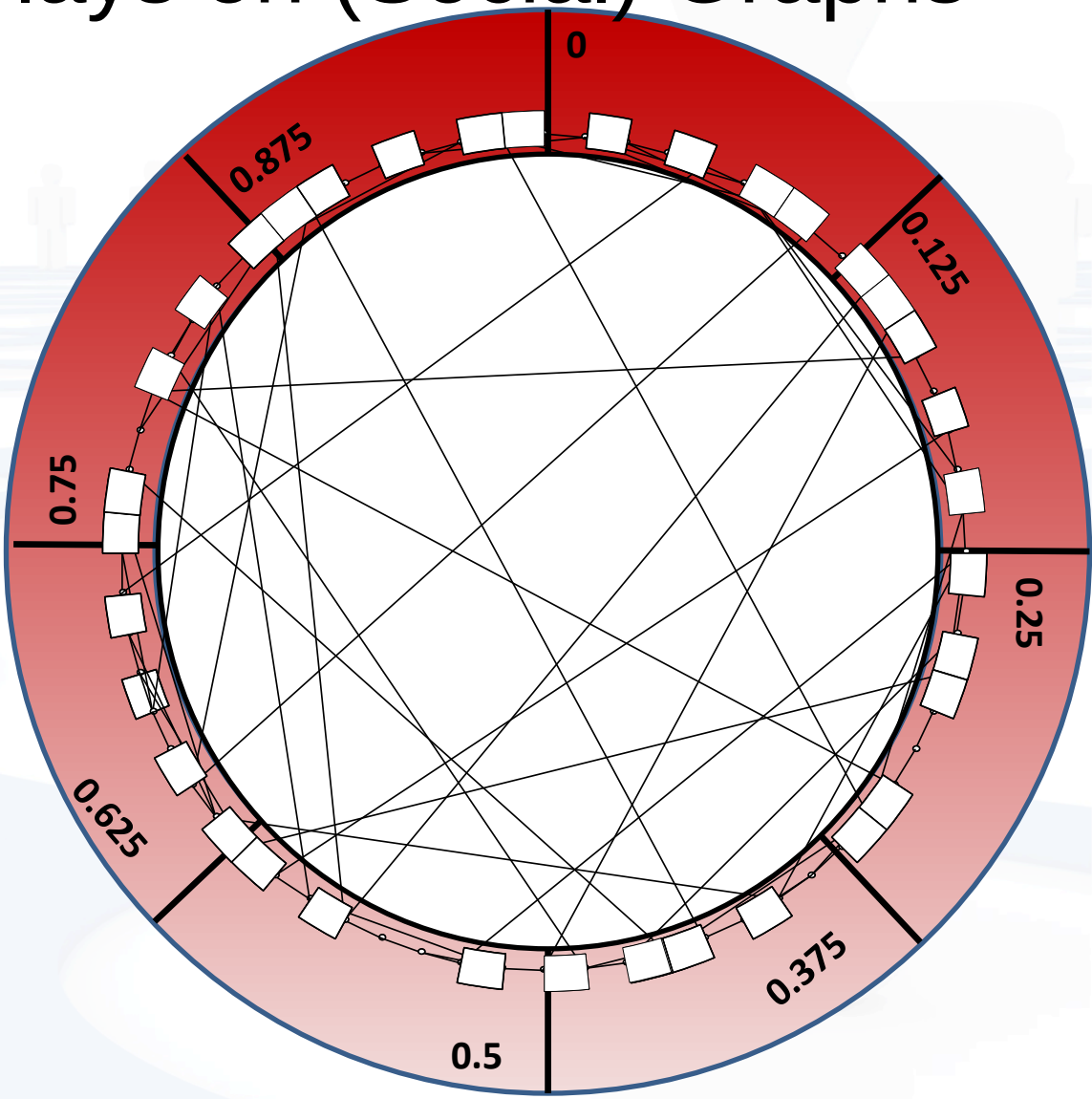


Search in Social Networks

- Depends on **knowledge** of a node about
  - **neighbors** position **relative to the destination**
- Does these properties hold for Social Networks?
- Can we route messages using the shortest path without having such knowledge?

