Topology Affects the Efficiency of Network Coding in Peer-to-Peer Networks

张志明

2010/11/24

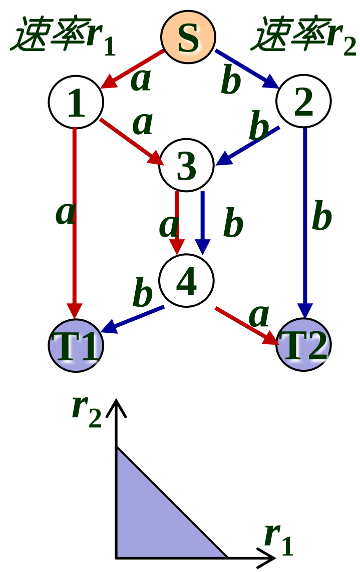
Outline

- Introduction
- The problem of linearly dependent blocks
- Topology effects on the efficiency of network coding
- Conclusion and discussion

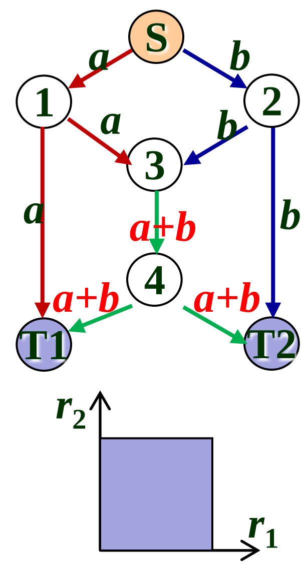
网络编码概述

- 网络编码（Network Coding，简称NC）（ Rudolf Ahlswede 等，于2000年提出）：
  - 定义：网络中的节点除了具有存储转发功能外，还可以对收到的信息进行处理；
  - 好处：可以使组播达到最小割最大流定理所确定的理论最大吞吐量；

一般组播



基于网络编码的组播



Application

Network coding can

- Achieve better network throughput
- Offer better scalability and superb resilience to peer failures and departures
- Shorten downloading times

The costs and trade-offs

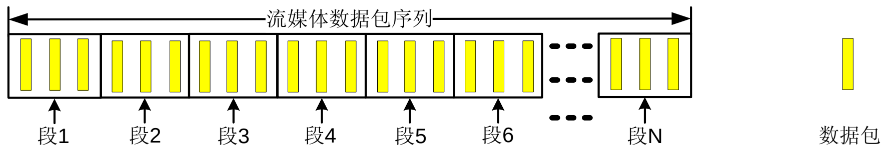
- Contray to claim from previous work, it is very likely to receive linearly dependent blocks when peers code outgoing blocks before they fully decode and recover original blocks in realistic P2P topologies.

Outline

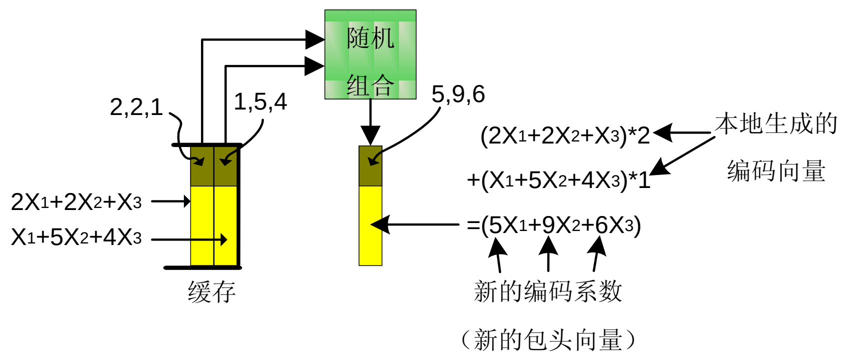
- Introduction
- The problem of linearly dependent blocks
- Topology effects on the efficiency of network coding
- Conclusion and discussion

