Overlay Networks Topology Management

Stefano Arteconi

Dipartimento di Scienze dell'Informazione
Università di Bologna

Project funded by the Future and Emerging Technologies arm of the IST Programme

Outline

- Peer-to-Peer (P2P) overlay networks
- Epidemic network topology
- Newscast algorithm
- Aggregation Protocol

Peer-to-Peer vs. Client-Server

- No distinction between Client and Server nodes
  - All nodes are *peers* (they both ask for and require services)
- High decentralization
  - No global knowledge or control
  - Each node has limited knowledge of the system and is able to interact with a (small) set of other nodes

P2P Overlay Networks

- Peer-to-Peer (P2P) networks are usually *overlay*
- Physical and communication level are assumed to exist
  - The Internet and TCP/IP protocol
  - Possibility to communicate between any pair of nodes
- Nodes: processes running at given locations
- Links: defined upon underlying physical topology
  - Logical label (i.e. IP,port pair)

Views and Topology

- The set of nodes that a peer knows about is called its *view*
- Typically, views are a (very) small subset of all nodes
- Views define the *overlay network* topology on top of the physical network
  - Overlay network nodes and links between them
- Overlay network topology easy to manage
  - No physical modification needed
  - Views exchange

Topology Management

- How to ensure that the overlay network topology satisfies certain properties:
  - has a desired structure (connected, random graph, ring, torus, binary tree, etc.)
  - *maintains* the desired structure in a dynamic setting (churn)
- Problem to be solved is *topology management*
- Solution inspired from the spreading of epidemics

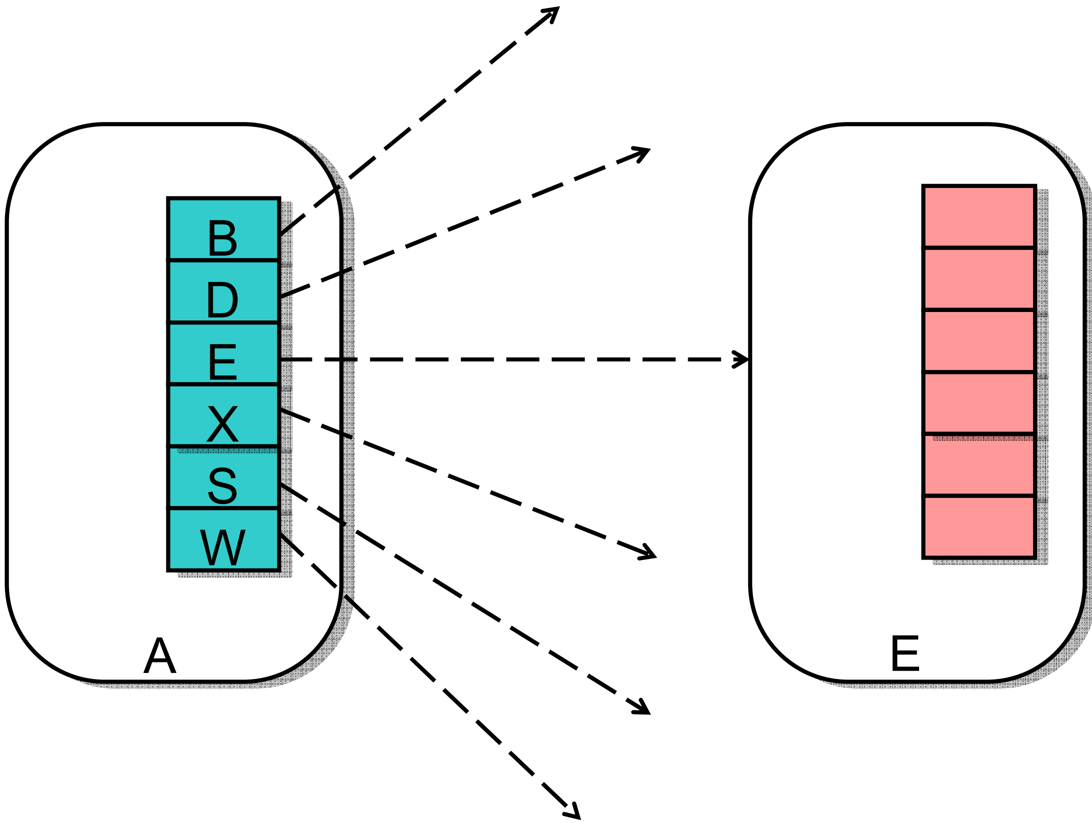
Epidemics Spreading Scheme

- Some people (nodes) in the population (network) is *infected* (has some information)
- Each person knows a set of other people (view)
- Each infected person transmits the infection to a subset of the people he knows
- Infection (information) propagates throughout the population

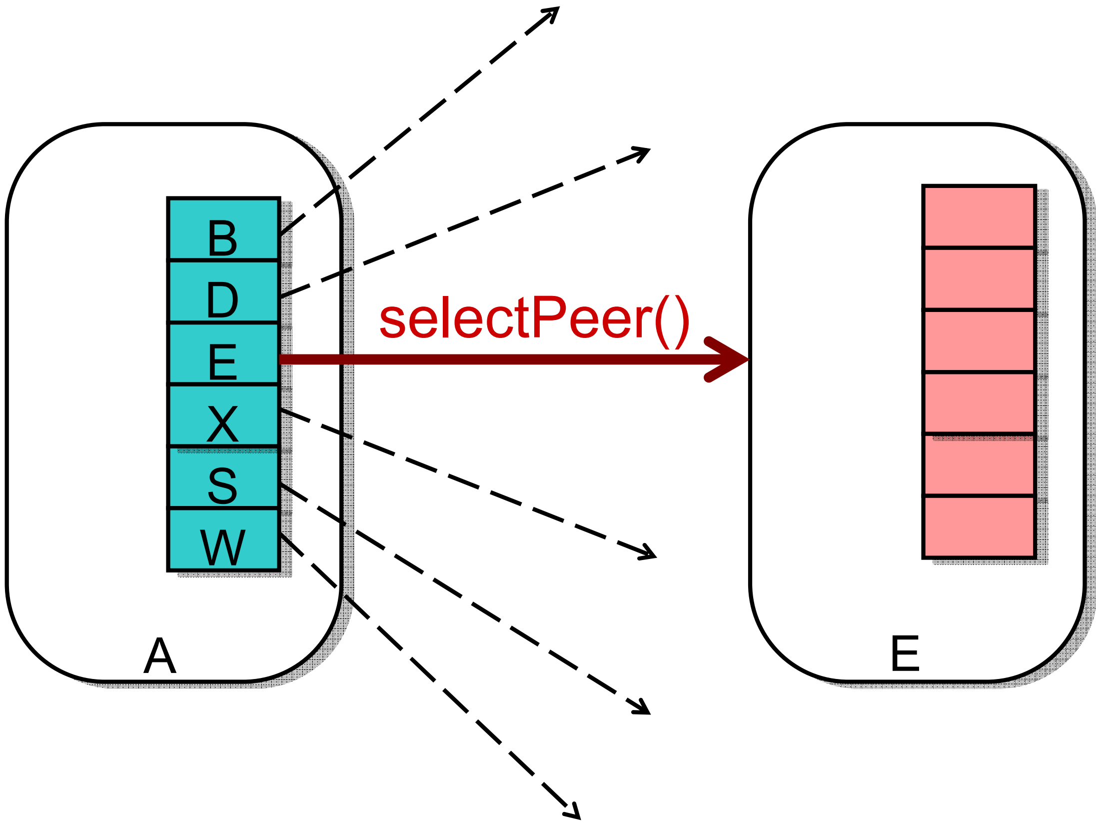
Topology Management: General Scheme

- Each node (say *i*) periodically:
  - Chooses a neighbor (say *j*) through selectPeer() function
  - Nodes *i* and *j* exchange their views
  - Views are merged/updated using updateState() function
- Through the definition of selectPeer() and updateState() functions it is possible to define different emerging global topology

