BATS Codes: Achieving the Capacity of Networks with Packet Loss

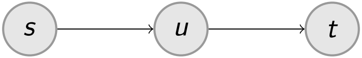
Shenghao Yang

Institute for Interdisplinary Information Sciences
Tsinghua University

March 6, 2014, Sino-German Workshop

Joint work with Raymond W. Yeung (INC, CUHK)

Achievable Rates

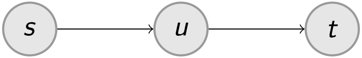


Both links have a packet loss rate 0.2.

The capacity of this network is 0.8.

Intermediate	End-to-End	Maximum Rate
forwarding	retransmission	0.64
forwarding	fountain codes	0.64

Achievable Rates

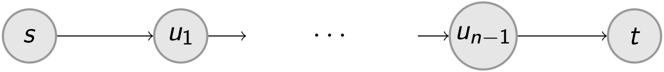


Both links have a packet loss rate 0.2.

The capacity of this network is 0.8.

Intermediate	End-to-End	Maximum Rate
forwarding	retransmission	0.64
forwarding	fountain codes	0.64
network coding	random linear codes	0.8

Achievable Rates: n-hop Line Network

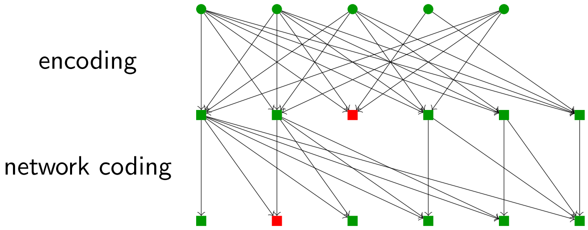


All links have a packet loss rate 0.2.

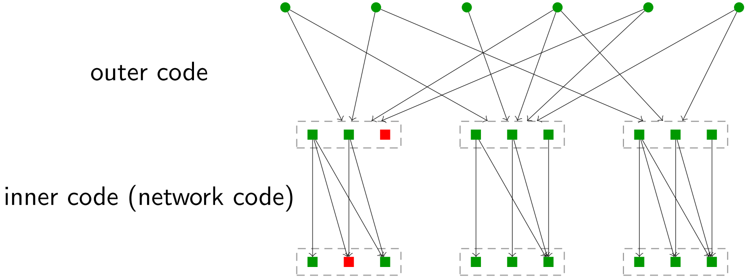
Intermediate Operation	Maximum Rate
forwarding	$0.8^n \to 0,\ n \to \infty$
network coding	0.8

Complexity of Linear Network Coding

- T: length of a packet; K: number of packets
- Encoding: $\mathcal{O}(TK)$ per packet.
- Decoding: $\mathcal{O}(K^2 + TK)$ per packet.
- Network coding: $\mathcal{O}(TK)$ per packet. Buffer K packets.



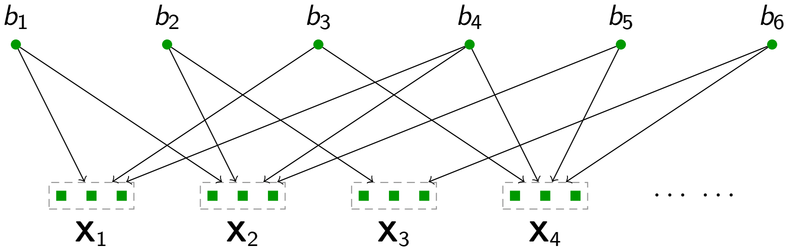
Batched Sparse (BATS) Codes



[YY11] S. Yang and R. W. Yeung. Coding for a network coded fountain. ISIT 2011, Saint Petersburg, Russia, 2011.

