


Advanced message passing techniques for distributed storage

http://www.ncrg.aston.ac.uk/

D. Saad, S. Bounkong, J. van Mourik

Aston University

20 January 2005

Optimization and Inference in Machine Learning and Physics

Outline

- What is distributed storage?
- Coding and segment distribution
- Message passing for optimal segment distribution
- Simulation results
- Future research directions

Distributed storage

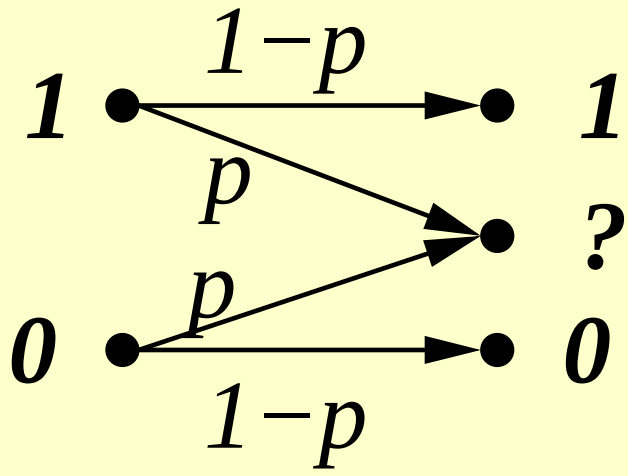
- A system of nodes (computers?) on a graph
- Main problems -
  - Limited storage and computing capability
  - Bottlenecks, speed of download
- Distributing file (segments), retrieval from neighbours
- Available methods
  - Replication
  - Cashing
  - Coding

Distributed storage - advantages

- Decentralised
- Computationally efficient
- Minimises latency
- Fault-tolerant; high availability of data even with some component failure
- Load balancing (both storage and traffic)
- Increased data security against eavesdropping over individual links

Coding - BEC

- The message is a binary vector
- Bits get lost while being transmitted via a Binary Erasure Channel
- Received bits are correct
- There are coding techniques for retrieving lost information, e.g., Low Density Parity Check Codes

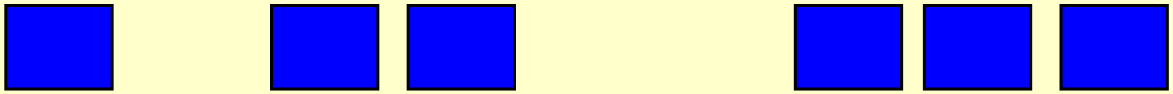


Coding - file segments

- Distributed storage can be viewed as file reconstruction from partial information
- Only part of the original segments are available