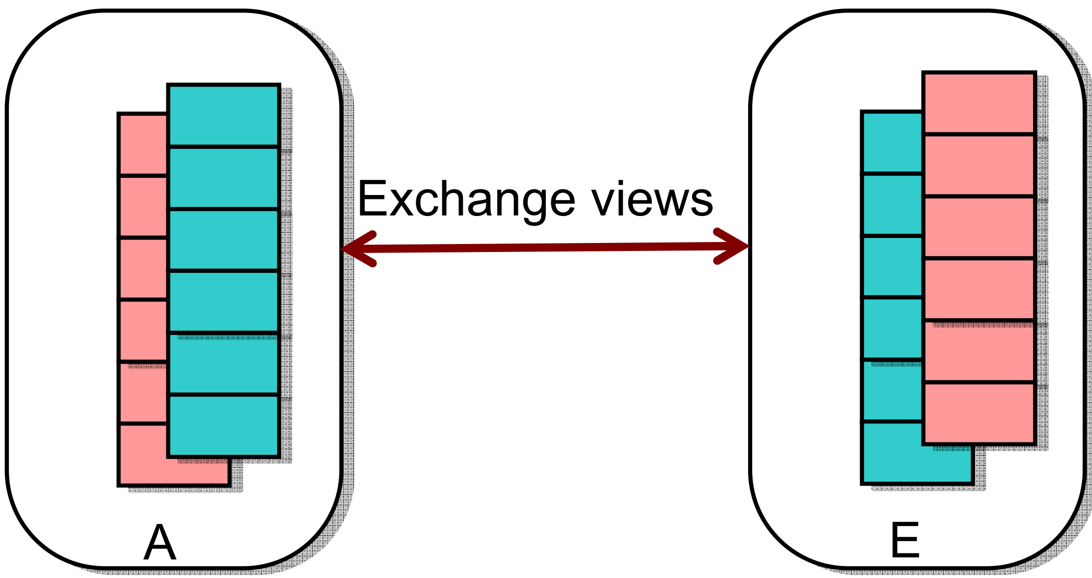
Topology management through Epidemics Views Exchange



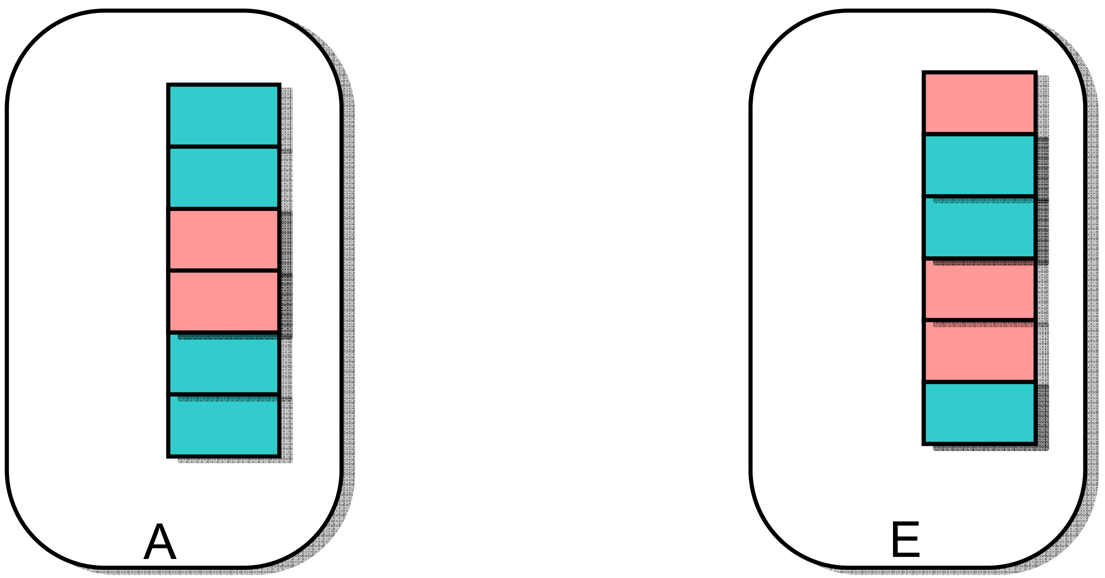
Topology management through Epidemics Views Exchange



Topology management through Epidemics Views Exchange



Topology management through Epidemics Views Exchange



Both peers apply updateState() thereby redefining topology

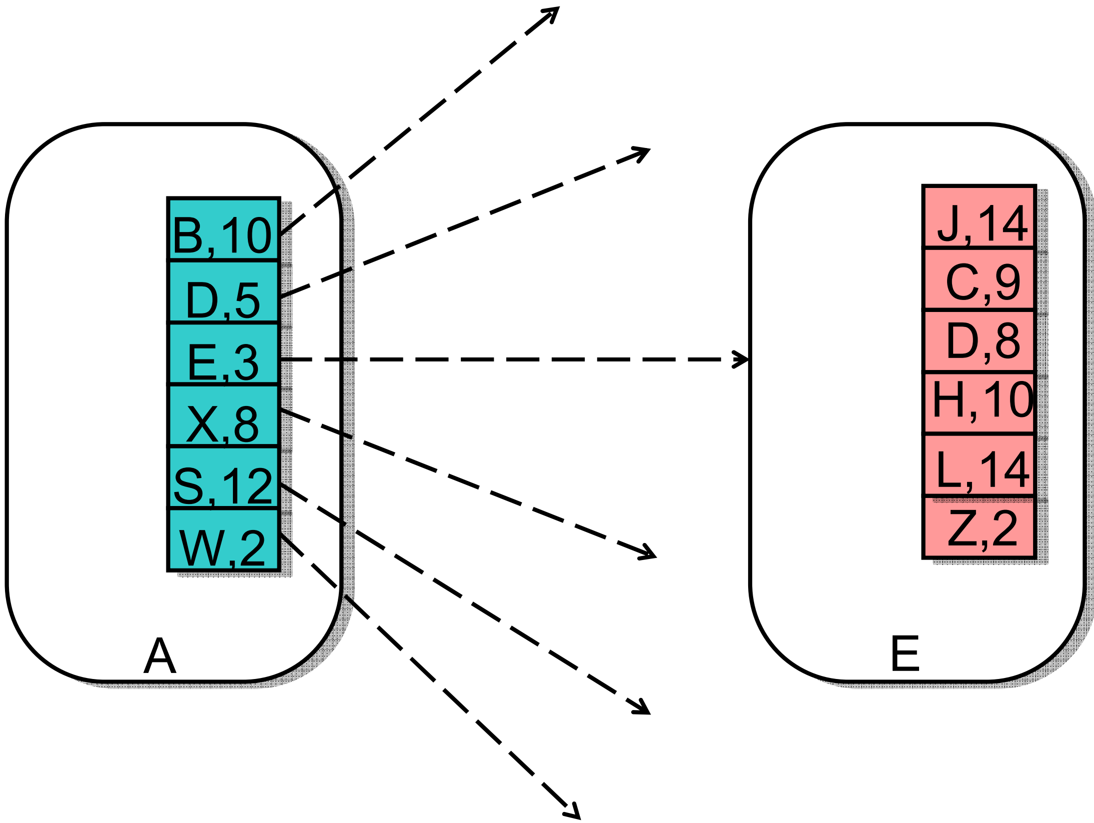
Using Epidemic Scheme

- Epidemic spreading model define the way information is exchanged
- Some basic steps are not specified
  - *SelectPeer()* method
  - *UpdateState()* method
- Many different target topology can be achieved by defining these two methods

