Maximum Likelihood Estimation for Multiple-Source Loss Tomography with Network Coding

Pegah Sattari
University of California, Irvine

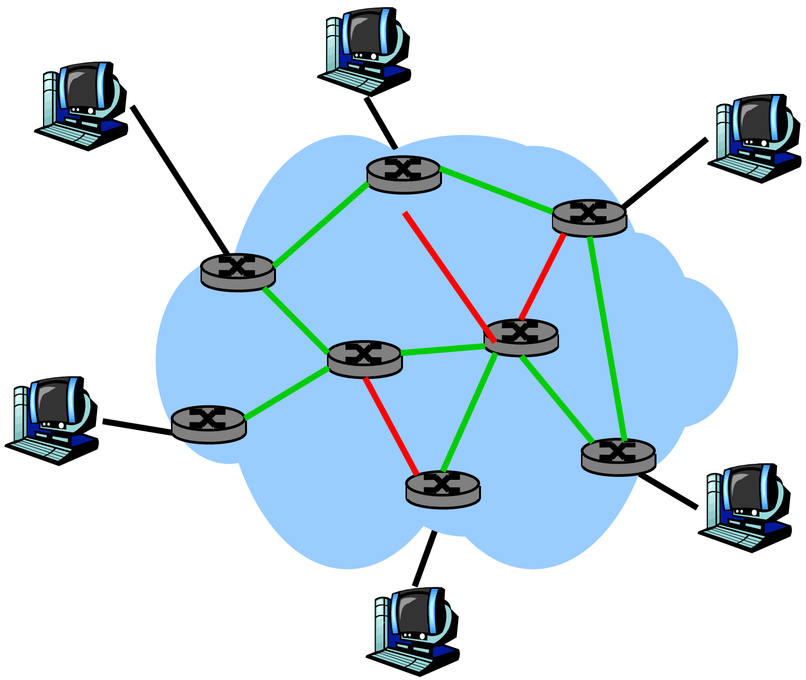
Joint work with:
Athina Markopoulou @ UC Irvine
Christina Fragouli @ EPFL

Outline

- Background and Motivation
- Model and Framework
- Likelihood Equation and Its Solution
  - Special cases: MINC and RMINC
  - Reversibility
- Maximum Likelihood Estimation
  - Reductions to multicast and reverse multicast
  - Estimating the coding edge
- Analysis of the MLE
  - Complexity and rate of convergence
  - Example

What is Network Tomography?

- Goal:
  - Obtain detailed picture of network from end-to-end views.
  - Design: sources/receivers, probes, routing/coding operations
- Characteristics of interest:
  - Topology or link-level characteristics, e.g, link loss, delay, etc



Link Loss Tomography
Related Work

- Single multicast tree:
  - MINC: Caceres et al., ToIT 1999
- Followed by:
  - Computationally efficient, suboptimal algorithms vs. MLE
  - General topologies: covering the network with several trees [Lo Presti and Duffield]
  - Unicast probes and packet-pairs
  - Joint topology and link loss inference [Nowak et al.]
- Network coding:
  - Active: suboptimal approaches for trees and general graphs [Markopoulou et al., arXiv:1005.4769]
  - Passive: failure patterns [Ho et al.], topology and error localization [Jaggi et al., Jafarisiavoshani et al.]

- We build on MLE for a multicast tree (MINC) and extend it to multiple-source trees with multicast and network coding