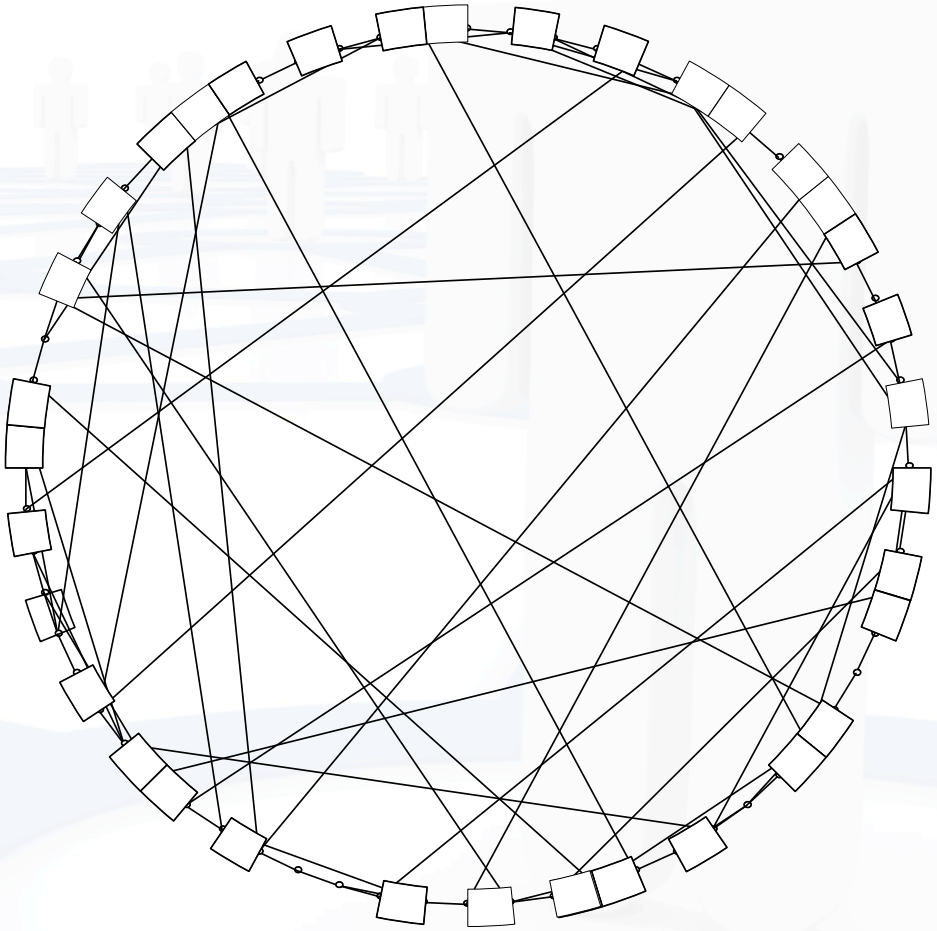
Structured Overlays on (Social) Graphs

- Take a set of peers
- Imbed into d-dimensional ID space
  - Each peer gets an ID
- Create connectivity based on their ID following certain rules
  - Results in a Navigable Small-World network
  - E.g., Chord, Symphony, Oscar etc
- **What if we do have a graph but no labels?**
  - **How to recreate them?**



Structured Overlays on (Social) Graphs

- Existing approach (Sandberg et al.)
  - Assign random IDs
  - Each peer periodically exchanges info of their IDs with a random peer and decides whether to swap the IDs.
  - All links are of the same “importance”
- Weak or Strong ties?
  - We want an algorithm that orders all the neighbors by the “strength of their ties”
  - Modified cost function is expected to give much better results.