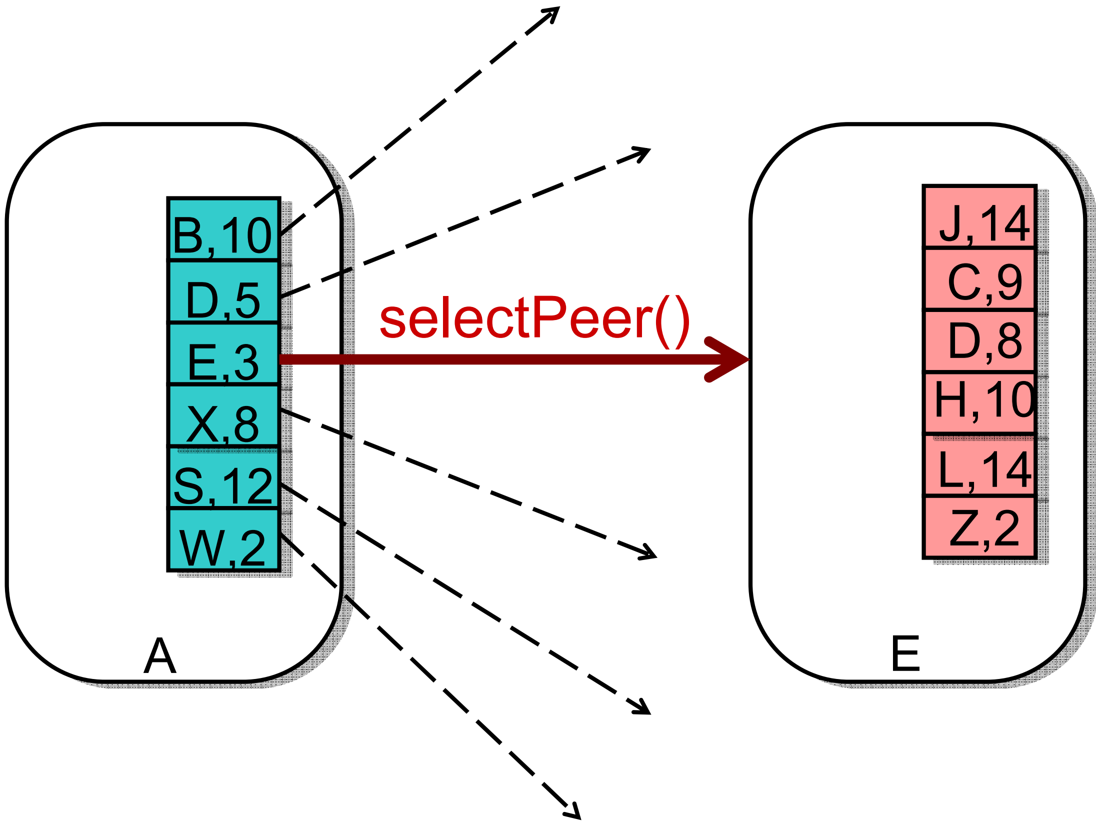
Random Topology: Newscast

- Node views
  - Limited size (maximum view size=*c*)
  - Node descriptors: timestamped links
- *SelectPeer()*
  - A random neighbor is chosen
- Each node adds a fresh descriptor of itself in its view
- Views are exchanged
- *UpdateState()*
  - Views are merged
  - Duplicated entries removed
  - To respect view size limit oldest descriptors are dropped