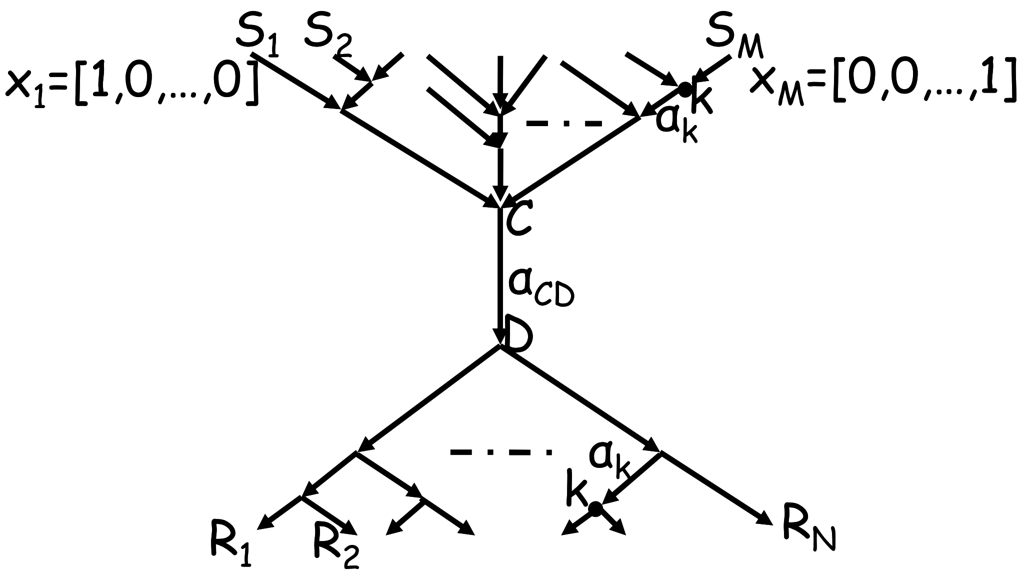
Network Coding Benefits

- Coded probe observes multiple paths, thus contains more information.
- One probe per link: overlapping probe paths "share" the bandwidth
- Optimal probe routing is NP-hard in traditional tomography, LP with network coding [Markopoulou et al., arXiv:1005.4769]
- Benefits are even more pronounced in general topologies.

Outline

- Background and Motivation
- Model and Framework
- Likelihood Equation and Its Solution
  - Special cases: MINC and RMINC
  - Reversibility
- Maximum Likelihood Estimation
  - Reductions to multicast and reverse multicast
  - Estimating the coding edge
- Analysis of the MLE
  - Complexity and rate of convergence
  - Example

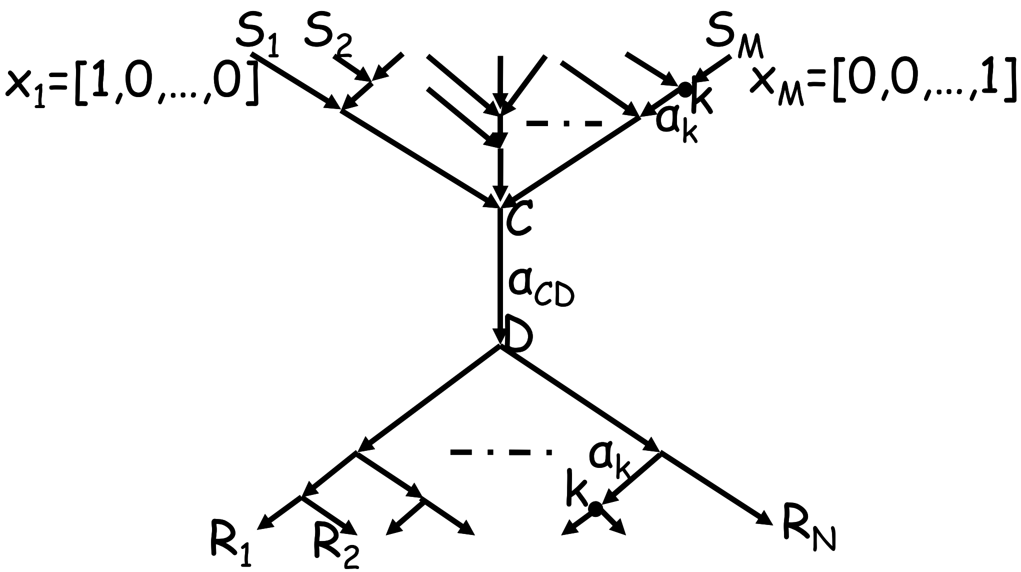
Model and Framework



- Link k denotes (k,j) if k>C, j>C
- Link k denotes (j,k) if j<D, k<D

- Logical directed tree
- M sources, N receivers
- f(j): parents of j
- d(j): children of j
- C has P parents and D has Q children
- K<k', k,k' in V, when k is a descendant of k' and vice versa

Model and Framework, Cont'd



- X_k: packet observed at node k
  - Binary vector of length M
  - $(X_k)_i$ represents the probe packet of source i
- $X=(X_k)$, k in V: the set of all X_k