Data Dissemination

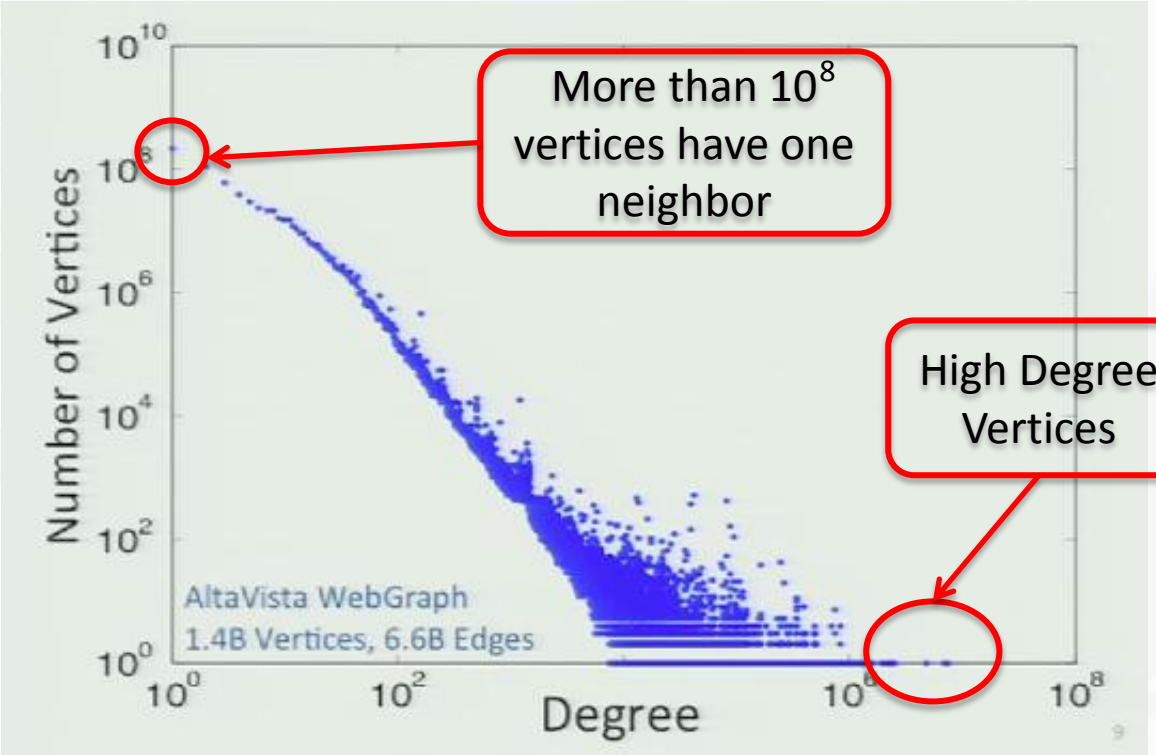
- Profile updates
- Status/Pictures/Videos/Music
- Wall Posts
- Comments
- Social Pages

Naïve Solution

- Each user stores his neighborhood information…
- …propagate updates directly to friends