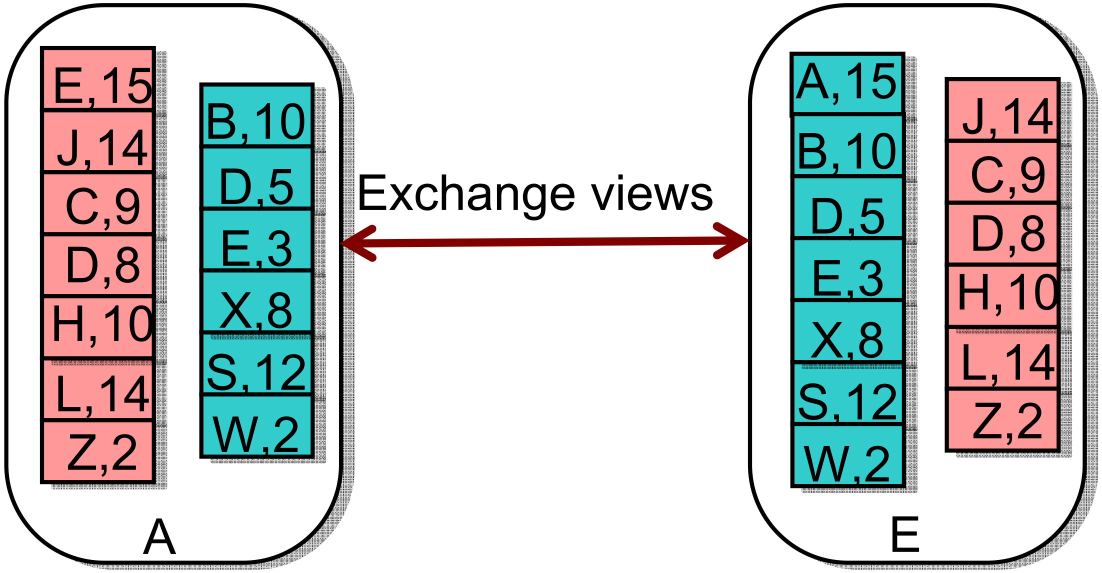
Newscast Example



Newscast Example

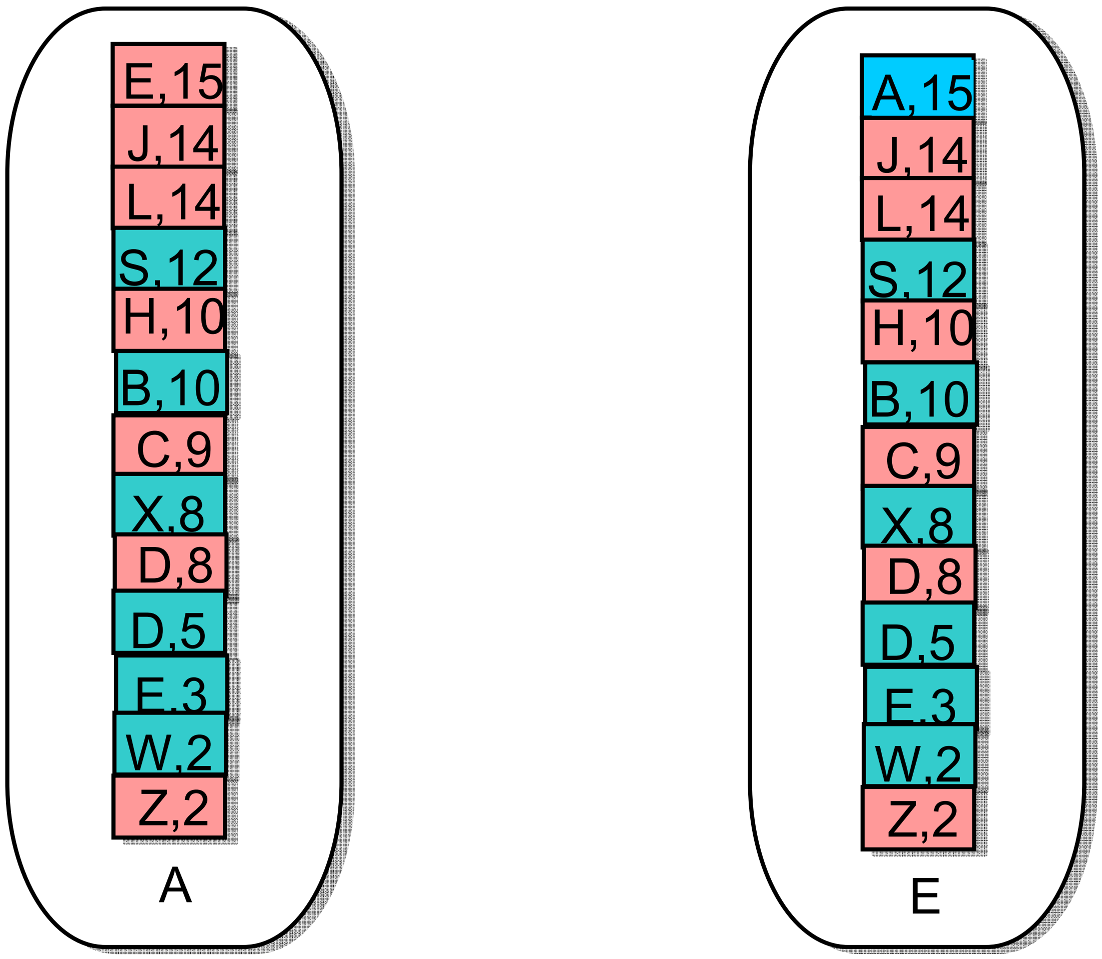


Newscast Example



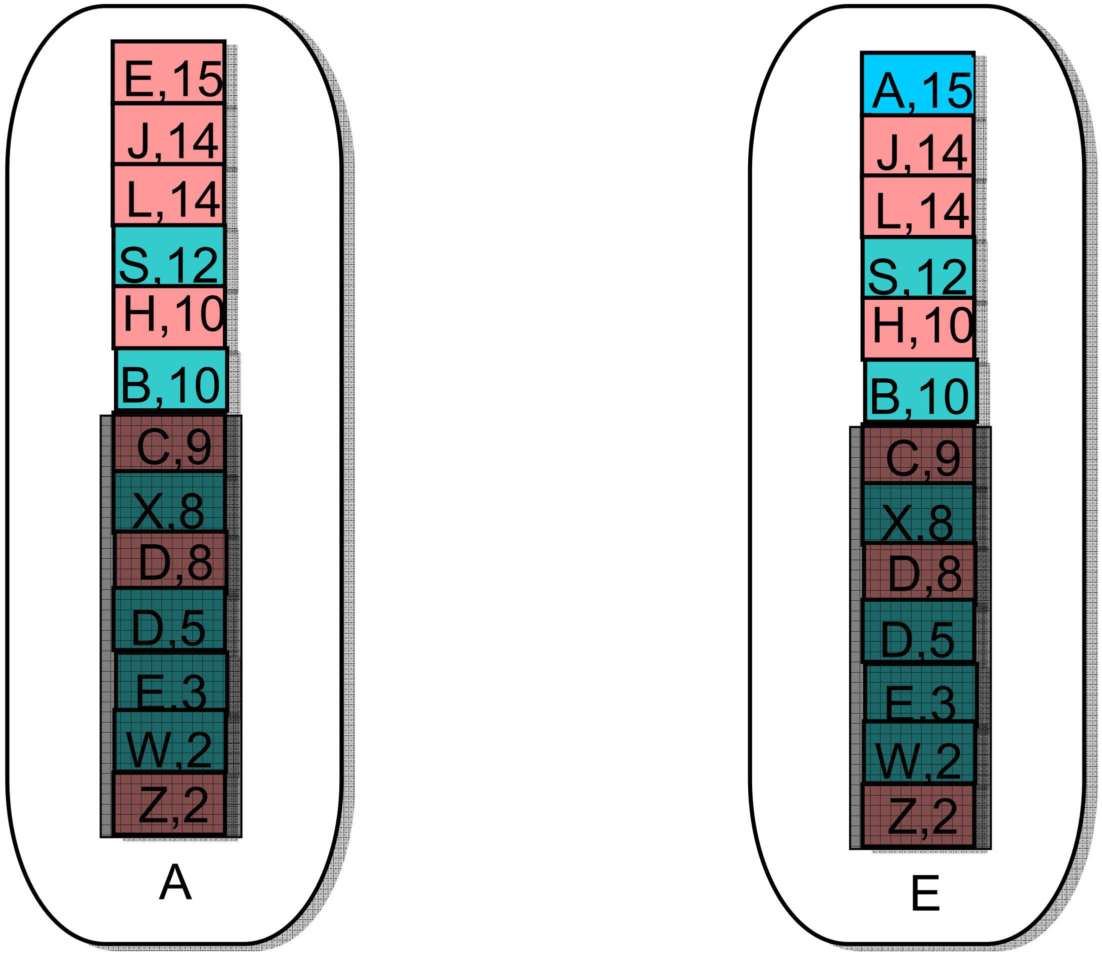
Each node add its fresh descriptor to its view (suppose t=15)

Newscast Example



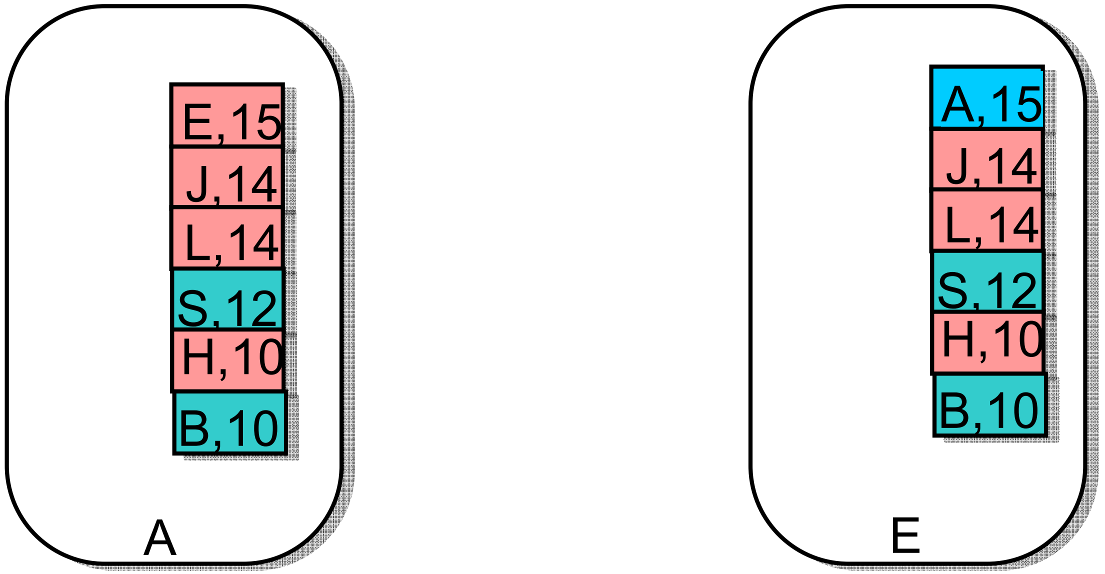
Views are merged

Newscast Example



Oldest descriptors discharged (suppose c=6)

