Information Flow on Trees : A Survey

Adway Patra
Dept. of Electrical and Computer Engineering
University of Maryland, College Park
Maryland, United States of America
apatra@umd.edu

***Abstract*—Consider a tree network T where an information bit is placed at the root. Each edge of the tree acts as an independent channel and the information "progresses downwards" from the root towards the leaves. What can we deduce about the original information bit from the knowledge of the bits at any depth n? This problem can be further extended to the domain of any set of networks where a fixed acyclic direction of information flow is maintained. In this short survey article, we try to summarize some of the major results relevant to this question of broadcasting in graphs. We shall mostly focus on tree networks, briefly describing the system model, important results and connection to various fields of statistical physics and biology. Then we shall move on to a very recent extension of this question to random directed acyclic graphs and summarize the answers to similar questions.**

***Index Terms*—broadcasting, network, tree, directed acyclic graphs**

I. Introduction

Consider a tree $T = (V, E)$ and a channel M on finite alphabet $\mathcal{A}$. For $i, j \in \mathcal{A}$, let $P(M(i) = j) = M_{ij}$ and M defines an ergodic Markov chain. At the root ρ, a symbol from $\sigma_\rho \in \mathcal{A}$ is chosen according to some initial distribution π. For an edge $e \equiv (x, y) \in E$, x being the parent having the symbol $\sigma_x = i$, probability that $\sigma_y = j$ is M_{ij}. Hence, each edge acts as an independent identical channel. We denote $\sigma_{L_n} = (\sigma_v)_{v \in L_n}$ where L_n is the set of nodes at level n of T and $c_n(i) = |\{v \in L_n : \sigma_v = i\}|$. Denoting the total variation distance between distributions P and Q as $D_V(P, Q)$, we say that the reconstruction problem for T and M is solvable if $\forall \quad i, j \in \mathcal{A}$,

$$\lim_{n\to\infty} D_V(P_n^i, P_n^j) > 0 \tag{1}$$

where P_n^l denotes the conditional distribution of σ_{L_n} given $\sigma_\rho = l$. Similarly, we say that the reconstruction problem is census solvable if $\forall \quad i, j \in \mathcal{A}$,

$$\lim_{n\to\infty} D_V(\tilde{P}_n^i, \tilde{P}_n^j) > 0 \tag{2}$$

where $\tilde{P}_n^l$ denotes the conditional distribution of c_n given $\sigma_\rho = l$. Equivalently, these definitions can also be defined in terms of the mutual information $I(\sigma_\rho; \sigma_{L_n})$ (respectively $I(\sigma_\rho; c_n)$). We denote $\Delta_n(\pi)$ to be the success probability of the maximum likelihood (optimal) reconstruction algorithm of the initial symbol.

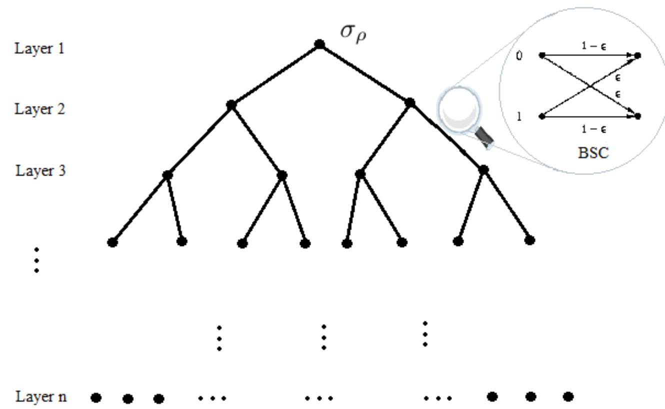

Fig. 1. A binary tree with each edge a BSC(ϵ)

II. Summary of Important Results

The following equivalent notions from [1] follow from the analysis based upon the fact that σ_{L_n} follows a Markov chain.