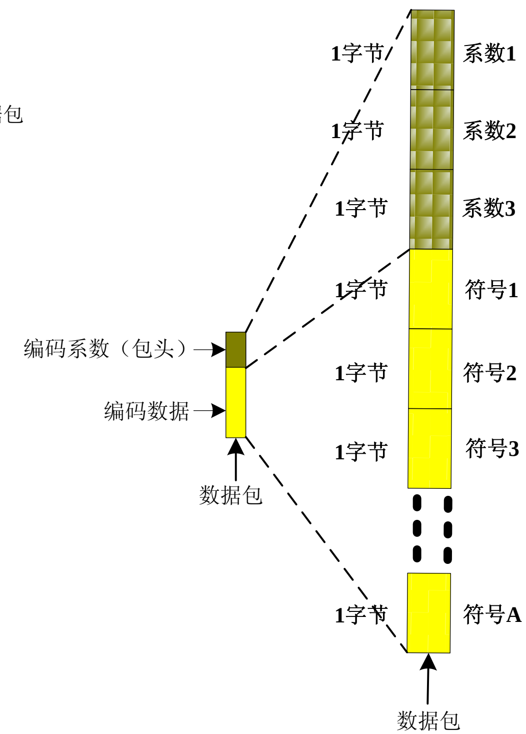
随机线性网络编码--编解码示例

分段：



编码：





解码： $X = A^{-1}Z$

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} = \begin{bmatrix} 5 & 9 & 6 \\ 7 & 10 & 8 \\ 3 & 4 & 3 \end{bmatrix} \quad Z = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$$

Definitions

Density

- Equals to m/n
  - m: The number of data packets used by encoder to encode new one
  - n: The original number of data packets in each segment

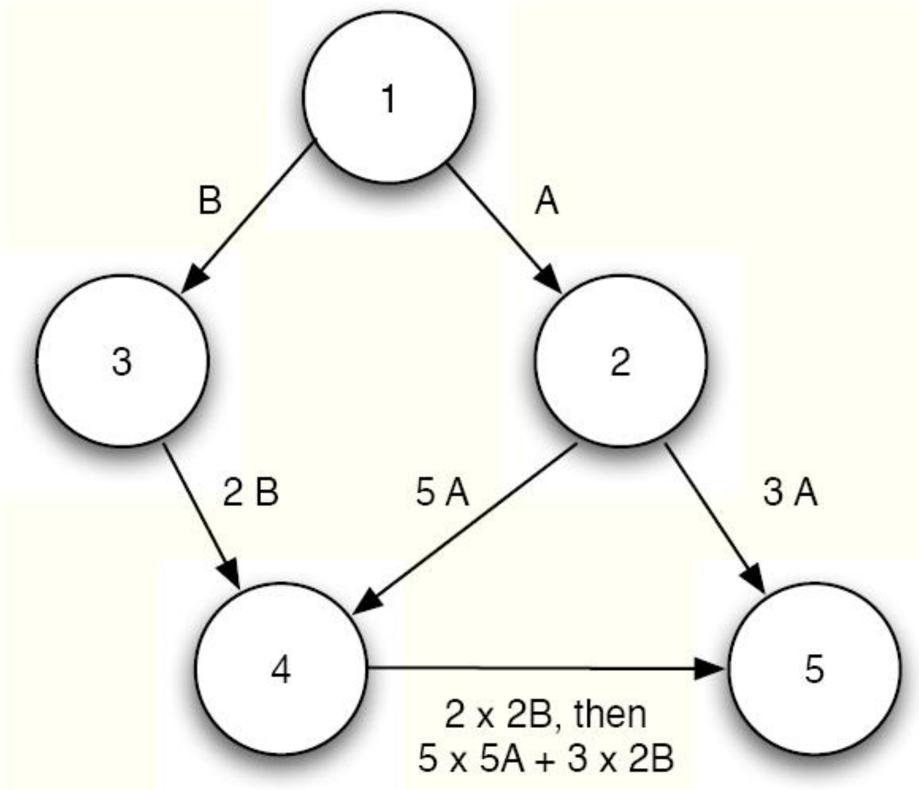
Aggressiveness

- Defined as α
  - α*n: The minimum number of coded data packets a peer need to hold before it can do encoding.