Newscast Example

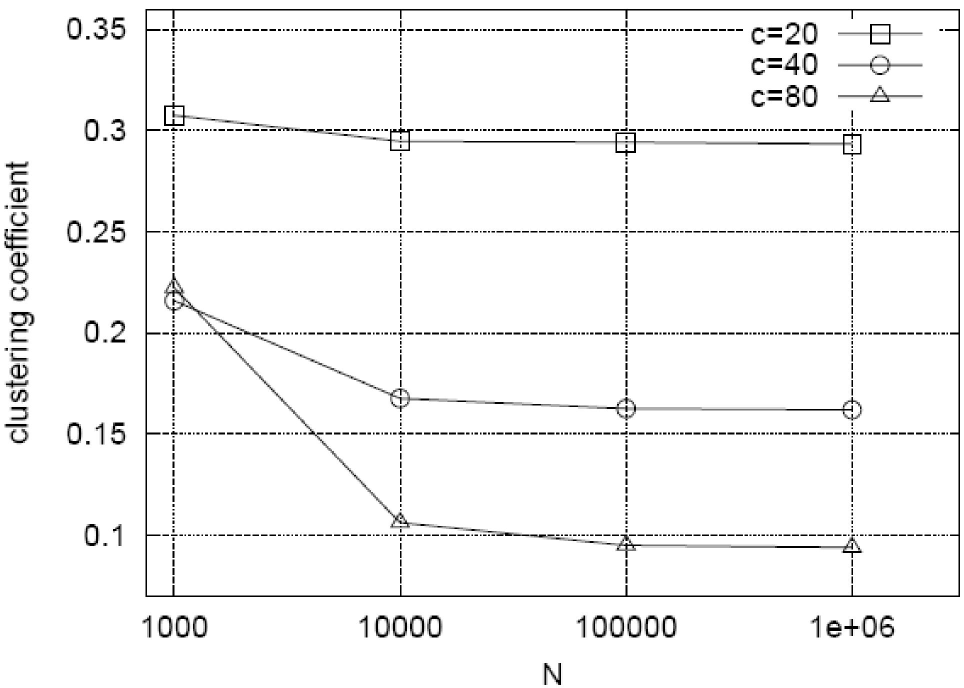


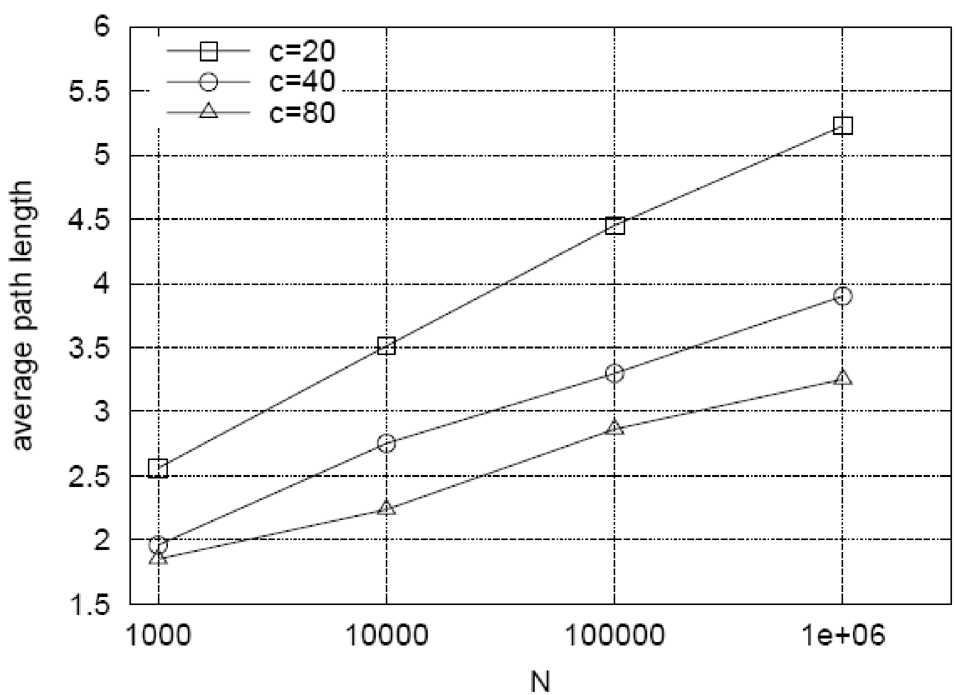
Newscast Resulting Topology

- Continuously changing
- Random-like
  - Low clustering coefficient
  - Low average path length
    - Logarithmic growth with respect to network size
- Robust to node removal
  - Dinamicity
  - Aging of older and not refreshed descriptors

Newscast Resulting Topology

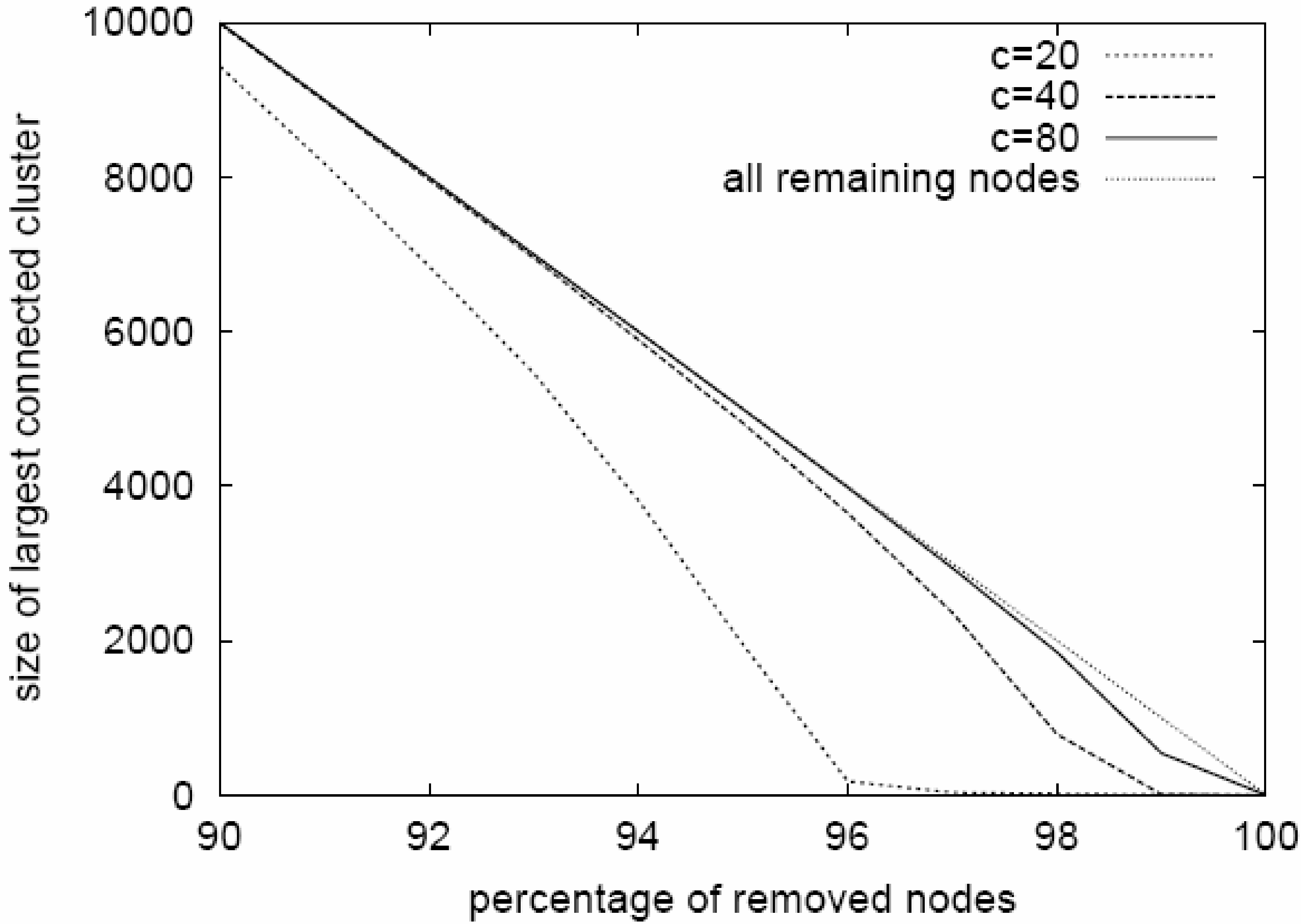
- Clustering coefficient (CC)
  - Proportion between links between neighbors of a node and the number of possible links
- Network CC
  - Average of nodes' CC