Online Social Networks

- follow power-law degree distribution

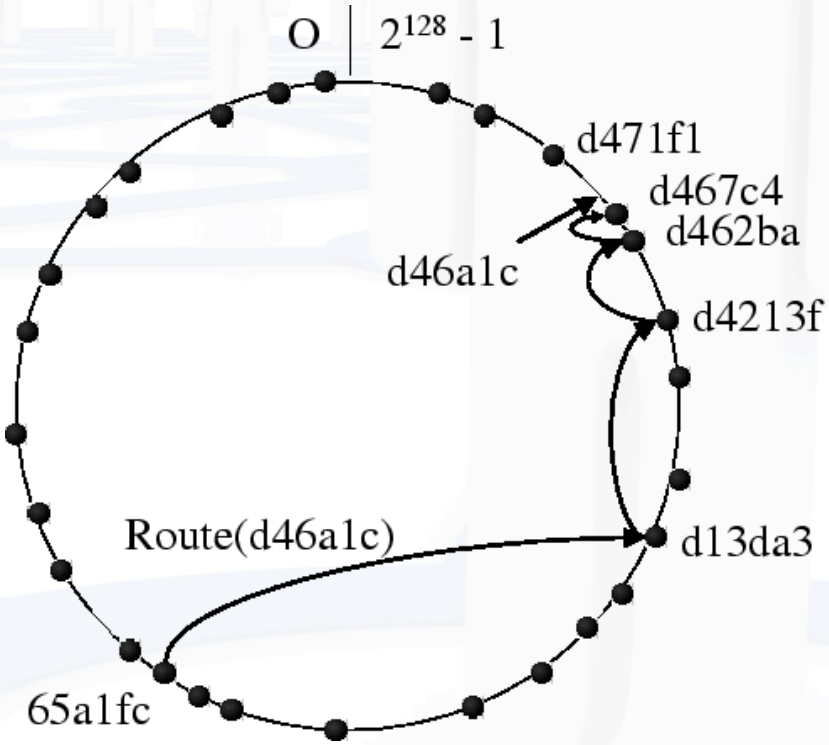


Issues

- Workload Imbalance
- Data availability
- Propagation time
- Scalability

Possible Solutions

- Structured Topology
  - Chord

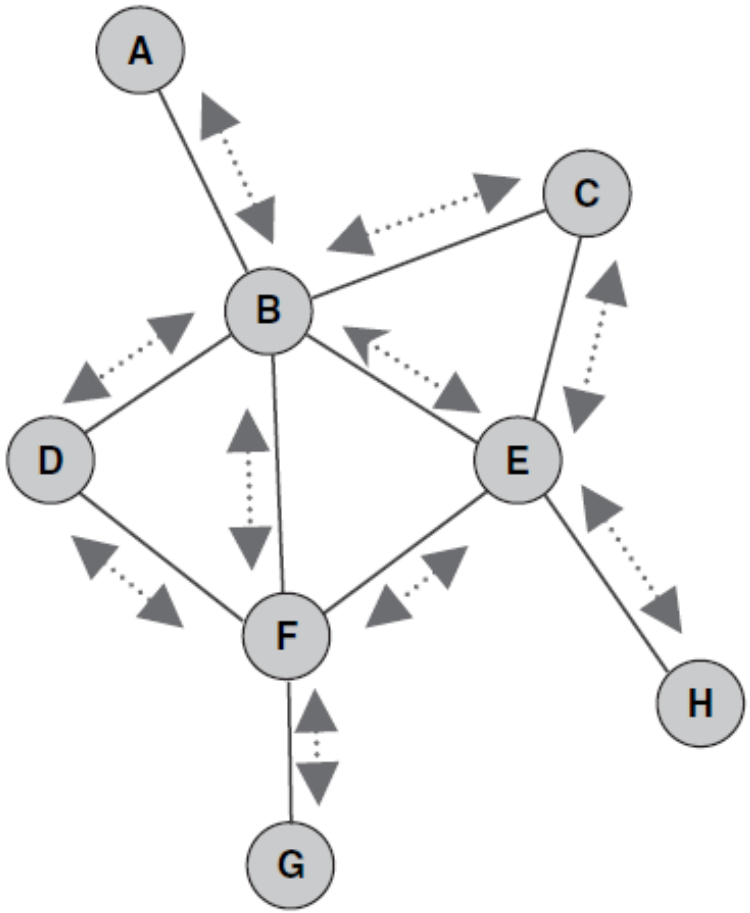


Challenges

- Structured Topology (maintenance overhead)
  - Chord (no social overlay awareness, long communication links, security issues)

Possible Solutions

- Unstructured Topology
  - No look up required
    - Flooding
    - Gossiping

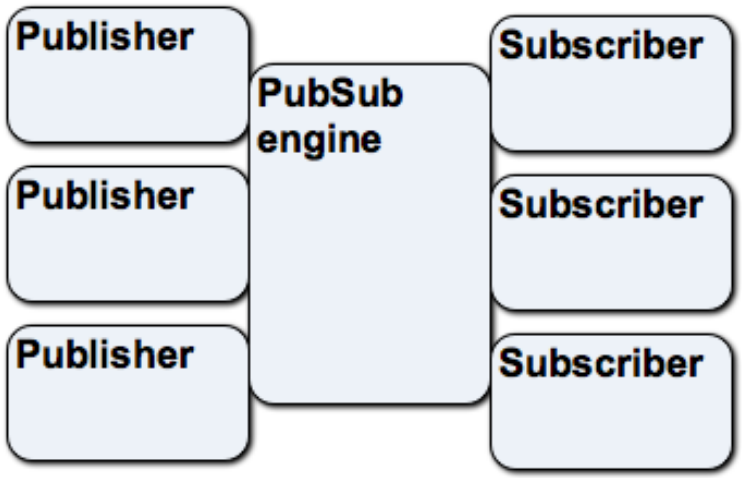


Challenges

- Unstructured Topology
  - No look up required (High Network Traffic)
    - Flooding (involves many uninterested peers)
    - Gossiping (involves many uninterested peers)