Network Coding Leads to Linearly Dependent Blocks-1

The first example with a small topology

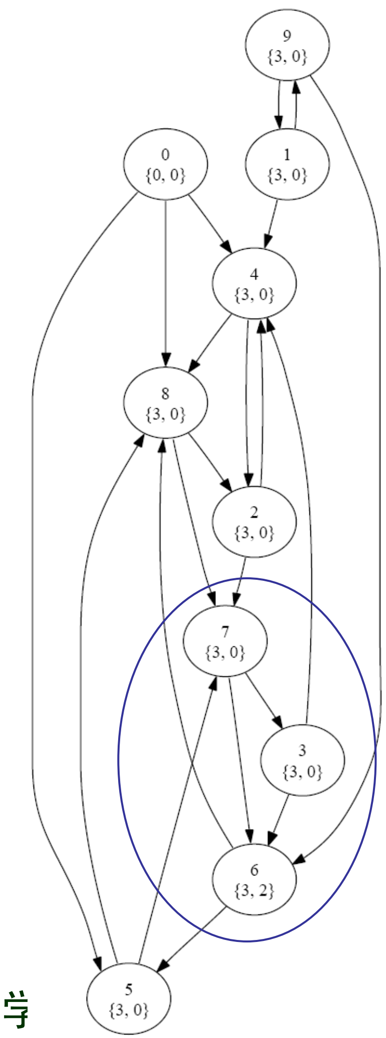
- Peers may easily receive linearly dependent blocks when aggressiveness is less than 1, such as node 5 in this topology.
- Aggressiveness equals to 1/n in this example.

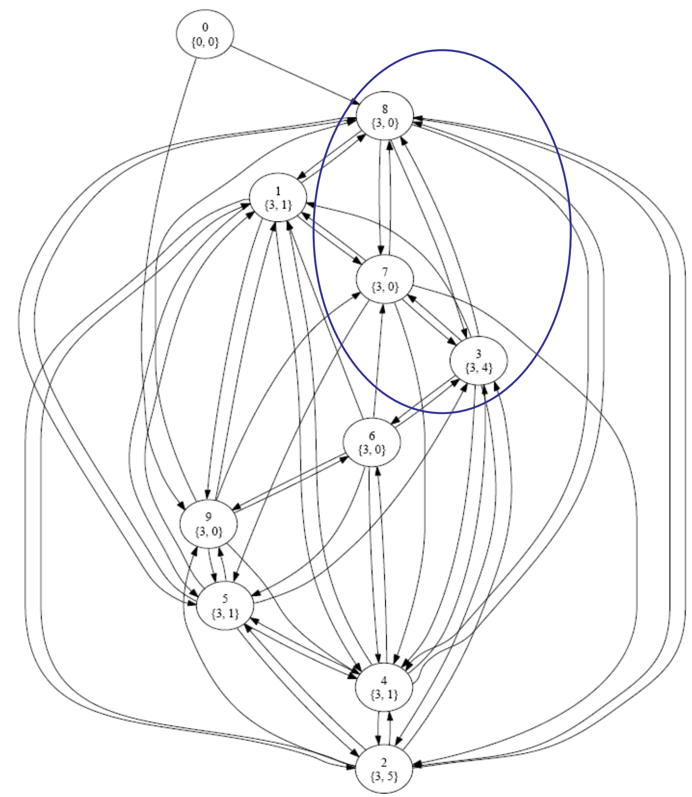


Network Coding Leads to Linearly Dependent Blocks-2

The second example with random topologies

- The same problem exists in larger topologies.
- Node 6 in left topology
- Nodes 3, 2, 4, 5 in right topology