- Average Path Length (APL)
  - Average hop-distance between any pair of nodes

Newscast Robustness



Aggregation

Collective Computation (Aggregation)

- Each node has a (numeric) local state
- Compute (global) aggregate function over the initial values at *all* nodes
- The aggregate value has to be known by each node
- Examples of aggregate functions:
  - Average
  - Min-max
  - Geometric mean
  - Variance
  - Network size

Biological Inspiration: Epidemic Spreading

- Each node periodically selects another (random) peer and exchanges local state information
  - Random peer sampling provided by Newscast
- Each node updates its local state based on the information exchanged
- System fully symmetric — all nodes act identically
- Communication is symmetric — “push-pull” epidemic

Aggregation Through Epidemics (Gossip)

- Local value S_p contains current estimate of the aggregate
- Suppose the (random) peer picked by node p is q
- Nodes p and q exchange current estimates
- Update local estimates depending on aggregation function
  - Average: $S_p \leftarrow \frac{(S_p + S_q)}{2}$
  - Geometric mean: $S_p \leftarrow \sqrt{(S_p S_q)}$
  - Maximum: $S_p \leftarrow max(S_p, S_q)$
- Other, more complex functions can be built combining elementary functions

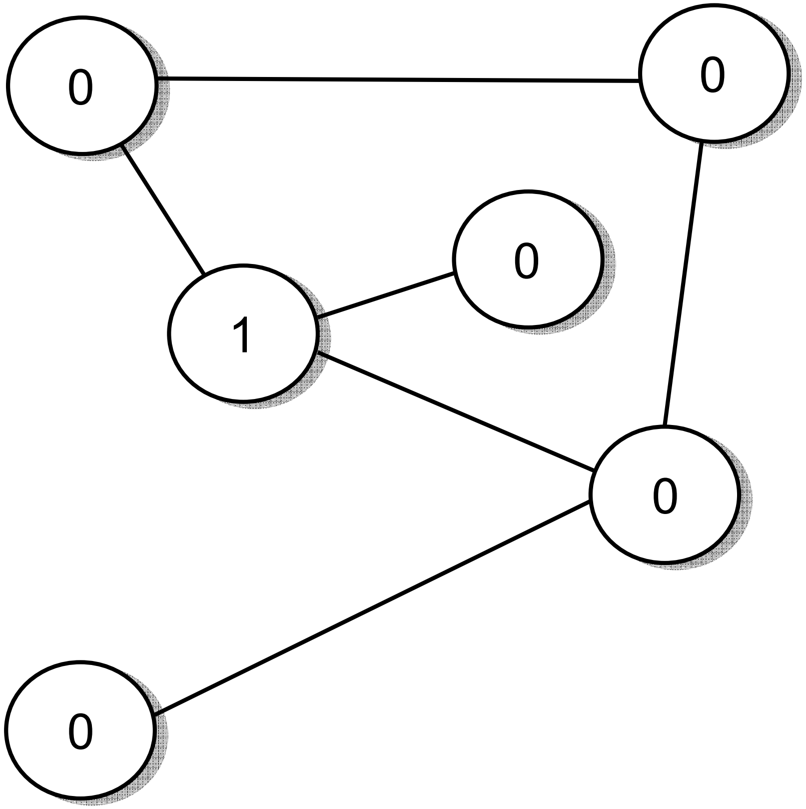
Properties of Epidemic-Based Aggregation

- If peers are sampled randomly sampled, then the variance of nodes' estimates decreases exponentially
  - Aggregation goal: variance=0
    - Same estimated value throughout the network
- Global random peer sampling provided by Newscast!

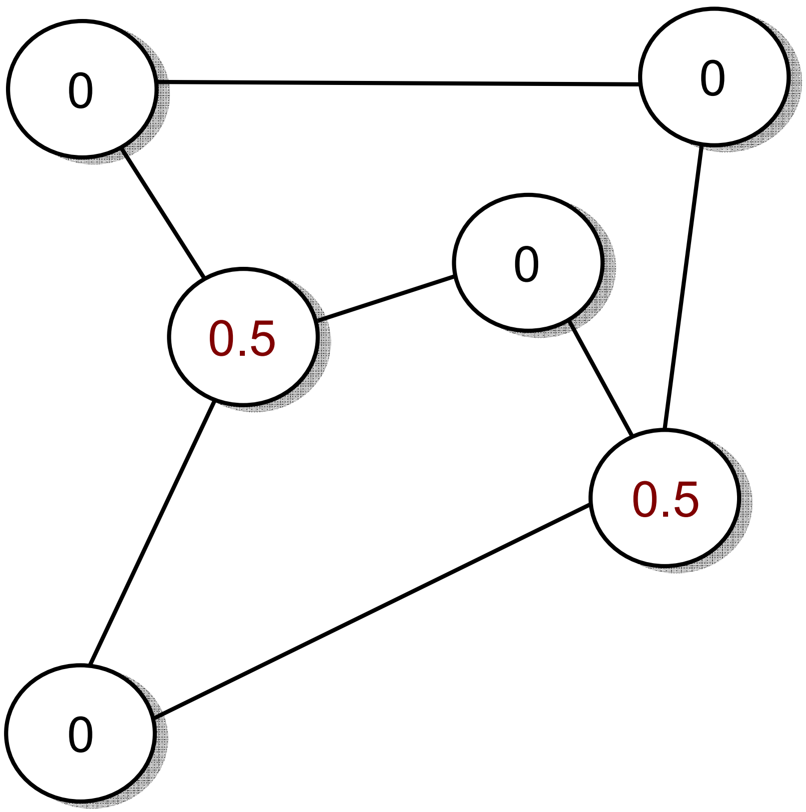
Aggregation Example: Network size

- Through averaging it is possible to evaluate network size
- Aggregation initialized with one node at value 1 and every other nodes at value 0
- Average aggregation performed
- Each nodes reaches value 1/N
  - Network size is simply 1/(1/N)