Proposition 1 (2.1 from [1]): For tree T and channel M, the following conditions are equivalent:

1) The reconstruction problem is solvable.
2) There exists a π for which $\lim_{n\to\infty} I(\sigma_\rho; \sigma_{L_n}) > 0$.
3) If π is uniform on $\mathcal{A}$, then $\lim_{n\to\infty} I(\sigma_\rho; \sigma_{L_n}) > 0$.
4) For any distribution π with $\min_i \pi_i > 0$, it holds that $\lim_{n\to\infty} I(\sigma_\rho; \sigma_{L_n}) > 0$.
5) There exists π for which $\lim_{n\to\infty} \Delta_n(\pi) > \max_i \pi_i$.

Analogous results also exist for c_n.

A special case of interest is when M is the Binary Symmetric Channel (BSC) with error probability ϵ. This is the ferromagnetic Ising Model in statistical physics. More specifically, if we associate bits $\{0, 1\}$ with $\{+1, -1\}$ spins, then the joint distribution of the vertices of any finite depth broadcasting sub-tree corresponds to the Boltzmann-Gibbs distribution on the sub-tree (see [2]). The Gibbs distribution is given by

$$\mathcal{G}(\sigma) = Z(t)^{-1} \exp(\frac{\sum_{u\sim v} J\sigma_u\sigma_v}{t}) \tag{3}$$

where t is the temperature, $Z(t)$ is a normalizing constant, $J > 0$ is the interaction strength and is related to ϵ as $\frac{\epsilon}{1-\epsilon} = \exp(-\frac{2J}{t})$. It was shown in [2] that the threshold for reconstruction depends on an inherent tree property called the Branching Number $br(T)$ as follows:

Definition 1 ([3]): For an edge $e \in E$, let $|e|$ denote the number of edges, including itself, on the path from ρ to e. The Branching Number $br(T)$ is defined as the supremum of real numbers $\lambda \geq 1$, such that T admits a positive flow from the root to infinity, if on every edge e, the flow is bounded by $\lambda^{-|e|}$.

Theorem 1 (Theorem 1.1 from [2]): Consider the problem of reconstructing σ_ρ from σ_{L_n} of T,

1) If $br(T)(1-2\epsilon)^2 > 1$, then $\inf_{n\to\infty} I(\sigma_\rho;\sigma_{L_n}) > 0$.
2) If $br(T)(1-2\epsilon)^2 < 1$, then $\inf_{n\to\infty} I(\sigma_\rho;\sigma_{L_n}) = 0$.

The proof is based upon upper and lower bounds on the quantity $I(\sigma_\rho;\sigma_{L_n})$. Notice that $(1-2\epsilon)$ is the second largest eigen-value of M in this case.

For census solvability, it turns out that the above threshold is applicable to all channels M. That is the reconstruction problem is census solvable when $br(T)|\lambda_2(M)|^2 > 1$ and not census solvable if $br(T)|\lambda_2(M)|^2 < 1$ [4], [5]. The non-solvability when equality holds is proven for some specific instance in [4].

The above theorems together show that, reconstruction by global majority is as good as maximum likelihood reconstruction. In evolutionary biology, given a bi-coloring of the boundary of a tree T, a parsimonious coloring of the internal nodes is a bi-coloring that minimizes total number of bi-colored edges. For a fixed tree and small $\epsilon > 0$, the maximum likelihood algorithm gives the same root value as one of the parsimonious colorings. But [6] showed that for $\epsilon \in (\frac{1}{\sqrt{2}}, \frac{3}{4}]$, maximum likelihood (i.e., majority rule) succeeds in reconstruction while parsimony fails.

III. Results on Directed Acyclic Graphs

We discuss a very recent extension of the above problem to random Directed Acyclic Graphs (DAG) [7]. Consider a directed graph with a single source vertex and every other non-source vertex having in-degree $d \geq 2$. Let L_k denote the set of vertices at layer k, i.e., the vertices at k-hop distance from the root. For every vertex in L_k, uniformly chose d vertices from L_{k-1} and construct the directed edges. Thus we have a DAG. Now a Bernoulli($\frac{1}{2}$) bit is chosen at the root. Each edge of the graph acts as an independent BSC(ϵ). Each node acts as a Boolean logic gate performing a (possibly node specific) Boolean operation on the noisy inputs to produce a single Boolean output which is then broad-casted downwards. One can identify two fundamental differences from the tree model:

- Unlike trees, the layer sizes do not need to scale exponentially.
- The in-degree is more than one so information processing is possible at nodes.