- Link loss rates: i.i.d Bernoulli
- Packet traversing link k above C:
  - lost with prob. $\bar{a}_k=1-a_k$ and arrives at j with prob. a_k
- Packet traversing link k below D:
  - lost with prob. $\bar{a}_k=1-a_k$ and arrives at k with prob. a_k
- Loss rate of CD: $\bar{a}_{CD}$

Model and Framework, Cont'd

Data, Likelihood, and Inference

- Outcome of a single trial:
  - Set of X_k's observed at receiver k: $X_{(R)}=(X_k)_{k \text{ in } R}$
  - Element of the space $\Omega=\{[\ldots,0,1,\ldots]\}^N$
- Given $\alpha=(\alpha_k)_{k \text{ in } V\backslash\{C,D\}} \cup \alpha_{CD}$:
  - P_α : distribution of the outcomes $X_{(R)}$ on Ω
- PMF for an outcome x in Ω: $p(x;\alpha)=P_\alpha(X_{(R)}=x)$
- n trials, n(x): #probes for which x is obtained.
  - n independent observations: $p(x^1,\ldots x^n;\alpha)=\prod_{x \text{ in } \Omega} p(x;\alpha)^{n(x)}$

- Goal: estimate α using ML, from data $(n(x))_{x \text{ in } \Omega}$
- log-likelihood:
  - $\mathcal{L}(\alpha) = \log p(x^1,\ldots x^n;\alpha) = \Sigma_{x \text{ in } \Omega}\, n(x)\log p(x;\alpha)$
- MLE $\breve{\alpha}$ is the α that maximizes $\mathcal{L}(\alpha)$:
  - $\breve{\alpha} = \arg\max_\alpha \mathcal{L}(\alpha)$

Outline

- Background and Motivation
- Model and Framework
- Likelihood Equation and Its Solution
  - Special cases: MINC and RMINC
  - Reversibility
- Maximum Likelihood Estimation
  - Reductions to multicast and reverse multicast
  - Estimating the coding edge
- Analysis of the MLE
  - Complexity and rate of convergence
  - Example

Likelihood Equation and Its Solution

- Candidates for the MLE are solutions $\hat{\alpha}$ of the likelihood equation:

$$\frac{\partial \mathcal{L}}{\partial \alpha_k}(\alpha) = 0, \quad k \in V$$

- Some additional variables:
  - For each node k>=D, $\Omega^r(k)$=set of outcomes x s.t. $(x_b)_j$ not 0 for at least one source j that is an ancestor of k and for any arbitrary set of receivers b, a subset of R
  - $\gamma^r_k = P_\alpha[\Omega^r(k)]$: the outcomes where link k has definitely worked

$$\hat{\gamma}^r_k = \sum_{x \in \Omega^r(k)} \hat{p}(x) \qquad \hat{p}(x) = \frac{n(x)}{n}$$

  - Similarly, $\Omega^m(k)$ for k<=C, $\gamma^m_k = P_\alpha[\Omega^m(k)]$, its estimate…
  - γ^m_k: prob. of outcomes $\Omega^m(k)$ where link k has definitely worked

Likelihood Equation and Its Solution, Cont'd

- Goal: compute $\hat{a}$ from $\hat{Y}=(\hat{Y}^r_k \cup \hat{Y}^m_k)_{k \text{ in } V}$
- Special cases: MINC and RMINC
- MINC [Caceres et al., 1999]:
  - M=1, root 0, α_k: k=endpoint, α_0=1, $\Omega=\{0,1\}^N$
  - For each node k: $\Omega^m(k)$, γ^m_k only
  - A^m_k: prob. that the path from root to node k works
  - MINC:

$$\hat{\alpha}_k = \frac{\hat{A}^m_k}{\hat{A}^m_{f(k)}}, \quad k \in V \backslash \{0\}$$

$$1 - \frac{\hat{\gamma}^m_k}{\hat{A}^m_k} = \prod_{j \in d(k)} \left(1 - \frac{\hat{\gamma}^m_j}{\hat{A}^m_k}\right)$$

Likelihood Equation and Its Solution

MINC example

[Caceres, ToIT1999]

- Example: 3-link tree
  - Estimate from observations γ_k : prob. that at least one receiver that is a descendant of k observes something.