Outline

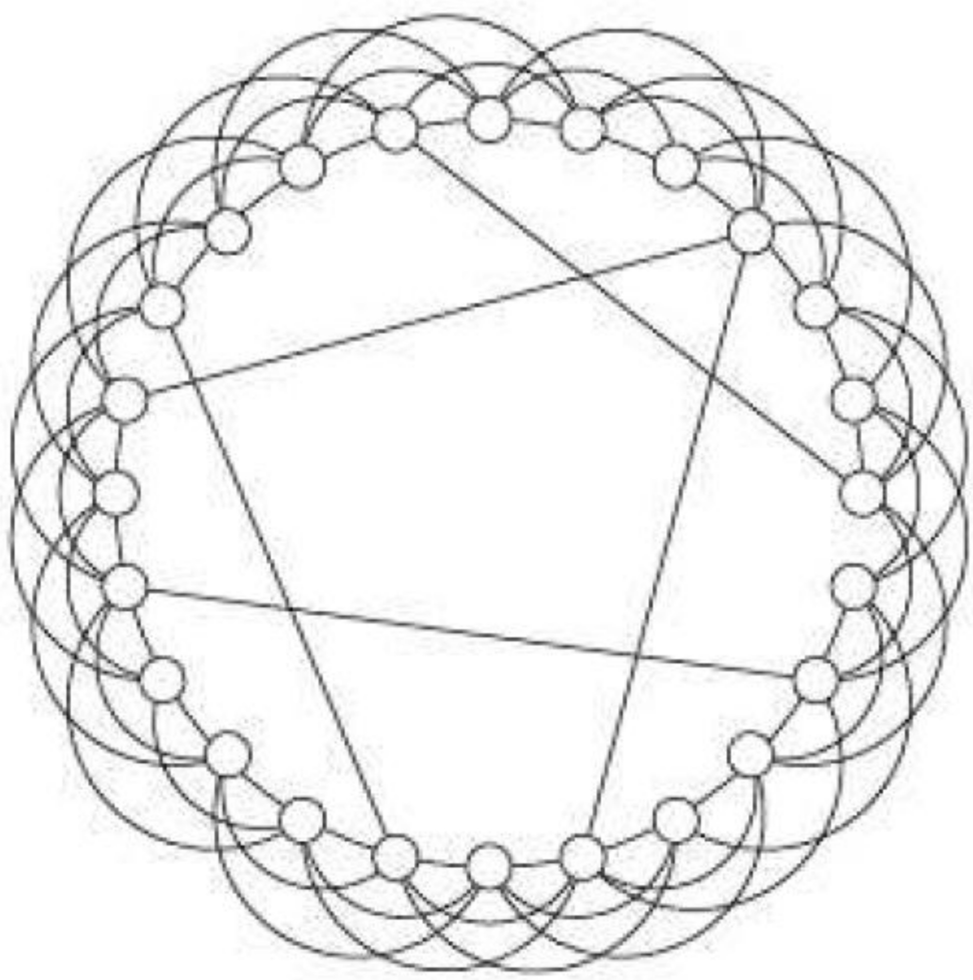
- Introduction
- The problem of linearly dependent blocks
- Topology effects on the efficiency of network coding
  - Randomness and network sparsity
- Conclusion and discussion

Definitions

- Block redundancy
  - The quotient of the number of coded blocks a peer receives and n (The number needed to successfully decode the segment).
- Distribution time
  - The time interval from initial forwarding of a block from the server to any of its downstream peers until all peers in the network have successfully received n independent coded blocks.
- Sever cost
  - The number of blocks forwarded from the server to any of its downstream peers.

Small-world Topology With Some Rewired Links

- The network topology randomness is varied by adjusting the rewiring probability *p* in small-world topologies
- *P*=0: Completely regular graph
- *P*=1: Random graph

