
Peer reviewed version

Link to published version (if available): 10.1109/ITWNIT.2009.5158583

Link to publication record in Explore Bristol Research

PDF-document

**University of Bristol - Explore Bristol Research**

**General rights**

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/pure/about/ebr-terms
AND-OR Tree Analysis of Distributed LT Codes

D. Sejdinovic, R.J. Piechocki, A. Doufexi

Centre for Communications Research, University of Bristol

2009 Information Theory Workshop, Volos, Greece

d.sejdinovic@bristol.ac.uk
1. Basic concepts
   - Digital fountain paradigm
   - Distributed LT codes

2. Main contribution
   - Selective combining
   - Numerical results
Outline

1. Basic concepts
   - Digital fountain paradigm
   - Distributed LT codes

2. Main contribution
   - Selective combining
   - Numerical results
Fountain coding for erasure recovery

- Efficient and robust solution for data transmission over packet erasure networks
- Rateless codes - can adapt their rate on the fly to suit different (or unknown) channel conditions
- FC avoids feedback implosion typical for ARQ broadcast schemes (stupendous number of retransmissions); this renders them attractive for broadcast/multicast
- A kind of sparse graph codes; typically decoded by belief propagation algorithm (low computational complexity)
- Increasingly commercialised, standardized by 3GPP and DVB-h
(Luby 2002), $LT(\Omega(x), k)$.

- Sample an output degree $d$ with probability $\Omega_d$.
- Sample $d$ data packets uniformly at random from the information sequence and XOR them.

There are (universal) capacity achieving ensembles oblivious of erasure probability of the channel with computational cost of $O(k \log k)$ - robust soliton distributions.
Raptor codes

• (Shokrollahi 2006), Keep $\Omega(x)$ capped at a maximum degree $d_{max}$ as $k \to \infty$ (lowers computational cost to $O(k)$ but introduces an error floor).

• Error floor is removed by an outer very high rate LDPC code.

• There are universal ensembles with linear encoding-decoding complexity.
\[ E = \Omega'(1)(1 + \varepsilon)k = \Lambda'(1)k, \text{ denote } \alpha := \Lambda'(1). \]

- **Factor graph degree distributions:**
  - Output nodes: \( \Omega(x) \) (node-perspective) and \( \omega(x) = \frac{\Omega'(x)}{\Omega'(1)} \) (edge-perspective) on output nodes
  - Input nodes: \( \Lambda(x) \sim \text{Binomial}(\frac{1}{k}, \alpha k) \), which converges pointwise to Poisson(\( \alpha \)), i.e., \( \Lambda(x) = e^{-\alpha(x-1)} \).
(Luby, Mitzenmacher, Shokrollahi 1998) AND-OR tree evaluation

Übergeneralised to density evolution techniques (Richardson, Urbanke, MCT, 2008)

**Theorem**

The erasure rate of an LT\((k, \Omega(x))\) at overhead \(\varepsilon\), as \(k \to \infty\), is given by \(y = \lim_{l \to \infty} y_l\), where \(y_l\) is given by:

\[
y_0 = 1,
\]

\[
y_l = \exp(-\alpha \omega(1 - y_{l-1})).
\]
Optimisation of $\Omega(x)$

Fix $\alpha$ and $\delta$ and minimise $\varepsilon$:

$$
\text{LP} : \quad \min 1 + \varepsilon (= \alpha \sum_{d}^{d_{\text{max}}} \frac{\omega_d}{d})
$$

$$
\alpha \sum_{d=1}^{d_{\text{max}}} \omega_d x_i^{d-1} \geq -\ln (1 - x_i), \quad i \in 1, 2, \ldots, m
$$

$$
\sum_{d}^{d_{\text{max}}} \omega_d = 1,
$$

- $0 = x_1 < x_2 < \cdots < x_m = 1 - \delta$ are $m$ equidistant points on $[0, 1 - \delta]$, $\delta$ is the desired erasure ratio, and $d_{\text{max}}$ is the maximum degree of the degree distribution.

- Looking at the dual program tells us that solution does not depend on $\alpha$!
Fix $\delta$ and minimise $\varepsilon$:

$$
\text{LP B:} \quad \min \sum_{d=1}^{d_{\text{max}}} \omega_d x_i^{d-1} \geq -\ln(1 - x_i), \quad i \in 1, 2, \ldots, m
$$
1 Basic concepts
   - Digital fountain paradigm
   - Distributed LT codes

2 Main contribution
   - Selective combining
   - Numerical results
Fountain codes in multicast
Can relay help by combining packets from two sources?