$$\widehat{\gamma}_1 = \widehat{p}(11) + \widehat{p}(10) + \widehat{p}(01)$$

$$\widehat{\gamma}_2 = \widehat{p}(11) + \widehat{p}(10)$$

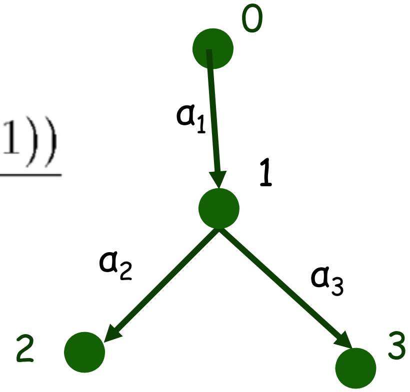
$$\widehat{\gamma}_3 = \widehat{p}(11) + \widehat{p}(01)$$

  - Find α_k from γ_k

$$\widehat{\alpha}_1 = \frac{\widehat{\gamma}_2\widehat{\gamma}_3}{\widehat{\gamma}_2 + \widehat{\gamma}_3 - \widehat{\gamma}_1} = \frac{(\widehat{p}(01) + \widehat{p}(11))(\widehat{p}(10) + \widehat{p}(11))}{\widehat{p}(11)}$$

$$\widehat{\alpha}_2 = \frac{\widehat{\gamma}_2 + \widehat{\gamma}_3 - \widehat{\gamma}_1}{\widehat{\gamma}_3} = \frac{\widehat{p}(11)}{\widehat{p}(01) + \widehat{p}(11)}$$

$$\widehat{\alpha}_3 = \frac{\widehat{\gamma}_2 + \widehat{\gamma}_3 - \widehat{\gamma}_1}{\widehat{\gamma}_2} = \frac{\widehat{p}(11)}{\widehat{p}(10) + \widehat{p}(11)}$$



Likelihood Equation and Its Solution, Cont'd

- RMINC:
  - N=1, receiver 0, α_k:k=starting point, α_0=1, $\Omega=\{0,1\}^M$
  - For each node k: $\Omega^r(k)$, γ^r_k only
  - A^r_k: prob. that path from node k to receiver works
  - RMINC:

$$\hat{\alpha}_k = \frac{\hat{A}^r_k}{\hat{A}^r_{d(k)}}, \quad k \in V \backslash \{0\}$$

$$1 - \frac{\hat{\gamma}^r_k}{\hat{A}^r_k} = \prod_{j \in f(k)} (1 - \frac{\hat{\gamma}^r_j}{\hat{A}^r_k})$$

  - MINC and RMINC have the same functional form
    - Reversibility [Markopoulou et al., arXiv 1005:4769]

Likelihood Equation and Its Solution

RMINC example

- By symmetry, exactly the same as multicast tree
  - Estimate γ'_k : prob. that at least one source that is an ancestor of k appears at receiver's observation.

$$\hat{\gamma}'_1 = \hat{p}(11) + \hat{p}(10) + \hat{p}(01)$$

$$\hat{\gamma}'_2 = \hat{p}(11) + \hat{p}(10)$$

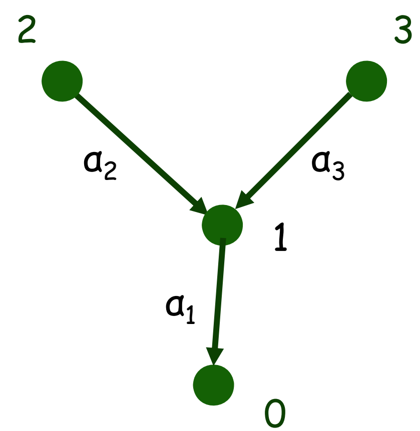
$$\hat{\gamma}'_3 = \hat{p}(11) + \hat{p}(01)$$

  - Can find α_k from γ'_k

$$\hat{\alpha}_1 = \frac{\hat{\gamma}'_2\hat{\gamma}'_3}{\hat{\gamma}'_2 + \hat{\gamma}'_3 - \hat{\gamma}'_1} = \frac{(\hat{p}(01) + \hat{p}(11))(\hat{p}(10) + \hat{p}(11))}{\hat{p}(11)}$$

$$\hat{\alpha}_2 = \frac{\hat{\gamma}'_2 + \hat{\gamma}'_3 - \hat{\gamma}'_1}{\hat{\gamma}'_3} = \frac{\hat{p}(11)}{\hat{p}(01) + \hat{p}(11)}$$

$$\hat{\alpha}_3 = \frac{\hat{\gamma}'_2 + \hat{\gamma}'_3 - \hat{\gamma}'_1}{\hat{\gamma}'_2} = \frac{\hat{p}(11)}{\hat{p}(10) + \hat{p}(11)}$$