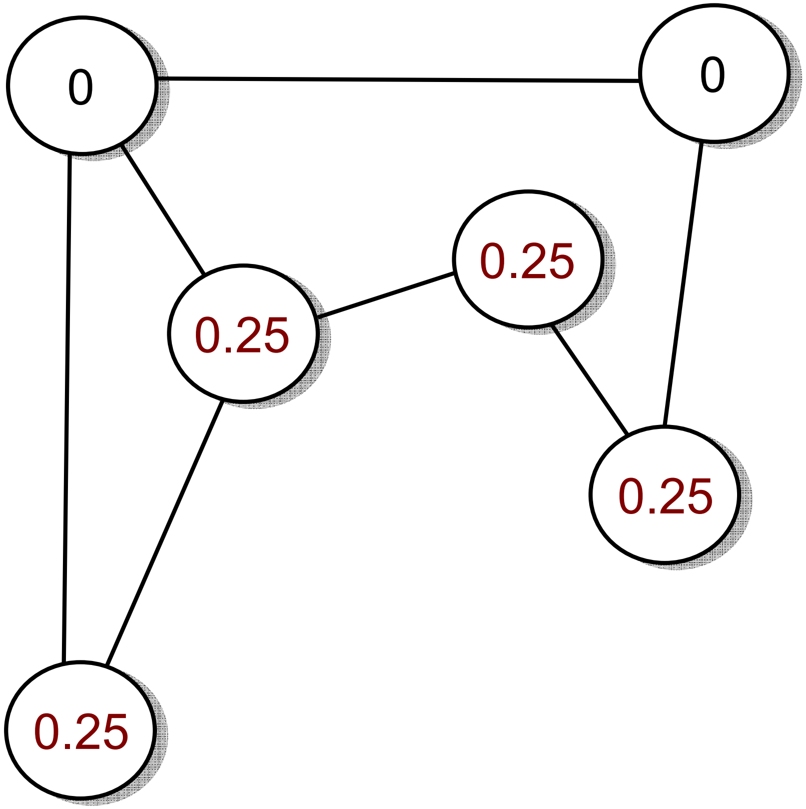
Network Size Estimation Using Average



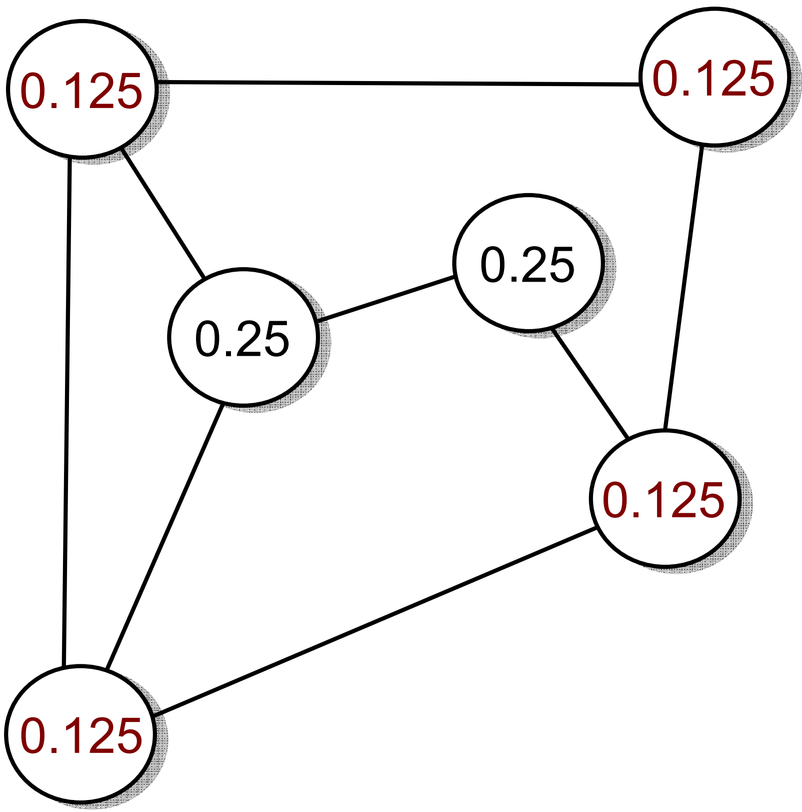
Network Size Estimation Using Average



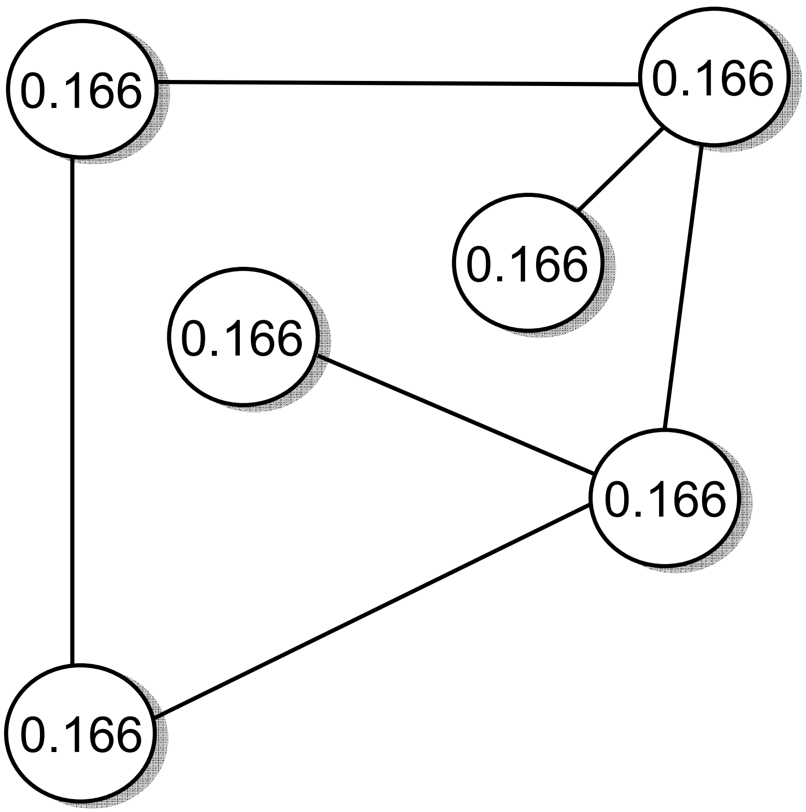
Network Size Estimation Using Average



Network Size Estimation Using Average

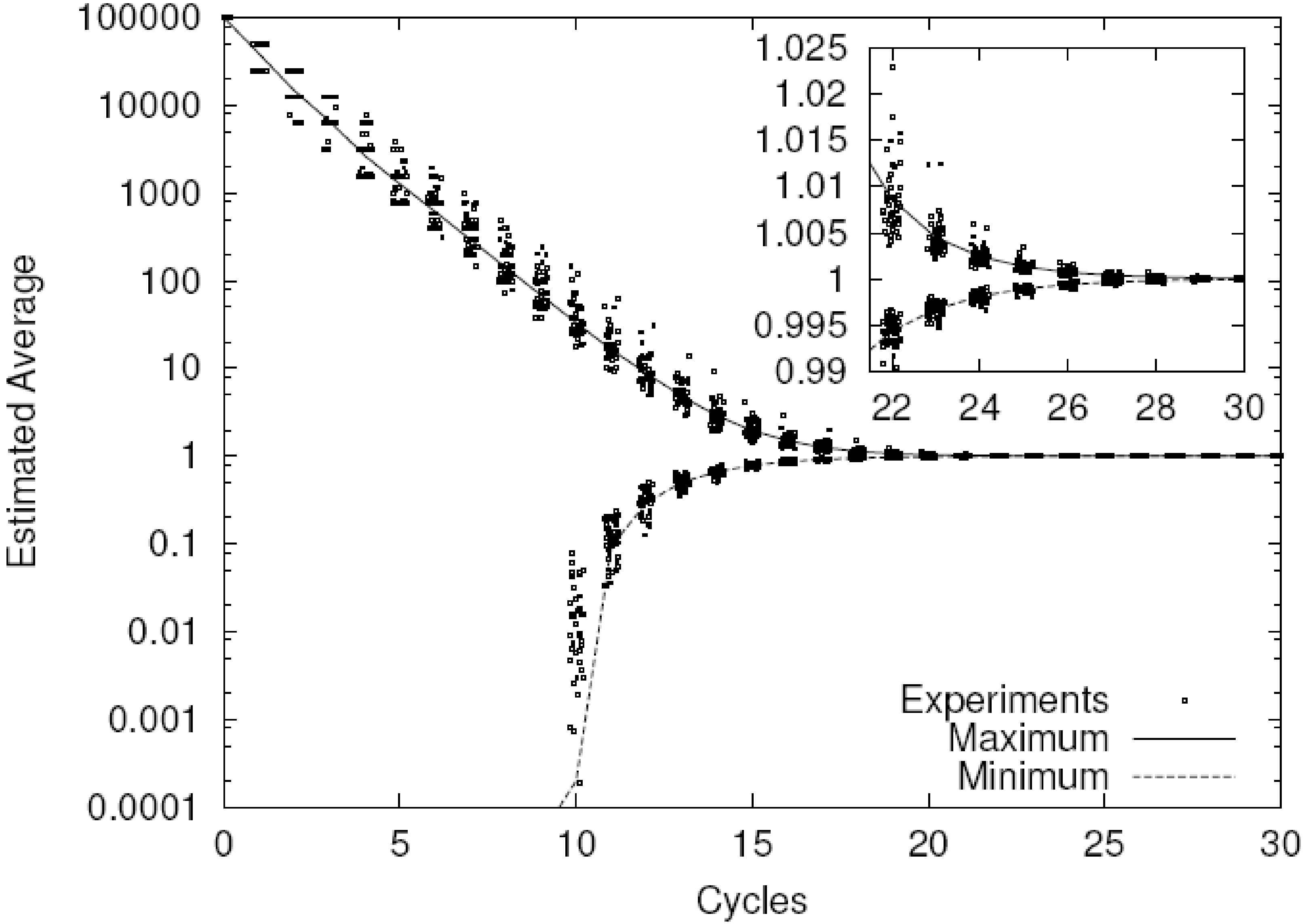


Network Size Estimation Using Average

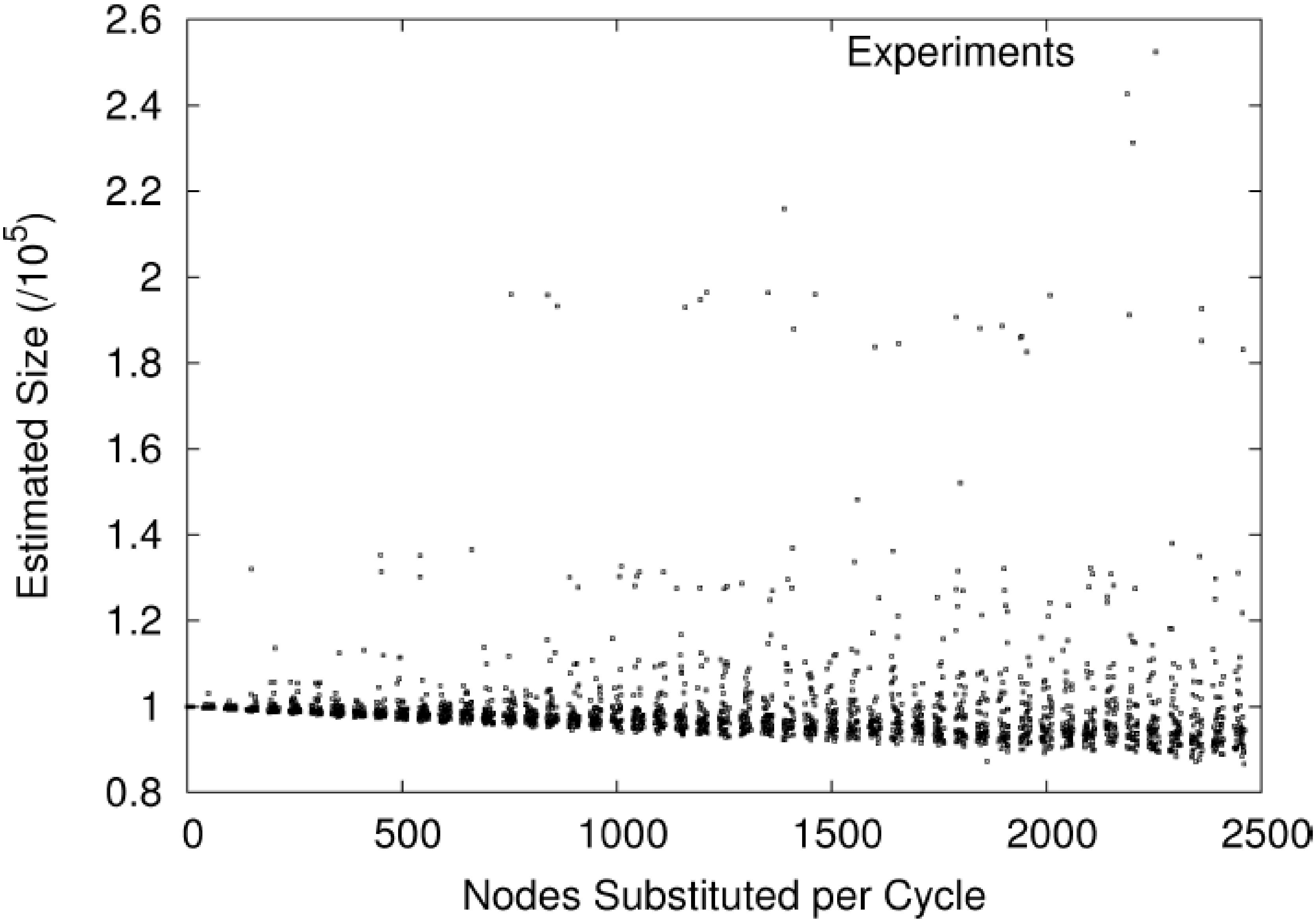


1/0.166 = 6.02 ≈ N

Exponential Convergence of Average



Robustness of Average



Bibliography

- BISON project homepage: http://www.cs.unibo.it/bison/
- M. Jelasity, W. Kowalczyk, and M. van Steen. Newscast computing. Technical Report IR-CS-006, Vrije Universiteit Amsterdam, Department of Computer Science, Amsterdam, The Netherlands, November 2003.
- Márk Jelasity, Alberto Montresor, and Ozalp Babaoglu. Gossip-based aggregation in large dynamic networks. *ACM Trans. Comput. Syst.*, 23(1):219-252, 2005.