Encoding of BATS Code: Outer Code

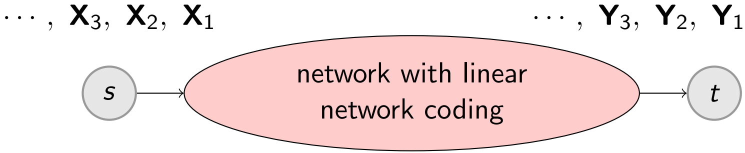
- Apply a "matrix fountain code" at the source node:
  1. Obtain a degree d by sampling a degree distribution Ψ.
  2. Pick d distinct input packets randomly.
  3. Generate a batch of M coded packets using the d packets.
- Transmit the batches sequentially.



$$\mathbf{X}_i = \begin{bmatrix} b_{i1} & b_{i2} & \cdots & b_{id_i} \end{bmatrix} \mathbf{G}_i = \mathbf{B}_i \mathbf{G}_i.$$

Encoding of BATS Code: Inner Code

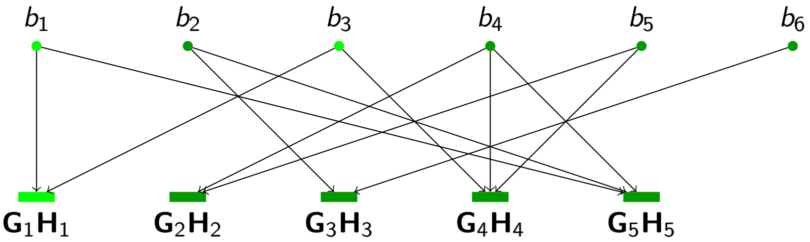
- The batches traverse the network.
- Encoding at the intermediate nodes forms the inner code.
- Linear network coding is applied in a causal manner within a batch.



$$\mathbf{Y}_i = \mathbf{X}_i \mathbf{H}_i, \; i = 1, 2, \ldots.$$

Belief Propagation Decoding

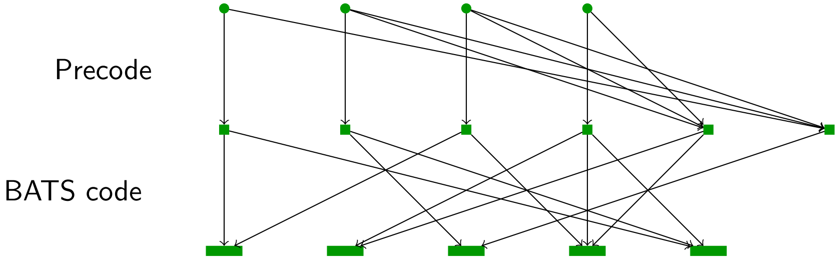
1. Find a check node i with $\text{degree}_i = \text{rank}(\mathbf{G}_i\mathbf{H}_i)$.
2. Decode the ith batch.
3. Update the decoding graph. Repeat 1).



The linear equation associated with a check node: $\mathbf{Y}_i = \mathbf{B}_i\mathbf{G}_i\mathbf{H}_i$.

Precoding

- Precoding by a fixed-rate erasure correction code.
- The BATS code recovers $(1 - \eta)$ of its input packets.



[Shokr06] A. Shokrollahi, Raptor codes, IEEE Trans. Inform. Theory, vol. 52, no. 6, pp. 25512567, Jun. 2006.

Complexity of Sequential Scheduling

Source node encoding		$\mathcal{O}(TM)$ per packet
Destination node decoding		$\mathcal{O}(M^2 + TM)$ per packet
Intermediate Node	buffer	$\mathcal{O}(TM)$
	network coding	$\mathcal{O}(TM)$ per packet

T: length of a packet

K: number of packets

M: batch size

Finite-Length Design

- For given a rank distribution of transfer matrices, BP decoding is asymptotically error free for BATS codes with very small coding overheads.
- For *short* block length, e.g., several thousands packets, the coding overhead for BP decoding can be large.

Finite-Length Design

- For given a rank distribution of transfer matrices, BP decoding is asymptotically error free for BATS codes with very small coding overheads.
- For *short* block length, e.g., several thousands packets, the coding overhead for BP decoding can be large.
- *Inactivation decoding* can be used to reduce the overhead when the block length is within several 10 thousands packets:
  - BP decoding stops when there is no decodable batch
  - Assign an undecoded input packet as decoded (inactive) to restart BP decoding
  - Alleviates the degree distribution optimization problem for files up to several ten megabytes.