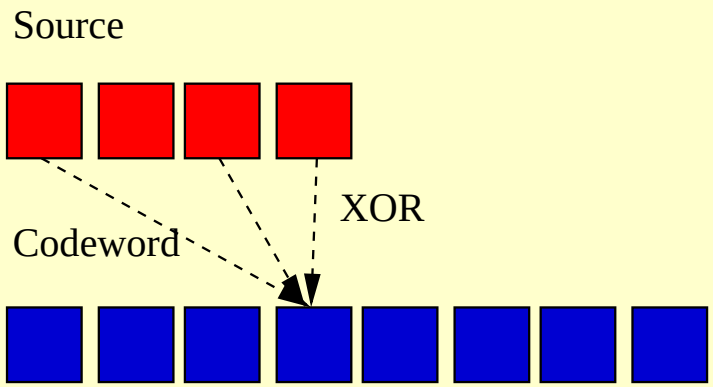
Source

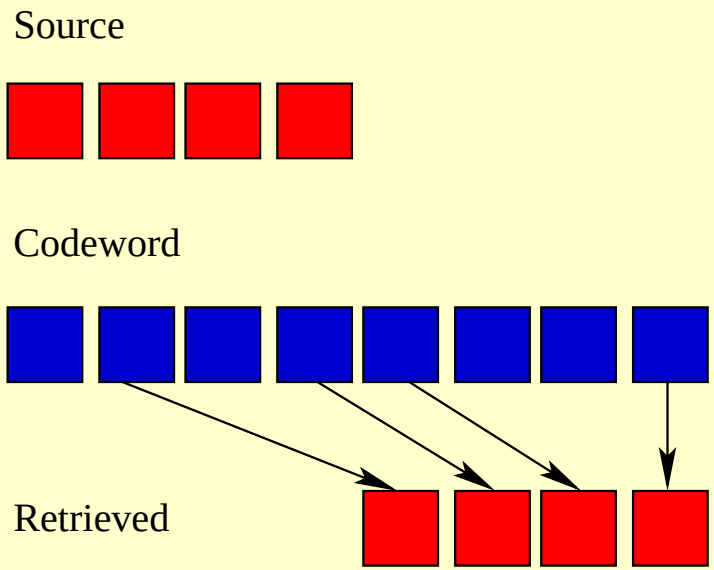
Received



Fountain codes

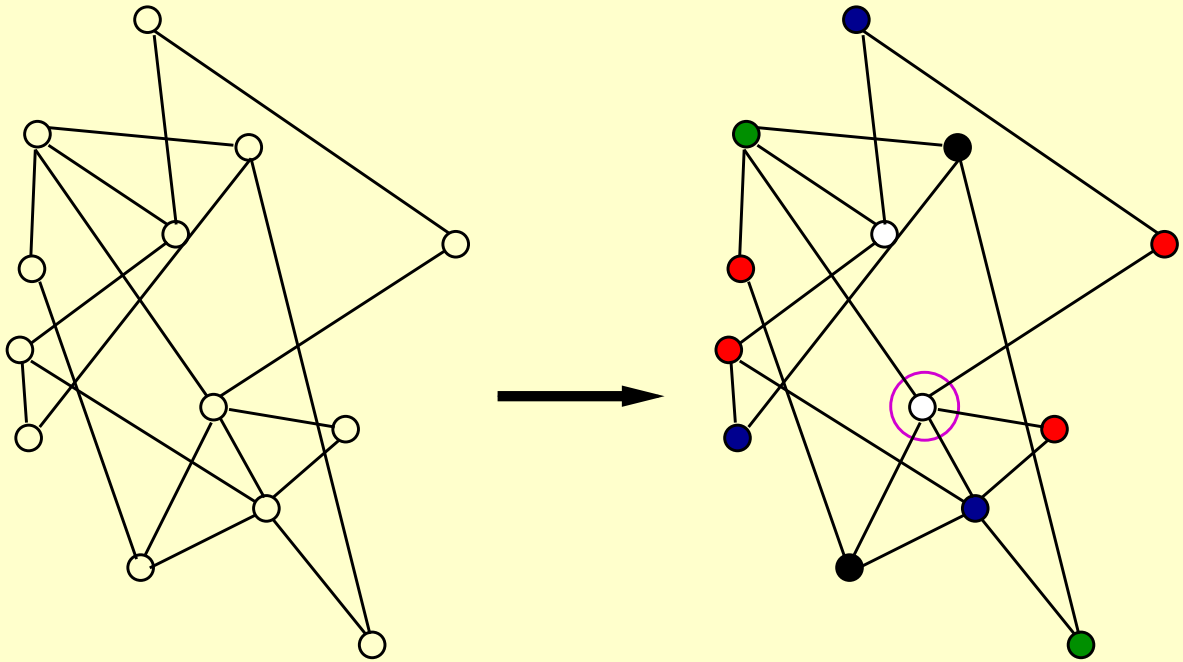
- Best suited for BEC, highly robust
- Generating a fountain of codewords from data packets, by XOR of d packets taken from some distribution of $\rho(d)$
- Decoding by message passing/deterministic





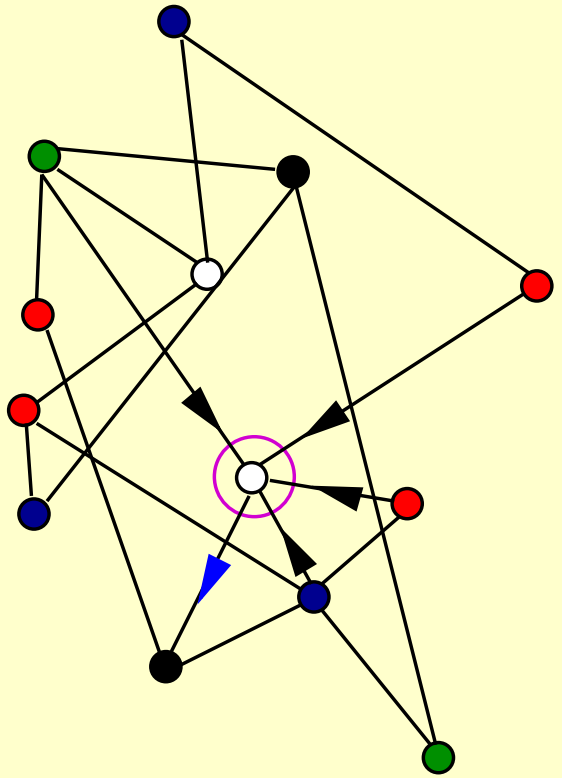
Segment distribution

- Different segments can be viewed as colours
- Main problem - how to maximise the number of different colours in a neighborhood
- File retrieval - when a sufficiently large subset of segments is collected



