STOC 2012 Poster Abstracts

Contents

The Stretch Factor of L_1- and L_∞-Delaunay Triangulations

N. Bonichon, C. Gavoille, N. Hanusse and L. Perkovic

In this paper we determine the stretch factor of the L_1-Delaunay and L_∞-Delaunay triangulations, and we show that this stretch is $\sqrt{4+2\sqrt{2}} \approx 2.61$. Between any two points x, y of such triangulations, we construct a path whose length is no more than $\sqrt{4+2\sqrt{2}}$ times the Euclidean distance between x and y, and this bound is best possible. This definitively improves the 25-year old bound of $\sqrt{10}$ by Chew (SoCG '86).

To the best of our knowledge, this is the first time the stretch factor of the well-studied L_p-Delaunay triangulations, for any real $p \geq 1$, is determined exactly.

The NOF Multiparty Communication Complexity of Composed Functions

Anil Ada, Arkadev Chattopadhyay, Omar Fawzi and Phuong Nguyen

We study the k-party 'number on the forehead' communication complexity of composed functions $f \circ g$, where $f : \{0,1\}^n \to \{\pm 1\}$, $g : \{0,1\}^k \to \{0,1\}$ and for $(x_1, \ldots, x_k) \in (\{0,1\}^n)^k$, $f \circ g(x_1, \ldots, x_k) = f(\ldots, g(x_{1,i}, \ldots, x_{k,i}), \ldots)$. We show that there is an $O(\log^3 n)$ cost simultaneous protocol for $\text{SYM} \circ g$ when $k > 1 + \log n$, SYM is any symmetric function and g is *any function*. Previously, an efficient protocol was only known for $\text{SYM} \circ g$ when g is symmetric and "compressible". We also get a non-simultaneous protocol for $\text{SYM} \circ g$ of cost $O(n/2^k \cdot \log n + k \log n)$ for any $k \geq 2$.

In the setting of $k \leq 1 + \log n$, we study more closely functions of the form $\text{MAJORITY} \circ g$, $\text{MOD}_m \circ g$, and $\text{NOR} \circ g$, where the latter two are generalizations of the well-known and studied functions Generalized Inner Product and Disjointness respectively. We characterize the communication complexity of these functions with respect to the choice of g. As an application of our results, we answer a question posed by Babai et al. (*SIAM Journal on Computing, 33:137–166, 2004*) and determine the communication complexity of $\text{MAJORITY} \circ \text{QCSB}_k$, where QCSB_k is the "quadratic character of the sum of the bits" function.

Spectral Norm of Symmetric Functions

Anil Ada, Omar Fawzi and Hamed Hatami

The spectral norm of a boolean function $f : 0,1^n - > -1,1$ is the sum of the absolute values of its Fourier coefficients, in other words, the l_1 norm of its Fourier coefficients. This quantity provides useful upper and lower bounds on the complexity of a function in areas like learning theory, circuit complexity and communication complexity, especially in settings where 'parity' represents an easy function. We give a combinatorial characterization for the spectral norm of all symmetric functions. We show that the logarithm of the spectral norm is essentially equal to $r(f) = maxr_0, r_1$, where r_0 and r_1 are the minimum integers less than $n/2$ such that $f(x)$ is a constant for all x with $|x|$ in $[r_0, n - r_1]$ or $f(x) = \text{parity}(x)$ for all x with $|x|$ in $[r_0, n - r_1]$.

A Uniform Min-Max Theorem and Its Applications

Colin Jia Zheng

I will describe a constructive proof of von Neumann's minmax theorem for 2-player zero-sum game — specifically, an algorithm which builds a near-optimal strategy for the first player from several best-responses of the first player to strategies of the second player. The idea is based on similar results of Freund and Schapire [FS99], however, our algorithm runs in poly(n) time even when a pure strategy for player 2 is a distribution (from a convex set of distributions) over $[N]$ for $N = 2^n$.

I will then highlight a few of its applications in computational complexity and cryptography: characterizing (uniform) conditional pseudoentropy [Vadhan and Zheng, 2011]; uniform hardcore lemma with optimal parameters; uniform dense model theorem; "separating SNARG from all falsifiable assumptions" [Gentry and Wichs, 2011] in the uniform model; etc.