Possible Solutions

- Unstructured Topology
  - No look up required
    - Flooding
    - Gossiping
  - Look up required
    - Direct Data Exchange
    - Publish/Subscribe System



Challenges

- Unstructured Topology
  - No look up required (High Network Traffic)
    - Flooding (involves many uninterested peers)
    - Gossiping (involves many uninterested peers)
  - Look up required (implementation overhead)
    - Direct Data Exchange (work imbalance)
    - Publish/Subscribe System (work imbalance, Scalability)

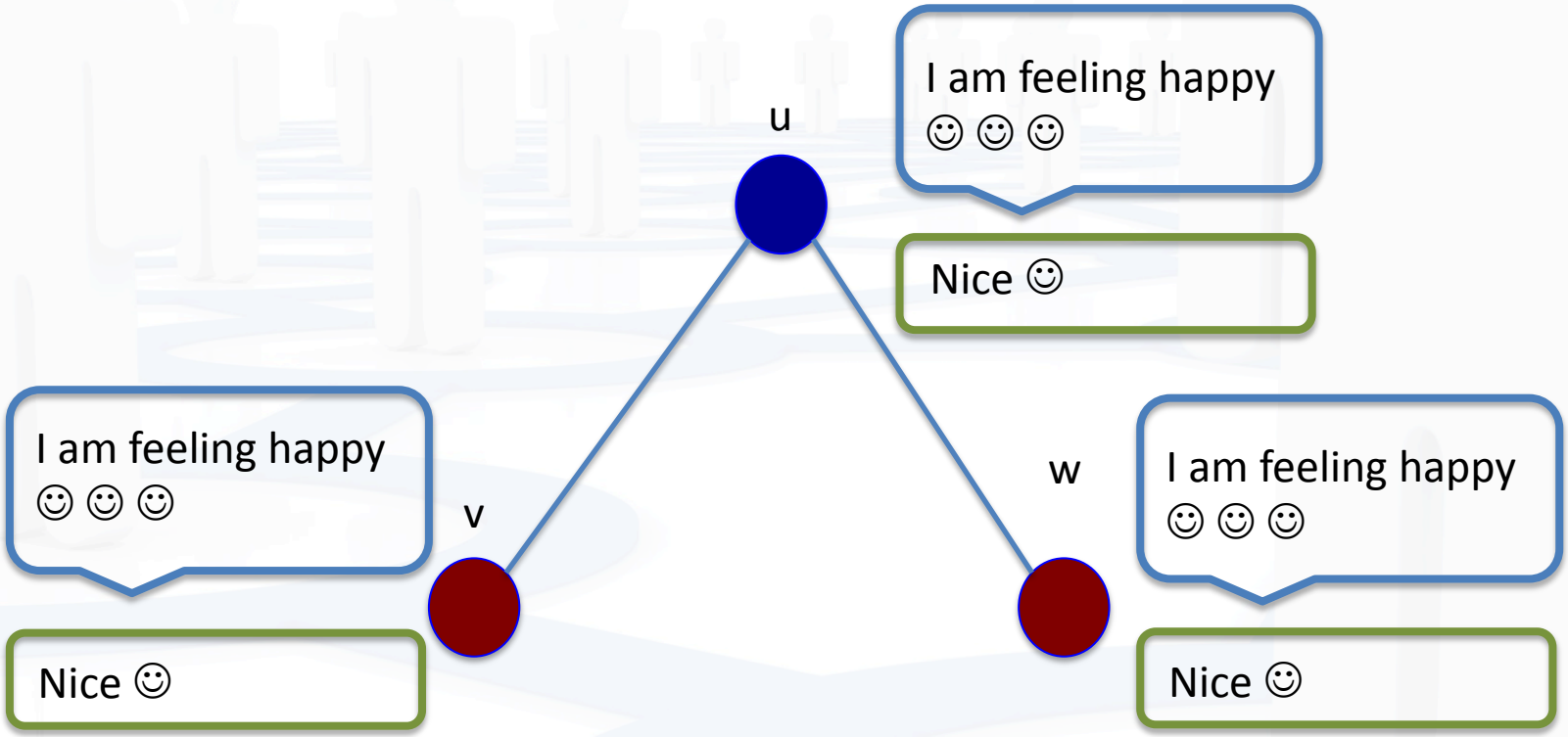
Constraints

- Minimize number of messages
- Minimize propagation time
- Minimize number of uninterested peers
- Workload Distribution

System Model

- A social network with undirected graph G=(V,E)
- For every uεV, f(u) defines friends of user u
- O(u) are all the updates generated by user u that should propagate to his friends f(u)