Message passing

- We use message passing technique for assigning segments to nodes
- Vector messages represent probabilities of having certain colours
- Requires messages from 2nd order neighbours
- Pseudoposteriors are averaged over a time window



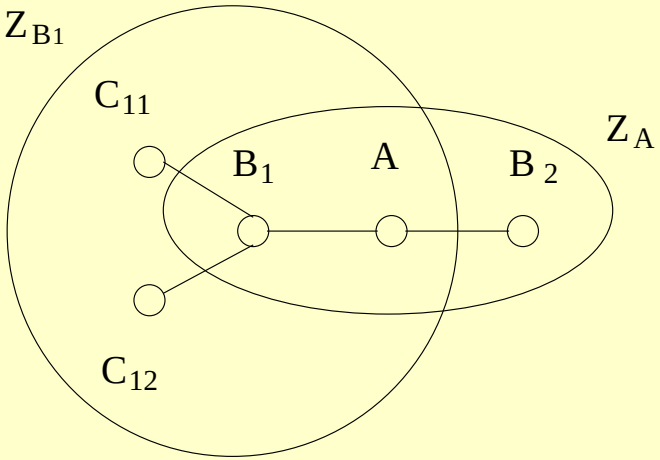
Message passing

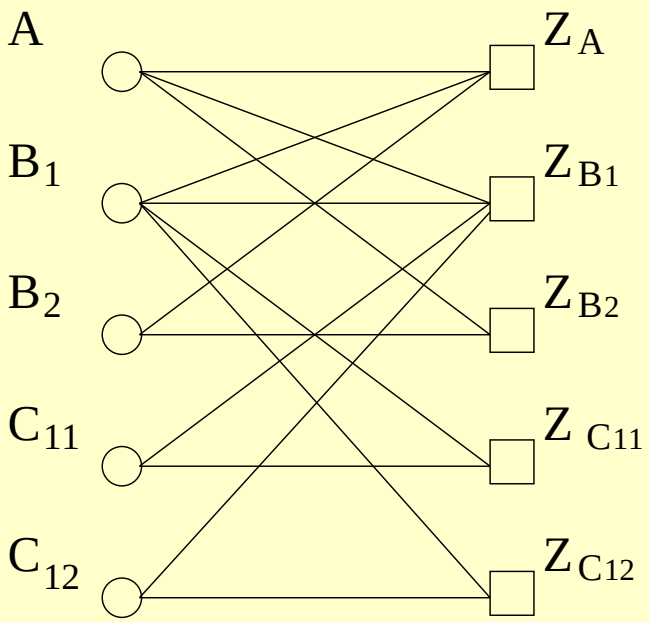
$$P(Z_A|A) = \frac{1}{K_A}\sum_{\{B\}} P(Z_A|A,\{B\})P(\{B\}|\{Z_B\},\{Z_C\})$$

$$= \frac{1}{K_A}\sum_{\{B\}} P(Z_A|A,\{B\})\prod_i P(B_i|Z_{B_i},\{Z_{C_i}\})$$

$$= \frac{1}{K_A}\sum_{\{B\}} P(Z_A|A,\{B\})\prod_i \frac{1}{K_{B_i}}\sum_{\{C\}_i} P(Z_{B_i}|B_i,\{C\}_i)P(B_i)\prod_j P(C_{ij})$$