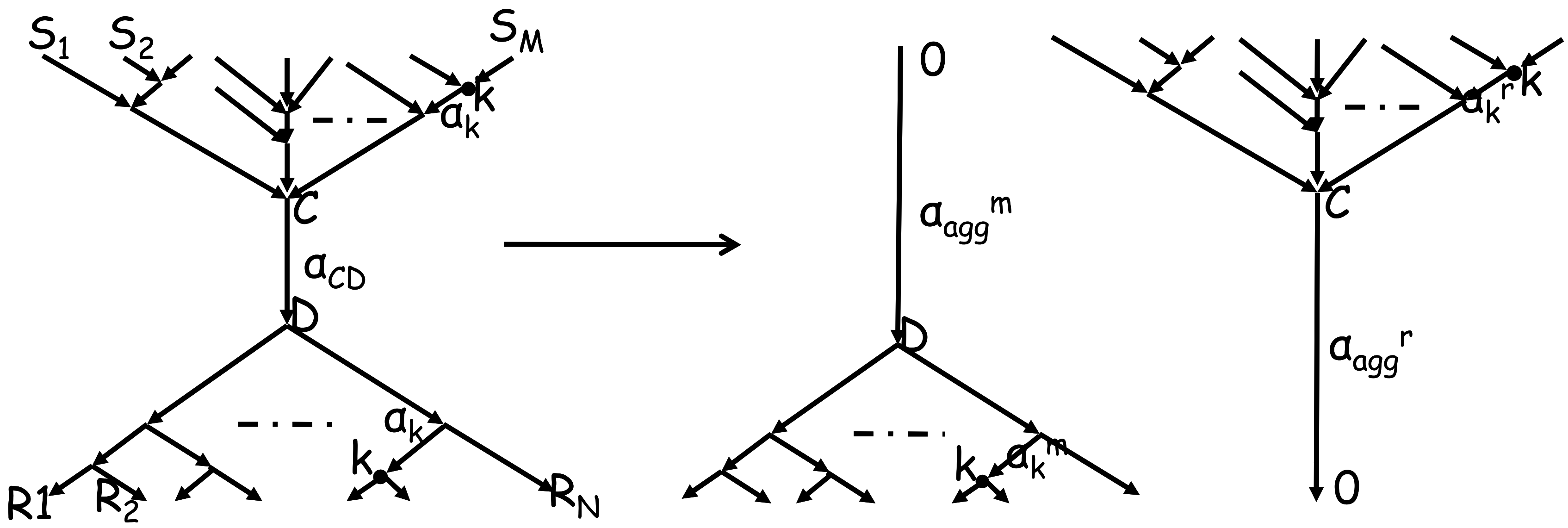
Outline

- Background and Motivation
- Model and Framework
- Likelihood Equation and Its Solution
  - Special cases: MINC and RMINC
  - Reversibility
- Maximum Likelihood Estimation
  - Reductions to multicast and reverse multicast
  - Estimating the coding edge
- Analysis of the MLE
  - Complexity and rate of convergence
  - Example