(Puducheri, Kliwer, Fuja, 2007)
Can relay help by combining packets from two sources?

(Puducheri, Kliwer, Fuja, 2007)
Distributed LT codes

- Introduced in (Puducheri et al, 2007) for encoding data from multiple sources independently
- common relay combines encoded packets from multiple sources
- resulting bit-stream approximates that of an LT code (deconvolution of Robust Soliton)
- substantial benefits in comparison with independent LT encoders and forwarding at the relay.
Main idea

- **Main insight:** Combining data from multiple sources produces decoding problem of larger size, and potentially a smaller code overhead.

- **Main question:** How to perform (small) decentralized encoding tasks such that the large decoding problem is well-behaved?
Main idea

- **Main insight:** Combining data from multiple sources produces decoding problem of larger size, and potentially a smaller code overhead.

- **Main question:** How to perform (small) decentralized encoding tasks such that the large decoding problem is well-behaved?
1 Basic concepts
   - Digital fountain paradigm
   - Distributed LT codes

2 Main contribution
   - Selective combining
   - Numerical results
Selective combining

- Fix an arbitrary number of sources \( t \).
- Greedy BP decoder requires an output node of degree 1, but smallest degree in \( \Phi(x)^t \) is \( t \) (allowing degree-0 packets is wasteful of resources).
- Relay needs to **selectively combine** incoming packets.
- DLT codes introduced in (Puducheri et al, 2007) define selective combining for 2 and 4 sources, naturally extendable to \( 2^m, m \in \mathbb{N} \).
- In our work, selective combining can be performed independently of the degrees of the incoming packets, and we can construct good codes for any number of sources.
LT encoding at the relay $\text{SDLT}(\Gamma(x), \Phi(x), t, k)$:

- Sample a combining degree $d$ with probability $\Gamma_d$
- Sample $d$ incoming data packets uniformly at random and XOR them.
- $tk(1+\varepsilon)$ recoded packets should suffice for decoding.
Decoding graph

- Collection of small graphs corresponding to encoding at the sources $\{G_i\}_{i=1}^t\ (\mathrm{LT}(\Phi(x), k))$
- $\mathcal{H} = \sqcup G_i$ (graphs $G_i$ share output nodes)
- $\mathcal{G}$- actual decoding graph, created after removing edges from $\mathcal{H}$, according to encoding operation at the relay.
Digital fountain paradigm penalises reception rather than transmission (Shamai, Telatar, Verdu 2007)

**Example**

\[ \text{SDLT}(\Gamma(x), \Phi(x), t, k): \] \( t \) sources are continually broadcasting data to a large number \( r \gg t \) of relays via lossy links. As all incoming packets are equally important descriptions of its source, any relay can tune into a desired number of ongoing broadcasts at any time, and can combine incoming packets from different time slots.
Theorem

The erasure rate of an SDLT(Γ(x), Φ(x), t, k), as \( k \to \infty \), is given by \( y = \lim_{l \to \infty} y_l \), where \( y_l \) is given by:

\[
\begin{align*}
y_0 &= 1, \\
y_l &= \exp\left(-\bar{\alpha} \phi(1 - y_{l-1}) \gamma(\Phi(1 - y_{l-1}))\right).
\end{align*}
\]

where \( \bar{\alpha} = \Gamma'(1)\Phi'(1)(1 + \varepsilon) \) is the average input degree on the decoding graph.

Corollary

The performance of a selective distributed LT code ensemble SDLT(Γ(x), Φ(x), t, k) is identical to the performance of an LT code ensemble LT(Γ(Φ(x)), tk) as \( k \to \infty \).
DE for SDLT

**Theorem**

The erasure rate of an SDLT(Γ(x), Φ(x), t, k), as k → ∞, is given by

\[ y = \lim_{l \to \infty} y_l, \]

where \( y_l \) is given by:

\[ y_0 = 1, \]
\[ y_l = \exp(-\bar{\alpha} \phi(1-y_{l-1}) \gamma(\Phi(1-y_{l-1}))). \]

where \( \bar{\alpha} = \Gamma'(1)\Phi'(1)(1+\varepsilon) \) is the average input degree on the decoding graph.

**Corollary**

The performance of a selective distributed LT code ensemble SDLT(Γ(x), Φ(x), t, k) is identical to the performance of an LT code ensemble LT(Γ(Φ(x)), tk) as k → ∞.
In collection of the graphs $\{G_i\}_{i=1}^t$, $D_{out} \sim \Phi(x)$, $D_{in} \sim \text{Poisson}(\alpha)$.