The question arises, can the benefits of the later overpower the shortcoming of the former and allow reconstruction of the root at any layer with sub-exponential size?

It was shown in [7] that for $d \geq 3$ and all Boolean processing functions to be d-input majority rule, ϵ_{maj} is a threshold for reconstruction where

$$\epsilon_{maj} = \frac{1}{2} - \frac{2^{d-2}}{\lceil \frac{d}{2} \rceil \binom{d}{\lceil \frac{d}{2} \rceil}} \tag{4}$$

Theorem 2: For a random DAG with $d \geq 3$ and majority processing functions (where ties are broken by outputting random equally likely bits), the following phase transition phenomenon occurs around ϵ_{maj}:

1) If $\epsilon \in (0, \epsilon_{maj})$ and the number of vertices per level satisfies $L_k \geq C(\epsilon, d) \log k$ for all sufficiently large k, then the reconstruction is possible in the sense that
$$\limsup_{k\to\infty} \mathbb{P}(Maj(L_k) \neq \sigma_\rho) < \frac{1}{2} \tag{5}$$
where $Maj(\cdot)$ is the majority decoder.
2) If $\epsilon \in (\epsilon_{maj}, \frac{1}{2})$ and the number of vertices per level satisfies $L_k o(D(\epsilon, d)^{-k})$, then reconstruction is impossible in the sense that
$$\lim_{k\to\infty} ||P^0_{k|G} - P^1_{k|G}|| = 0 \qquad G - a.s. \tag{6}$$
where $P^l_{k|G}$ is the conditional distribution of the bits at layer k given the structure of G and $\sigma_\rho = l$. Here, $C(\epsilon, d)$ and $D(\epsilon, d)$ are functions of ϵ and d described in detail in [7].

Thus, for $\epsilon < \epsilon_{maj}$, the majority decision rule can asymptotically recover the root-bit where as for $\epsilon > \epsilon_{maj}$ even knowing the exact structure of the graph and using maximum likelihood decision rule, it is not possible to recover the root. The paper also established the same threshold value for using single vertex reconstruction instead of majority rule reconstruction under the same assumptions. An explicit way to construct such DAGs for which the reconstruction is possible has also been given by using the ideas of regular bipartite lossless *expander graphs*.

IV. Conclusion

The purpose of this survey is to give a brief introduction to the latest and most important results in the domain of information broadcasting on graph networks. Such models are frequently encountered in fields of phylogeny, statistical physics and very recently in learning theory. The reader is referred to the corresponding papers for further details and various open problems.

References

[1] E. Mossel, “Survey: Information flow on trees,” in *Graphs, Morphisms and Statistical Physics*, 2001.
[2] W. Evans, C. Kenyon, Y. Peres, and L. J. Schulman, “Broadcasting on trees and the ising model,” *Ann. Appl. Probab.*, vol. 10, no. 2, pp. 410–433, 05 2000. [Online]. Available: https://doi.org/10.1214/aoap/1019487349
[3] R. Lyons, “Random walks and percolation on trees,” *Ann. Probab.*, vol. 18, no. 3, pp. 931–958, 07 1990. [Online]. Available: https://doi.org/10.1214/aop/1176990730
[4] E. Mossel and Y. Peres, “Information flow on trees,” *Ann. Appl. Probab.*, vol. 13, no. 3, pp. 817–844, 08 2003. [Online]. Available: https://doi.org/10.1214/aoap/1060202828
[5] H. Kesten and B. P. Stigum, “Additional limit theorems for indecomposable multidimensional galton-watson processes,” *Ann. Math. Statist.*, vol. 37, no. 6, pp. 1463–1481, 12 1966. [Online]. Available: https://doi.org/10.1214/aoms/1177699139
[6] M. A. Steel, “Distributions on bicoloured evolutionary trees,” *Bulletin of the Australian Mathematical Society*, vol. 41, no. 1, p. 159–160, 1990.
[7] A. Makur, E. Mossel, and Y. Polyanskiy, “Broadcasting on random directed acyclic graphs,” *IEEE Transactions on Information Theory*, vol. 66, no. 2, pp. 780–812, 2020.