Noisy Broadcast and Distributed Computation

Shouvik Ganguly

ECE 255C
Final Project Presentation

Introduction

- Shannon's paradigm :

  small number of users; each has amount of information going to infinity

- Alternative approach :

  each user has small amount of information; number of users growing large

Problem Statement

- Undirected graph $G = (V, \mathcal{E})$
- $i \in V$: $x_i \in \mathcal{X}$
- j : ONE receiver
- Goal : Compute $f(x^N)$ at j with $\Pr[\text{error}] \leq \delta$

Order-optimal protocol ??

Noisy Broadcast

- G is complete
- $\mathcal{X} = \{0, 1\}$
- f is the *parity* or *identity* function
- i.i.d. BSC(ϵ) links

First posed by El-Gamal, 1987 (El-Gamal, [1])

Noisy Broadcast : Simple Scheme

- Each node transmits n times
- Error probability $\leq N(4\epsilon(1-\epsilon))^{n/2}$
- $\mathcal{O}(N \log N)$ transmissions

Can we do better?

Noisy Broadcast : Gallager's Scheme for Parity (Gallager, [2])

- Partition the sender nodes into subsets of size $\approx k$
- Each sender node in a subset transmits n times
- Each sender node forms an estimate of the overall parity of its *own subset*
- Sender node adds its own parity to the estimate and transmits once more

Noisy Broadcast : Gallager's Scheme for Parity

- Majority decision at receiver node to determine parity of each subset
- $\Pr[\text{sender node incorrect}] = \epsilon_n < [4\epsilon(1-\epsilon)]^{n/2}$
- $\Pr[\text{incorrect subset parity received at receiver}]$

$$\leq \beta = \frac{1}{2}[1-(1-2\epsilon_n)^{k-1}(1-2\epsilon)]$$

- $\Pr[\text{overall decision incorrect}] \leq N[4\beta(1-\beta)]^{k/2}$

$$= N[1-(1-2\epsilon_n)^{2k-2}(1-2\epsilon)^2]^{k/2}$$

Noisy Broadcast : Gallager's Scheme for Parity

- $P \leq \delta$ can be achieved by $k = \mathcal{O}(\log N)$ and $n = \mathcal{O}(\log k)$
- $\mathcal{O}(N \log \log N)$ total transmissions
- Can also compute identity using $\mathcal{O}(N \log \log N)$ transmissions

Can we do better?

Converse for Noisy Broadcast (Goyal et. al. [3])

Theorem

$\forall\ \epsilon \in (0, 1/2)$, $\exists\ n_0(\epsilon)$ such that $\forall N \geq n_0$, any protocol for computing identity in noisy broadcast problem with BSC(ϵ) links requires at least $\dfrac{N \log\log N}{20 \log(1/\epsilon)}$ transmissions.

- Gallager's scheme is order-optimal for identity
- However, $\exists\ \mathcal{O}(N)$ scheme for computing *any* symmetric function (Goyal et. al. [3])

General Networks

- G not necessarily complete
- Example : *Random Geometric Network*
- Place N nodes uniformly at random inside a $\sqrt{N} \times \sqrt{N}$ square
- Each node connected to all nodes within a distance $r(N)$
- Reduces to broadcast network for $r \geq \sqrt{2N}$

Random Geometric Network : Symmetric Function Computation

- Aim : Compute symmetric function of inputs (Ying et. al. [4])
- Can be done in $\mathcal{O}(N \log\log N)$ transmissions if $r \geq \Theta(\sqrt{\log N})$
- Similar strategy : divide the square into $\mathcal{O}\left(\frac{N}{\log N}\right)$ cells

Random Geometric Network : Converse for Parity (Dutta et. al. [5])

Theorem

If $r = \mathcal{O}(N^{1/2-\beta})$ for some $\beta > 0$, any δ-error protocol for computing parity in a random geometric network with BSC(ϵ) links requires $\Omega(N \log\log N)$ transmissions.

- $\exists\ \mathcal{O}(N)$ protocol for computing the *maximum* function (Kanoria and Manjunath, [6])

Deterministic Geometric Networks

- $\sqrt{N} \times \sqrt{N}$ **grid** of N nodes
- Each node connected to all nodes within a distance $r(N)$
- $r \geq \sqrt{2N} \implies$ broadcast network
- Has properties similar to RGN for $r \geq \Theta(\sqrt{\log N})$ (Franceschetti et. al. [7])
- Connected even for $r < \Theta(\sqrt{\log N})$

Deterministic Geometric Networks : Computing Symmetric Functions

- How does it behave when $r < \Theta(\sqrt{\log N})$??
- Need $\Theta(N \log\log N)$ transmissions for $r > \Theta\left(\sqrt{\frac{\log N}{\log\log N}}\right)$
- Need $\Theta\left(\frac{N \log N}{r^2}\right)$ for $r < \Theta\left(\sqrt{\frac{\log N}{\log\log N}}\right)$ [7]

Deterministic Geometric Networks : Further Studies

- Nodes not having a common neighbour can share a time slot
- Target : Reduce number of time slots as well as number of transmissions
- Need $\max\{\Theta\left(\frac{\sqrt{N}}{r}\right), \Theta(r^2 \log\log N)\}$ time slots (Franceschetti et. al. [8])

References I

Abbas El-Gamal.
Reliable communication of highly distributed information.
Open Problems in Communication and Computation, T.M. Cover and B. Gopinath, eds., Springer-Verlag, New York, 1987, pp. 60–62.

R. Gallager.
Finding parity in simple broadcast networks.
IEEE Transactions on Information Theory, 34(2), pp 176–180, Mar 1988.

References II

- N. Goyal, G. Kindler and M. Saks.
  Lower bounds for the noisy broadcast problem.
  SIAM Journal of Computing, 37(6), pp. 1806–1841, March 2008.
- L. Ying, R. Srikant and G.E. Dullerud.
  Distributed symmetric function computation in noisy wireless sensor networks.
  IEEE Transactions on Information Theory, 53(12), pp 4826–4833, Dec 2007.

References III

- C. Dutta, Y. Kanoria, D. Manjunath and J. Radhakrishnan.
  A tight lower bound for parity in noisy communication networks.
  Proc. 19th Annual ACM-SIAM Symposium on Discrete Algorithms, pp 1056–1065, 2008.
- Y. Kanoria and D. Manjunath.
  On distributed computation in noisy random planar networks.
  Proc. International Symposium on Information Theory, pp 626–630, 2007.

References IV

- N. Karamchandani, R. Appuswamy and M. Franceschetti.
  Distributed computation of symmetric functions with binary inputs.
  Proc. IEEE Information Theory Workshop, pp 76–80, 2009.

- N. Karamchandani and M. Franceschetti.
  Distributed Function Computation in Networks: a Joint Delay-Energy Perspective
  Proc. 20th International Symposium on Optimization of Mobile, Ad Hoc and Wireless Networks, 2011.