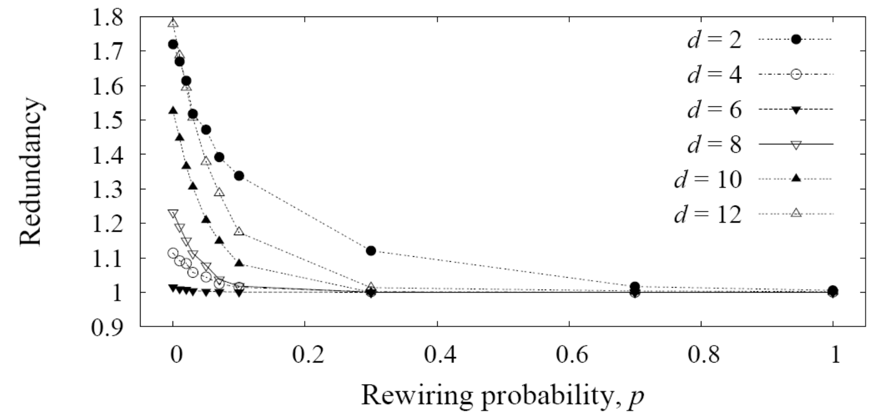


Small-world Topology With Some Rewired Links

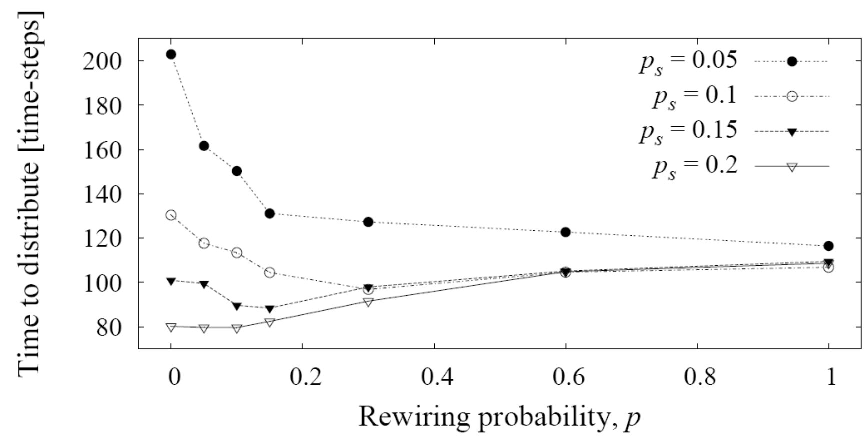
- Random graphs have low clustering and it is likely that any two peers will have short path length between them.
- Regular graphs are likely to have long paths between peers and significant clustering.

Performance Impact of Randomness and Sparsity

- P=0: regular graph
  - sharing neighbors leads to high redundancy.
  - Long path results in long distribution time.
- d: the degree of peers
  - Too few neighbors leads to infrequent new information, resulting in high redundancy
  - Too many nbrs -> Many common downstream peers-> High redundancy
- Ps: percentage of peers directly connected to server
  - Bigger Ps ->shorter distribution time



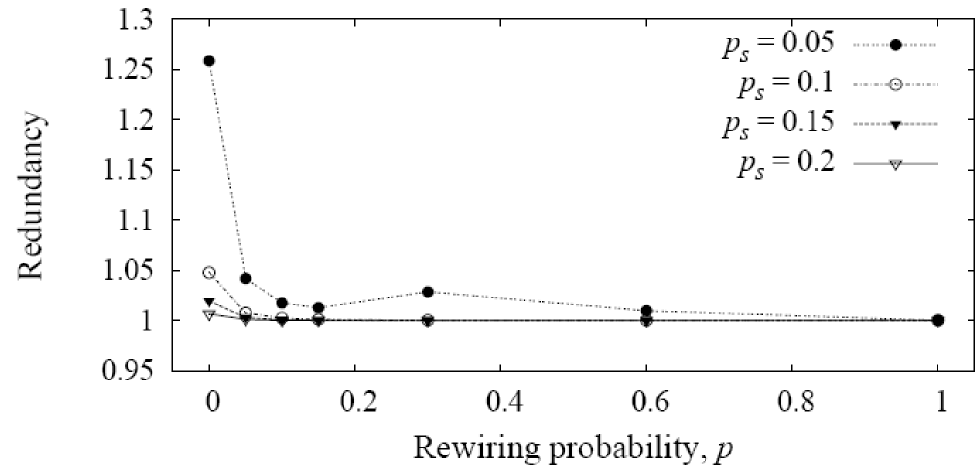
(a) Average redundancy experienced at a peer



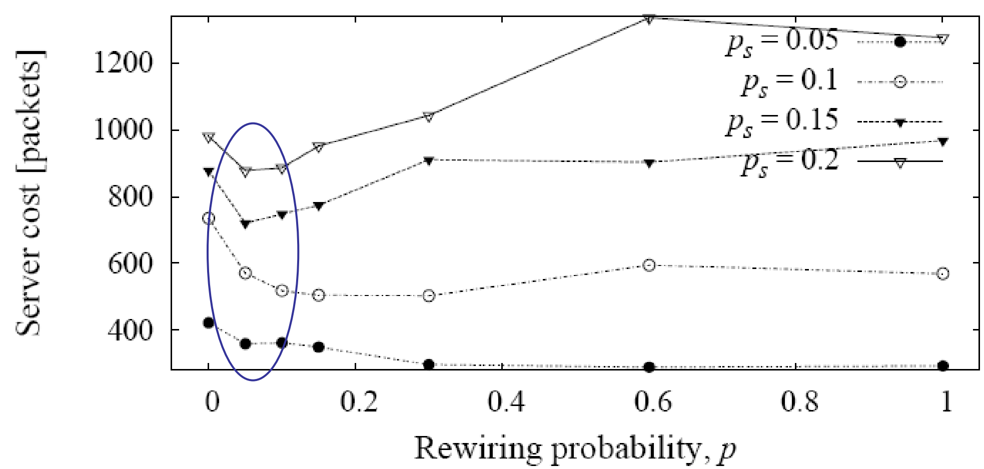
(b) Time to complete block forwarding

Performance Impact of Randomness and Sparsity

- Server cost
  - Bigger Ps ->lower redundancy
  - Smaller Ps -> lower server cost
- We see a minimum in the server cost when redundancy is low and the distribution time is low.