In their output-shared union $\mathcal{H}$, $D_{out} \sim \Phi(x)^t$, $D_{in} \sim \text{Poisson}(\alpha)$.

In the overall decoding graph $\mathcal{G}$, $\bar{D}_{out} \sim \Gamma(\Phi(x))$, $\bar{D}_{in} \sim \sum_{i=1}^{D_{in}} X_i$, where $X_i \sim \text{Bernoulli}(\Gamma'(1)/t)$, i.e., $\bar{D}_{in}$ is a thinning of $D_{in}$, and thinning conserves the Poisson law.
LP B:\quad \text{min} \sum_{d=1}^{d_{\text{max}}} \frac{\gamma_d}{d} \\
\sum_{d=1}^{d_{\text{max}}} \gamma_d [\Phi'(x_i) \Phi(x_i)^{d-1}] \geq -\ln(1-x_i), \quad i \in 1, 2, \ldots, m
Outline

1. Basic concepts
   - Digital fountain paradigm
   - Distributed LT codes

2. Main contribution
   - Selective combining
   - Numerical results
Extreme cases

- **Coding only at source nodes:** $\text{SDLT}(x, \Phi(x), t, k)$
- **Coding only at relay nodes:** $\text{SDLT}(\Gamma(x), x, t, k)$
- **If $k$ is relatively small ($k \sim 1000$), coding only at relay performs better than coding only at sources even for small values of $t$, even though $\Gamma(x)$ has a very low maximum degree (problem is of a larger size).**
- **If we encode both at the sources and at the relay, we benefit from both:**
  - larger maximum degree on output degree distribution of the overall graph,
  - larger problem size.
Some nice distributions for $t = 10$

- $\Phi(x) = x$ and $\Gamma(x) = 0.0048x + 0.4422x^2 + 0.3075x^3 + 0.2455x^{10}$ reaches $\delta = 0.02$ at $1 + \varepsilon^* = 1.029$.

- $\Phi(x) = 0.05x + 0.5x^2 + 0.4x^3 + 0.05x^4$ and $\Gamma(x) = 0.7741x + 0.0025x^2 + 0.149x^3 + 0.026x^4 + 0.0718x^{10}$ reaches $\delta = 0.02$ at $1 + \varepsilon^* = 0.9983$. 
Numerical results

Figure: Simulation results for various selective distributed LT scenarios

- SDLT($\Gamma_R(x)$, $x$, $t=10$, $k=\infty$)
- SDLT($\Gamma^*(x)$, $\Phi^*(x)$, $t=10$, $k=\infty$)
- SDLT($\Gamma_R(x)$, $x$, $t=10$, $k=10^3$)
- SDLT($\Gamma^*(x)$, $\Phi^*(x)$, $t=10$, $k=10^3$)
- LT($\Gamma^*(\Phi^*(x))$, $k=10^4$)
- SDLT($x$, $\Phi_\rho(x)$, $t=10$, $k=10^3$)
Distributed and decentralised fountain codes can be studied by standard density evolution (AND-OR tree analysis) techniques.

It may be advantageous to perform fountain encoding at source nodes and intermediate nodes (recoding).

Outlook

- More rigorous analysis of behaviour of underlying LPs for optimisation.
- Joint optimisation of $\Phi$ and $\Gamma$.
- Layers of relays (recoding of recoding of recoding)
- Raptor scenario: how to seamlessly embed precodes to remove error floors?
Summary

- Distributed and decentralised fountain codes can be studied by standard density evolution (AND-OR tree analysis) techniques.

- It may be advantageous to perform fountain encoding at source nodes and intermediate nodes (recoding).

Outlook

- More rigorous analysis of behaviour of underlying LPs for optimisation.
- Joint optimisation of Φ and Γ.
- Layers of relays (recoding of recoding of recoding)
- Raptor scenario: how to seamlessly embed precodes to remove error floors?
Distributed and decentralised fountain codes can be studied by standard density evolution (AND-OR tree analysis) techniques.

It may be advantageous to perform fountain encoding at source nodes \textit{and} intermediate nodes (recoding).

Outlook

- More rigorous analysis of behaviour of underlying LPs for optimisation.
- Joint optimisation of $\Phi$ and $\Gamma$.
- Layers of relays (recoding of recoding of recoding)
- Raptor scenario: how to seamlessly embed precodes to remove error floors?