$$P(A|\{Z_B\}) = \frac{1}{K_Z}\prod_{\{Z_B\}} P(Z_B|A)\ P(A)$$



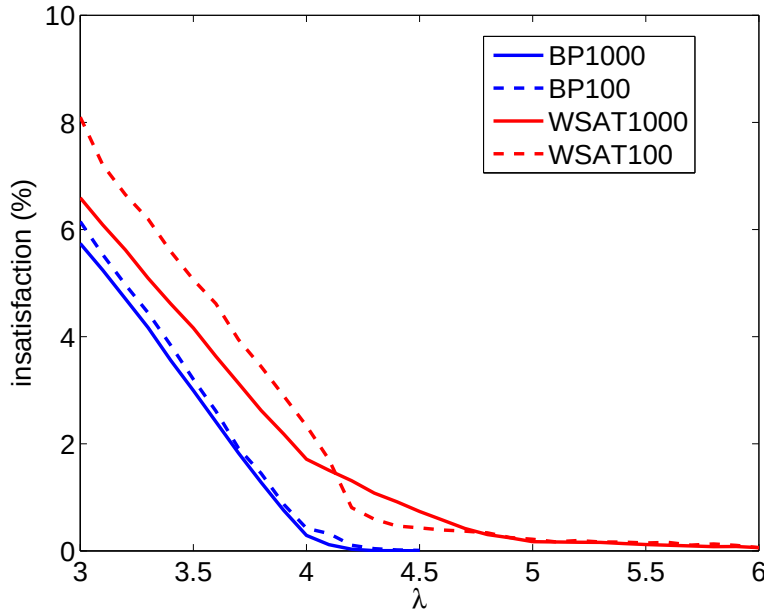


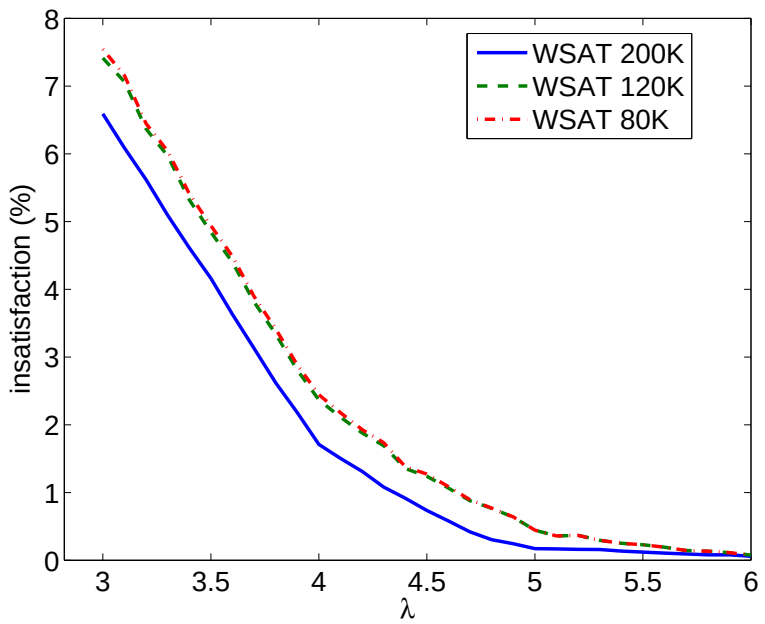
Experiments

- Graph structure
  - Poisson connectivity (limited)
  - Linear connectivity - between two integer values
- Network size - 100, 1000 nodes
- Number of segments - 4
- Averaged over - 100 graphs
- Comparison with WalkSAT -
  - GSAT/random with probability 1/2
  - Very high limit for the number of search steps

Results - BP vs. WSAT

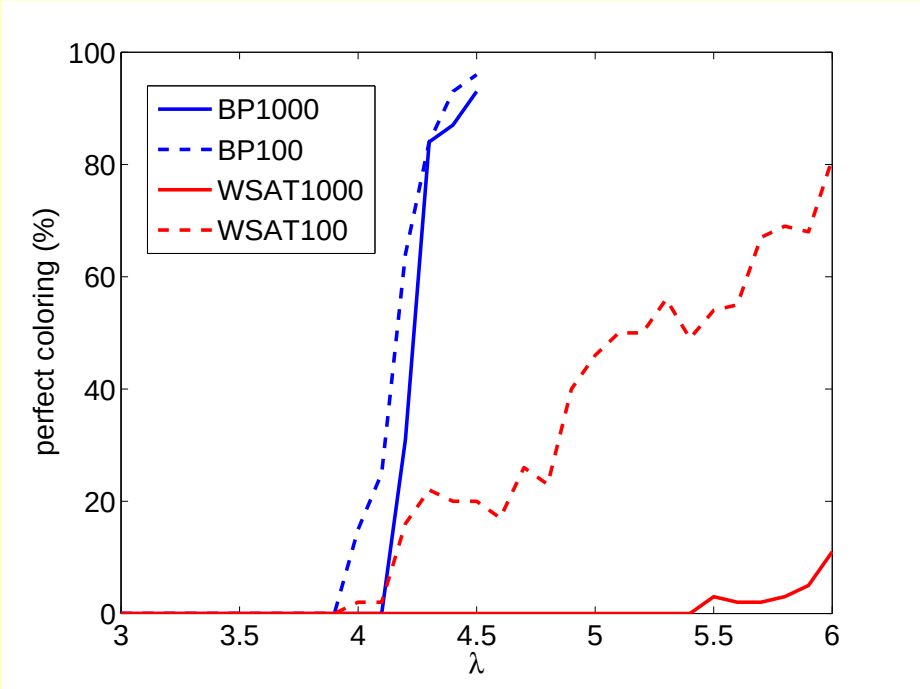
For the case of linear connectivity we get





Connectivity transition point

One can also obtain the transition point in terms of connectivity



Summary & future research

- We use a message passing algorithm for distributed storage on a graph
- The suggested variant of belief propagation gives superior results to other approaches on the graphs studied
- Future Research:
  - Multiple segments per node
  - Multiple hops
  - Dynamical allocation of segment