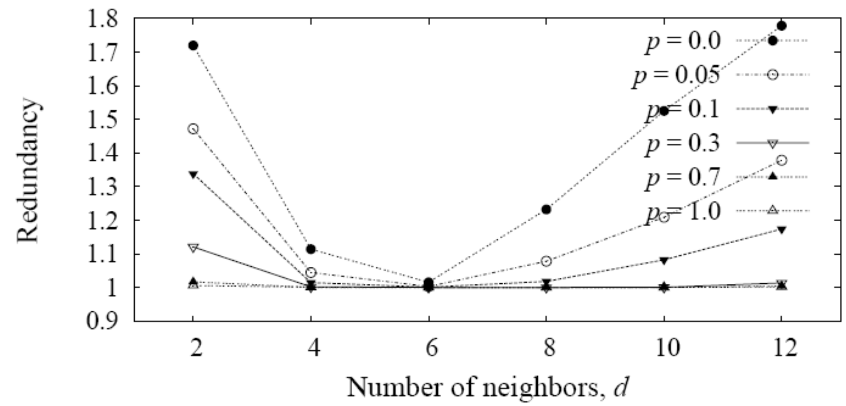
(a) Average redundancy experienced at a peer



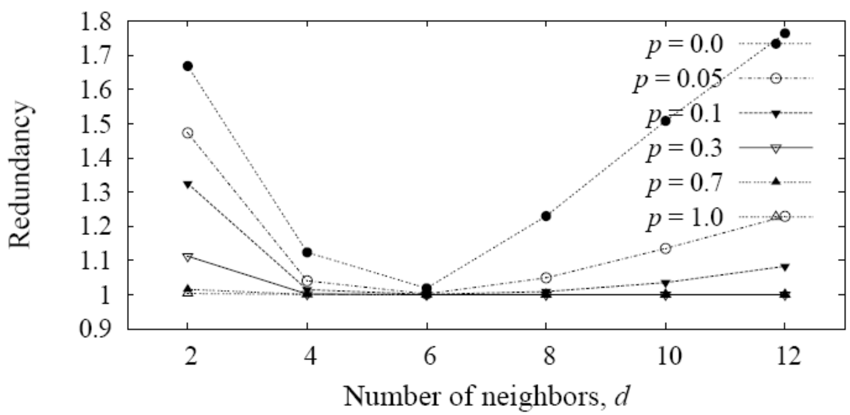
(c) Cost to the server

Impact of Network Size-1

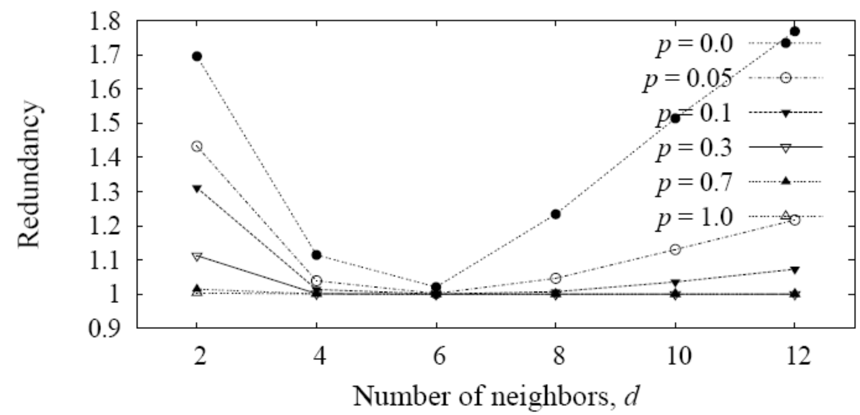
- It is not the global connectivity that has the most significant impact on network coding redundancy. It is instead the local connectivity (the peer degree)
- The best d=6