Maximum Likelihood Estimation

Reductions

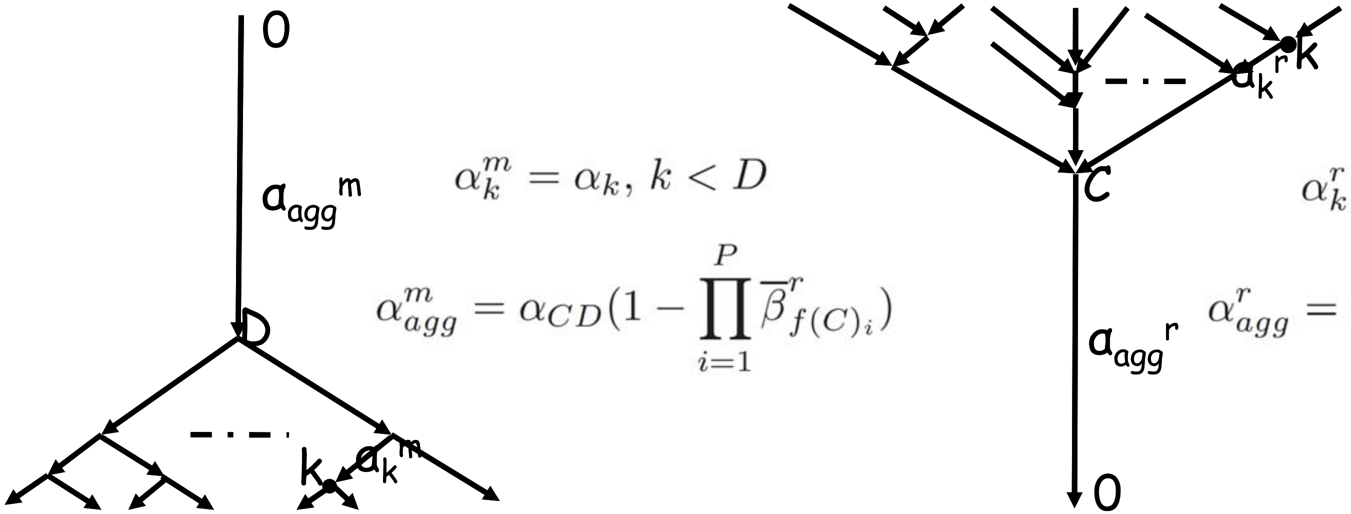
- Reduction to multicast (m): take upper part and consider it as a big aggregate link
  - D receivers a packet if at least one path up to C work and link CD works
- Reduction to reverse multicast (r)



Maximum Likelihood Estimation

Reductions, Cont'd

- Can map outcomes of the original tree to outcomes of the reduced tree
  - E.g., the prob. of an outcome in the reduced multicast tree is the sum of a set of outcomes in the original tree



$$\alpha_k^m = \alpha_k,\ k < D$$

$$\alpha_{agg}^m = \alpha_{CD}(1 - \prod_{i=1}^{P} \overline{\beta}^r_{f(C)_i})$$

$$\alpha_k^r = \alpha_k,\ k > C$$

$$\alpha_{agg}^r = \alpha_{CD}(1 - \prod_{j=1}^{Q} \overline{\beta}^m_{d(D)_j})$$

- Can show that: $\hat{\gamma}_C^r = \hat{\gamma}_D^m = 1 - \hat{p}([0, 0, \cdots, 0])$

Maximum Likelihood Estimation

General Statements

- Likelihood functions of general tree vs. reduced multicast/reverse multicast trees:
  - Different but maximized over the same values of their common parameters
  - These values can be proved to be:

$$\hat{\alpha}_k = \hat{\alpha}_k^r, k > C$$

$$\hat{\alpha}_k = \hat{\alpha}_k^m, k < D$$

Outline

- Background and Motivation
- Model and Framework
- Likelihood Equation and Its Solution
  - Special cases: MINC and RMINC
  - Reversibility
- Maximum Likelihood Estimation
  - Reductions to multicast and reverse multicast
  - Estimating the coding edge
- Analysis of the MLE
  - Complexity and rate of convergence
  - Example

Maximum Likelihood Estimation

Estimating the coding edge

- Can prove that:

$$\hat{\alpha}_{CD} = \frac{\hat{A}_C^r \cdot \hat{A}_D^m}{\hat{\gamma}_C^r} = \frac{\hat{A}_C^r \cdot \hat{A}_D^m}{\hat{\gamma}_D^m}$$

- MLE algorithm:
  - For all links k, k<D, reduce the original tree to multicast and use MINC:

$$\hat{\alpha}_k = \hat{\alpha}_k^r, \, k > C$$

  - For all links k, k>C, reduce the original tree to reverse multicast and use RMINC:

$$\hat{\alpha}_k = \hat{\alpha}_k^m, \, k < D$$

  - For link CD:

$$\hat{\alpha}_{CD} = \frac{\hat{A}_C^r \cdot \hat{A}_D^m}{\hat{\gamma}_C^r} = \frac{\hat{A}_C^r \cdot \hat{A}_D^m}{\hat{\gamma}_D^m}$$

Outline

- Background and Motivation
- Model and Framework
- Likelihood Equation and Its Solution
  - Special cases: MINC and RMINC
  - Reversibility
- Maximum Likelihood Estimation
  - Reductions to multicast and reverse multicast
  - Estimating the coding edge
- Analysis of the MLE
  - Complexity and rate of convergence
  - Example

