My recent research

Yashodhan Kanoria
Note prepared for Jennifer Chayes

September 22, 2011

I will first touch informally on two projects close to my heart, before giving a more formal description of two recent accepted papers.

Much effort has gone into a Facebook game called 'Treasure Traders' which collects network bargaining data. The game has been well received and we have been getting large volumes of data. It will be some time before we collect and clean all the data we are aiming for, and then analyze it. But this is a very exciting direction for me, and I intend to continue in future this thread of running web-based experiments on networks.

Another recent project (to appear in SODA 2012 [1]) obtains a rigorous understanding of the clustering phase transition in random XORSAT. The XORSAT problem requires finding an assignment of n Boolean variables that satisfies m exclusive OR (XOR) clauses. We consider random XORSAT instances, with clauses of size k. We prove a complete characterization of the clustering phase transition for random k-XORSAT. In particular, we prove that the clustering threshold is sharp and determine its exact location. We prove that the set of solutions is well connected below this threshold and that each of the clusters is well connected above the same threshold. Our proof constructs a very sparse basis for the set of solutions. This is the first proof of the clustering phase transition in any random constraint satisfaction problem. The reason I'm not listing this work in my two best recent papers is that it does not relate to my focus areas of application.

Now a formal description of two recent accepted papers.

Distributed Storage for Intermittent Energy Sources: Control Design and Performance Limits

This paper [2], with Andrea, David Tse and Baosen Zhang will be presented at Allerton 2011.

The problem: One of the most important challenges in the integration of renewable sources into the energy infrastructure lies in their intermittent nature. More precisely, the power output of these sources varies with time and location due to factors that cannot be controlled by the provider. For instance, in the case of solar and wind energy, these exogenous factors are mainly meteorological. Three strategies have been proposed to hedge against this variability:

- Use energy storage to effectively average the produced power over time.
- Exploit distributed generation to average production over locations. This requires transmission infrastructure.
- Overprovision the grid, so that available supply exceeds demand on average.

How can we optimally design and operate the grid using these strategies?

Our solution: We assume a discrete time model with Gaussian net demand at each node, at each time. Given a network with fixed transmission and storage infrastructure, we want to minimize the amount of 'dirty energy' ϵ that must be produced using fossil fuels to meet the unfulfilled demand. We provide a systematic methodology using a (heuristic) linear-quadratic-Gaussian (LQG) formulation that leads to a linear control scheme to determine transmission and dirty energy production at each time.

To evaluate the performance of our scheme we analyze the symmetric 1-D and 2-D grids, with storage S and $\mathcal{N}(0,1)$ demand at each node, and transmission capacity C on each edge. We find that our heuristic methodology leads to surprisingly good performance in each case: we find that $\log(1/\epsilon) \sim \sqrt{CS}$ is achieved in the 1-D case and $\log(1/\epsilon) \sim CS$ is achieved in the 2-D case. Further, we find that these performances are the *best achievable* up to constants assuming a reasonable conjecture in probability theory.

We generalize our results to the case with overprovisioning (i.e., supply exceeds demand). Thus, our results provide both a control scheme, and a prescription for (near) optimal allocation of a given budget towards improving energy infrastructure. Qualitatively, we find that adding energy storage leads to dramatic improvements in performance.

We consider our modeling and our technical contributions in this work to be of equal importance. To the best of our knowledge, there have been no previous theoretical studies that analyze the role of storage in a network setting. As such, our model and results represent first steps in this core problem.

Subexponential convergence for information aggregation on regular trees

This paper [3], with Andrea, will be presented at CDC-ECC 2011 in December.

The problem. We consider the decentralized binary hypothesis testing problem on trees of bounded degree and increasing depth. For a regular tree of depth t and branching factor $k \geq 2$, we assume that the leaves have access to independent and identically distributed noisy observations of the 'state of the world' s. Starting with the leaves, each node makes a decision in a finite alphabet $\mathcal{M}$, that it sends to its parent in the tree. Finally, the root decides between the two possible states of the world based on the information it receives.

Context: This problem of decentralized hypothesis testing has been studied by Cover, Tsitsiklis and others since the 1970s, but progress has been slow, with no results on trees of bounded degree, or degree growing slowly with n. The only results have been on the line, and on trees of bounded *depth*, which are considerably easier to deal with.

Lower bounds on the error probability in our model have implications to social learning, since we obtain a universal upper bound on the rate at which agents can learn from their peers, independent of the update rule(s) agents employ.

Our solution. We prove that the error probability vanishes only subexponentially in the number of available observations, under quite general hypotheses. In the case of binary messages, decay is subexponential for any decision rule. Moreover, the simple 'majority' rule is nearly optimal. For general (finite) message alphabet $\mathcal{M}$, decay is subexponential for 'node-oblivious' decision rules, that satisfy a mild irreducibility condition. In the latter case, we propose a family of decision rules with close-to-optimal asymptotic behavior.

References

[1] M. Ibrahimi, Y. Kanoria, A. Montanari and M. Kraning, "The Set of Solutions of Random XORSAT Formulae," to appear in Proc. ACM-SIAM Symp. Discrete Algorithms, 2012.

[2] Y. Kanoria, A. Montanari, D. Tse and B. Zhang, "Distributed Storage for Intermittent Energy Sources: Control Design and Performance Limits," Proc. Annual Allerton Conference on Communication, Control, and Computing, 2011.

[3] Y. Kanoria, A. Montanari, "Subexponential convergence for information aggregation on regular trees," Proc. IEEE Conference on Decision and Control and European Control Conference (CDC-ECC), 2011.