(a) Network of 100 peers



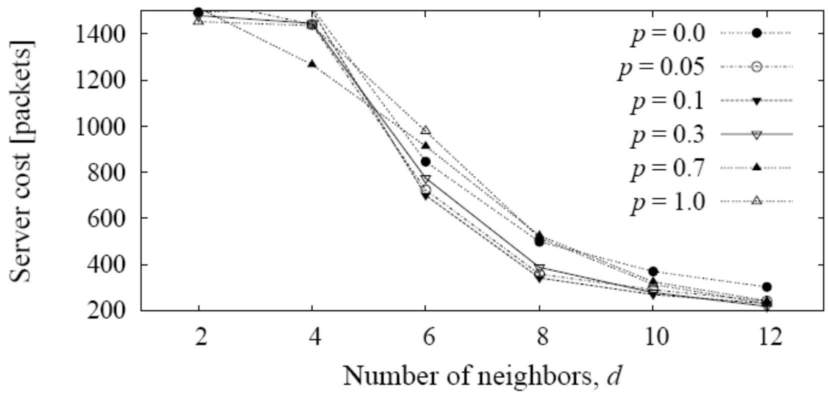
(b) Network of 500 peers



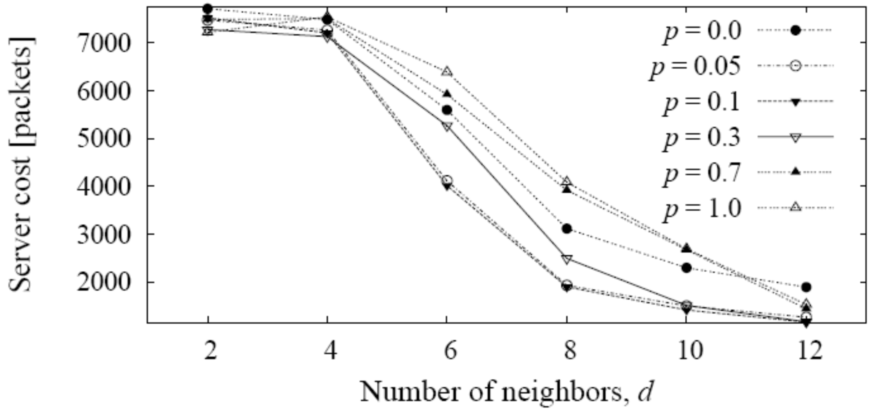
(c) Network of 1000 peers

Impact of Network Size-2

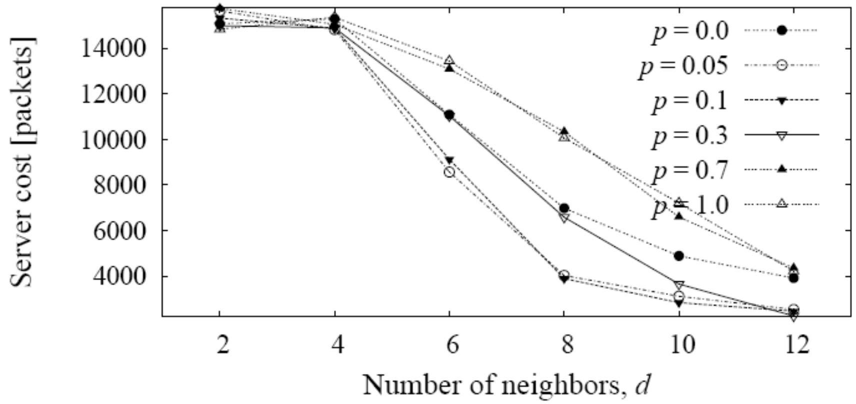
- Regular topologies with a few long-distance links are preferred.
- The distribution time increases sublinearly as network size increases.



(a) Network of 100 peers



(b) Network of 500 peers



(c) Network of 1000 peers

Outline

- Introduction
- The problem of linearly dependent blocks
- Topology effects on the efficiency of network coding
- Conclusion and discussion

Conclusion and Discussion

- The P2P topologies offering the best overall performance are small-world topologies with low rewiring probability (around p = 0.1) with peer degree of six.
  - Peers have a sufficient number of neighbors for effective distribution, without too many peers sharing downstream neighbors.
    - Low redundancy and lower distribution times
  - Messaging overhead is only needed for a small fraction of the links
    - Better overhead
- Small-world networks with low rewiring probability p exhibit significantly lower server cost than corresponding networks with more randomness

Conclusion and Discussion

- Changing the rewired links for different segments of the data stream.
  - Better load balance

Questions?