Maximum Likelihood Estimation

- Very efficient
  - First two steps: MINC and RMINC
    - Computation at each node at worst proportional to the depth of the node
  - Last step (α_{CD}) uses estimates already computed in the first two steps
- From the asymptotic properties of MLEs:
  - $I(\alpha)$ is non-singular, and as n approaches infinity:

$$\sqrt{n}(\hat{\alpha} - \alpha) \xrightarrow[\text{in distribution}]{} \mathcal{N}(0, \mathcal{I}^{-1}(\alpha))$$

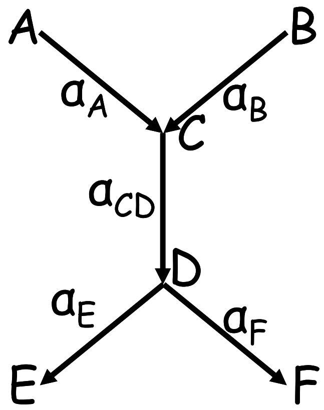
- Confidence interval asymptotically for large n:

$$\alpha_k \pm z_{\delta/2}\sqrt{\frac{\mathcal{I}_{kk}^{-1}(\alpha)}{n}}$$

Outline

- Background and Motivation
- Model and Framework
- Likelihood Equation and Its Solution
  - Special cases: MINC and RMINC
  - Reversibility
- Maximum Likelihood Estimation
  - Reductions to multicast and reverse multicast
  - Estimating the coding edge
- Analysis of the MLE
  - Complexity and rate of convergence
  - Example

Example: A 2-source 2-receiver Tree



#	Original (5-link) tree			Reduced multicast tree			Reduced reverse multicast tree	
	E	F	Prob.	E	F	P^m_α	EF	P^r_α
1	-	-	$\hat{p}_0$	0	0	$\hat{p}_0$	[0,0]	$\hat{p}_0$
2	x_1	-	$\hat{p}_1$				[1,0]	$\hat{p}_1 + \hat{p}_4 + \hat{p}_7$
3	x_2	-	$\hat{p}_2$	1	0	$\hat{p}_1 + \hat{p}_2 + \hat{p}_3$	[0,1]	$\hat{p}_2 + \hat{p}_5 + \hat{p}_8$
4	$x_1 \oplus x_2$	-	$\hat{p}_3$				[1,1]	$\hat{p}_3 + \hat{p}_6 + \hat{p}_9$
5	-	x_1	$\hat{p}_4$				[1,0]	$\hat{p}_1 + \hat{p}_4 + \hat{p}_7$
6	-	x_2	$\hat{p}_5$	0	1	$\hat{p}_4 + \hat{p}_5 + \hat{p}_6$	[0,1]	$\hat{p}_2 + \hat{p}_5 + \hat{p}_8$
7	-	$x_1 \oplus x_2$	$\hat{p}_6$				[1,1]	$\hat{p}_3 + \hat{p}_6 + \hat{p}_9$
8	x_1	x_1	$\hat{p}_7$				[1,0]	$\hat{p}_1 + \hat{p}_4 + \hat{p}_7$
9	x_2	x_2	$\hat{p}_8$	1	1	$\hat{p}_7 + \hat{p}_8 + \hat{p}_9$	[0,1]	$\hat{p}_2 + \hat{p}_5 + \hat{p}_8$
10	$x_1 \oplus x_2$	$x_1 \oplus x_2$	$\hat{p}_9$				[1,1]	$\hat{p}_3 + \hat{p}_6 + \hat{p}_9$

$$\hat{\gamma}^r_A = \hat{p}_1 + \hat{p}_3 + \hat{p}_4 + \hat{p}_6 + \hat{p}_7 + \hat{p}_9$$