Inactivation Decoding

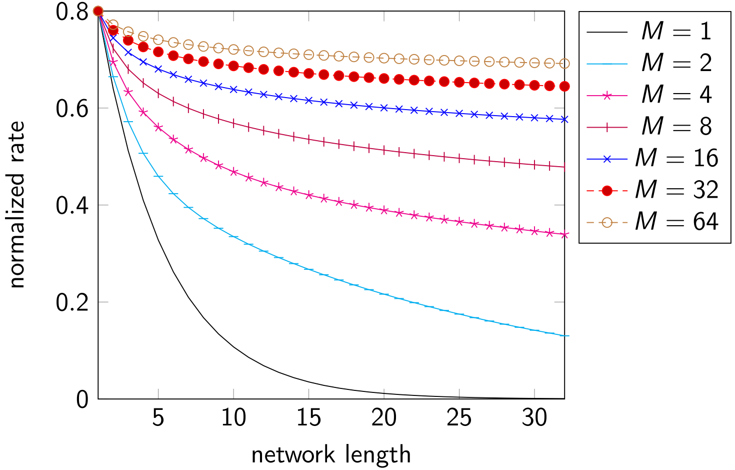
BATS codes with $M = 32$ and $q = 256$.

K	coding overhead			inactivation no.		
	average	max	min	average	max	min
1600	2.04	16	0	94.0	119	72
8000	6.30	77	0	215.5	268	179
16000	26.58	1089	0	352.2	379	302

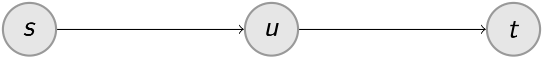
Tradeoff

- $M = 1$: BATS codes degenerate to Raptor codes.
  - Low complexity
  - No benefit of network coding
- $M = K$ and $\mathrm{degree} \equiv K$: BATS codes becomes RLNC.
  - High complexity
  - Full benefit of network coding.
- Exist parameters with moderate values that give very good performance

Tradeoff for Line Networks



Experiment setting



- Packet loss rate 0.2.
- Node s encodes K packets using a BATS code.
- Node u caches only one batch.
- Node t sends one feedback after decoding.

Experiment setting

Experiment setting

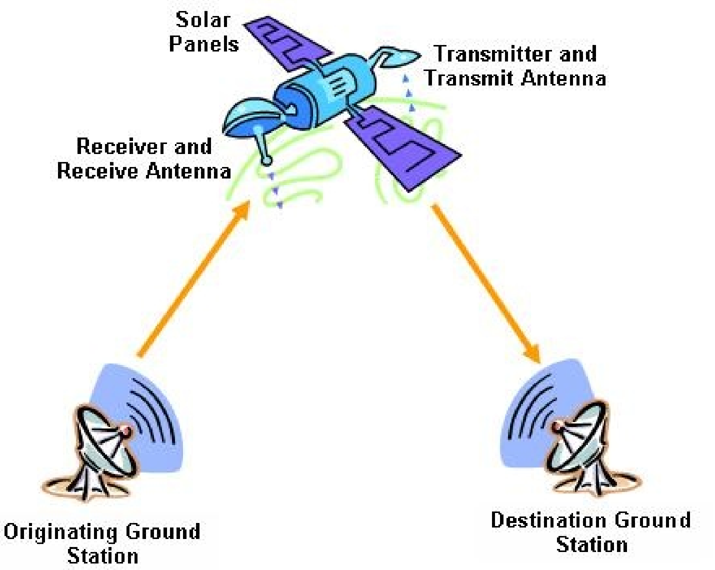
Experiment setting

- Sender/receiver: a laptop with open source Atheros wireless drivers.
- Relay: one wireless router with Atheros chipset running OpenWrt (about 150HKD/20USD)
- WiFi 802.11 b/g/n at 2.4GHz
- Sender's rate is set to 1 Mb/s to reduce the effect of the router's low computation power.

Compare with TCP

	Average rate (Kb/s)
BATS w/ recoding	592.86
BATS w/o recoding	530.65
TCP (normal 802.11)	420.33

Application: Satellite Communication



Summary

- BATS codes provide a digital fountain solution with linear network coding:
    - Outer code at the source node is a matrix fountain code.
    - Linear network coding at the intermediate nodes forms the inner code.
    - Prevents BOTH packet loss and delay from accumulating along the way.
- The more hops between the source node and the sink node, the larger the benefit.