Problem


u
I am feeling happy
☺ ☺ ☺
Nice ☺
v
I am feeling happy
☺ ☺ ☺
Nice ☺
w
I am feeling happy
☺ ☺ ☺
Nice ☺

Approach

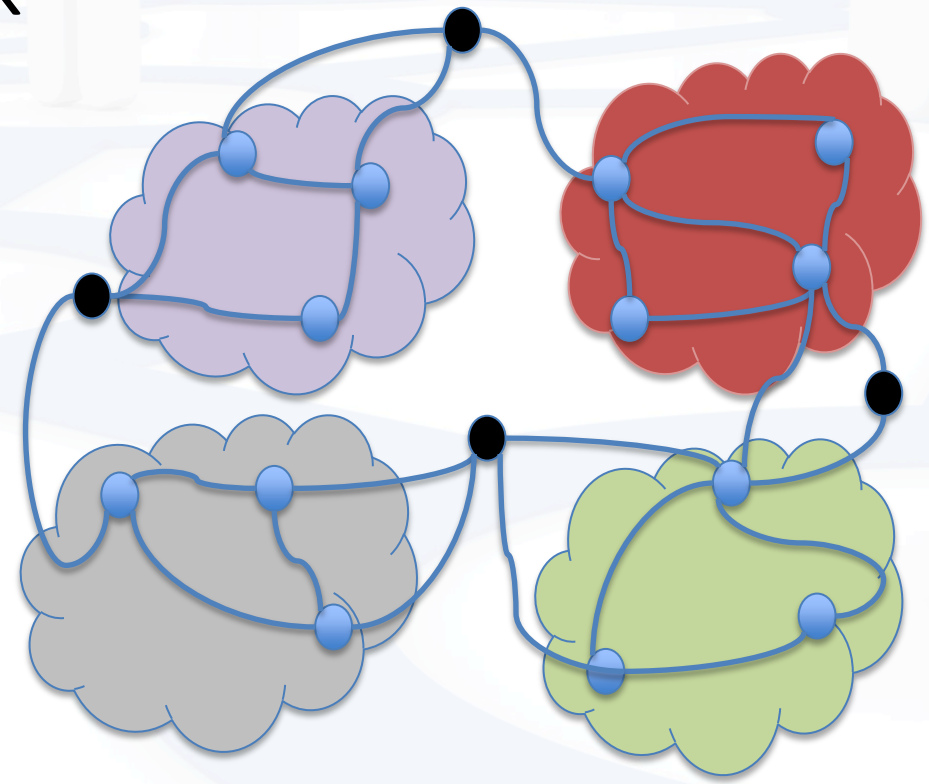
- An overlay for dissemination for online friends
- A overlay for dissemination for offline friends
- A overlay for dissemination of social groups/pages

Approach

- **An overlay for dissemination for online friends**
- A overlay for dissemination for offline friends
- A overlay for dissemination of social groups/pages

Approach

- A Gossip-based Hybrid Overlay for data dissemination in decentralized online social network



Approach

- Group nodes in fix sized clusters
- Select a gateway node for each cluster
- Relay between clusters using rendezvous routing

Search/Data Dissemination for Decentralized Online Social Networks

Anis Nasir