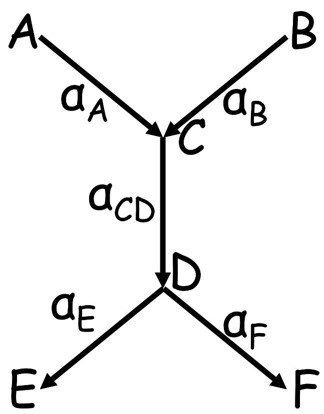
$$\hat{\gamma}^r_B = \hat{p}_2 + \hat{p}_3 + \hat{p}_5 + \hat{p}_6 + \hat{p}_8 + \hat{p}_9$$

$$\hat{\gamma}^r_C = \hat{\gamma}^m_D = \hat{p}_1 + \hat{p}_2 + \hat{p}_3 + \hat{p}_4 + \hat{p}_5 + \hat{p}_6 + \hat{p}_7 + \hat{p}_8 + \hat{p}_9 = 1 - \hat{p}_0$$

$$\hat{\gamma}^m_F = \hat{p}_4 + \hat{p}_5 + \hat{p}_6 + \hat{p}_7 + \hat{p}_8 + \hat{p}_9$$

$$\hat{\gamma}^m_E = \hat{p}_1 + \hat{p}_2 + \hat{p}_3 + \hat{p}_7 + \hat{p}_8 + \hat{p}_9$$

Example: A 2-source 2-receiver Tree



#	Original (5-link) tree			Reduced multicast tree			Reduced reverse multicast tree	
	E	F	Prob.	E	F	P^m_α	EF	P^r_α
1	-	-	$\hat{p}_0$	0	0	$\hat{p}_0$	[0,0]	$\hat{p}_0$
2	x_1	-	$\hat{p}_1$	1	0	$\hat{p}_1+\hat{p}_2+\hat{p}_3$	[1,0]	$\hat{p}_1+\hat{p}_4+\hat{p}_7$
3	x_2	-	$\hat{p}_2$				[0,1]	$\hat{p}_2+\hat{p}_5+\hat{p}_8$
4	$x_1 \oplus x_2$	-	$\hat{p}_3$				[1,1]	$\hat{p}_3+\hat{p}_6+\hat{p}_9$
5	-	x_1	$\hat{p}_4$	0	1	$\hat{p}_4+\hat{p}_5+\hat{p}_6$	[1,0]	$\hat{p}_1+\hat{p}_4+\hat{p}_7$
6	-	x_2	$\hat{p}_5$				[0,1]	$\hat{p}_2+\hat{p}_5+\hat{p}_8$
7	-	$x_1 \oplus x_2$	$\hat{p}_6$				[1,1]	$\hat{p}_3+\hat{p}_6+\hat{p}_9$
8	x_1	x_1	$\hat{p}_7$	1	1	$\hat{p}_7+\hat{p}_8+\hat{p}_9$	[1,0]	$\hat{p}_1+\hat{p}_4+\hat{p}_7$
9	x_2	x_2	$\hat{p}_8$				[0,1]	$\hat{p}_2+\hat{p}_5+\hat{p}_8$
10	$x_1 \oplus x_2$	$x_1 \oplus x_2$	$\hat{p}_9$				[1,1]	$\hat{p}_3+\hat{p}_6+\hat{p}_9$

$$\hat{\alpha}_A = \frac{\hat{\gamma}^r_A + \hat{\gamma}^r_B - \hat{\gamma}^r_C}{\hat{\gamma}^r_B} \qquad \hat{\alpha}_B = \frac{\hat{\gamma}^r_A + \hat{\gamma}^r_B - \hat{\gamma}^r_C}{\hat{\gamma}^r_A}$$

$$\hat{\alpha}_E = \frac{\hat{\gamma}^m_E + \hat{\gamma}^m_F - \hat{\gamma}^m_D}{\hat{\gamma}^m_F} \qquad \hat{\alpha}_F = \frac{\hat{\gamma}^m_E + \hat{\gamma}^m_F - \hat{\gamma}^m_D}{\hat{\gamma}^m_E}$$

$$\hat{\alpha}_{CD} = \frac{\hat{\gamma}^r_A \hat{\gamma}^r_B \hat{\gamma}^m_E \hat{\gamma}^m_F}{\hat{\gamma}^m_D(\hat{\gamma}^r_A + \hat{\gamma}^r_B - \hat{\gamma}^r_C)(\hat{\gamma}^m_E + \hat{\gamma}^m_F - \hat{\gamma}^m_D)}$$

Conclusion

- Low complexity MLE for link loss rates in trees with multicast and network coding
  - So far, MLE only for single-source multicast
  - suboptimal approaches for multiple-source trees

- Tomography vs. NC:
  - Context: in networks that already employ NC
  - NC allows for rapid inference
  - Internal nodes can still be "simple"
    - NC not more complex than forwarding or multicasting
    - other processing delegated to "special" nodes

Thank You!

questions?

psattari@uci.edu