On the Complexity of Trial and Error

Xiaohui Bei, Ning Chen and Shengyu Zhang

Motivated by certain applications from physics, biochemistry, economics, and computer science in which the objects under investigation are unknown or not directly accessible because of various limitations, we propose a trial-and-error model to examine search problems with *unknown* inputs. Given a search problem with a hidden input, we are asked to find a valid solution. The way to find such a solution is to propose candidate solutions, i.e., *trials*, and, using observed violations, i.e., *errors*, to prepare future proposals. In accordance with our motivating applications, we consider a fairly broad class of constraint satisfaction problems, and assume that errors are signaled by a verification oracle in the format of the index of a violated constraint (with the exact content of the constraint still hidden). The objective is to design time- and/or trial-efficient algorithms that will find a valid solution or alternatively, to show that the problem is intrinsically hard.

On one hand, despite the seemingly very little information provided by the verification oracle, we show that efficient algorithms do exist for a number of important problems. For the Nash, Core, Stable Matching, and SAT problems, the unknown-input versions are as hard as the corresponding known-input versions, up to a factor of polynomial. We further conduct a closer study of the latter two problems and give almost tight bounds on their trial complexities. The techniques employed to prove these results vary considerably, including, e.g., order theory and the ellipsoid method with a strong separation oracle.

On the other hand, there are problems whose complexities are substantially increased in the unknown-input model. For Graph Isomorphism and Group Isomorphism, in particular, although there are trial-efficient algorithms, no time-efficient algorithms exist (unless PH collapses and P = NP, respectively). These results also imply lower bounds on the tradeoff between time and trial complexities. The proofs use quite nonstandard reductions, in which an efficient simulator is carefully designed to simulate a desirable but computationally unaffordable oracle.

Our model investigates the value of information, and our results demonstrate that the lack of input information can introduce various levels of extra difficulty. The model accommodates a wide range of combinatorial and algebraic structures, and exhibits intimate connections with (and we hope can also serve as a useful supplement to) certain existing learning and complexity theories.

Online Bottleneck Matching

Barbara Anthony and Christine Chung

We consider the online bottleneck matching problem, where k server-vertices lie in a metric space and k request-vertices that arrive over time each must immediately be permanently assigned to a server-vertex. The goal is to minimize the maximum distance between any request and its server. It has been shown that no algorithm can have a competitive ratio better than $O(k)$ for this problem. The prohibitive general lower bound on the problem and the exceedingly poor performance of the simple and natural Greedy algorithm, motivate us to consider a benchmark that is less formidable than the optimal solution, in order to attain a more informative analysis of proposed algorithms for the bottleneck matching problem. Specifically, we employ a weak adversary model of analysis. Results obtained under this model of analysis can be viewed as “bicriteria" results.

In our setting with resource augmentation, we ask how well the online algorithm performs when it has multiple servers (namely two) per server-vertex, while the optimal offline solution only has one; thus the online algorithm can service twice as many request-vertices with each server-vertex. We show that while the competitive ratio of the naive Greedy algorithm improves from exponential (when each server-vertex has one server) to linear (when each server-vertex has two servers). We also show the competitive ratio of Permutation, which is the optimally competitive algorithm for the related min-weight matching problem, remains linear when an extra server is introduced at each server-vertex. And finally, we show that the competitive ratio of Balance, a modified form of Greedy that is more judicious in its use of the additional

server at each server-vertex, is also linear with an extra server at each server-vertex, even though it has been shown that an extra server makes it constant-competitive for the min-weight matching problem. These results suggest that in some sense the bottleneck objective is more difficult than the standard min-weight objective. Resource augmentation greatly helps Greedy for the min-weight objective, but none of the three algorithms break the $\Omega(k)$ barrier for the bottleneck objective. Our results also suggest that the basic Greedy algorithm can be a reasonable choice of algorithm for the bottleneck matching problem, due to its relative simplicity, and its comparable performance to Balance and Permutation against a weakened adversary.

Testing Lipschitz Functions on Hypergrid Domains

Pranjal Awasthi, Madhav Jha, Marco Molinaro and Sofya Raskhodnikova

A function $f(x_1, ..., x_d)$, where each input x_i is an integer from 1 to n, is Lipschitz if changing one of the inputs by 1 changes the output by at most 1. In other words, Lipschitz functions are not very sensitive to small changes in the input. Our main result is an efficient tester for the Lipschitz property of functions $f : 1, ..., n^d \to Z$. A property tester is given an oracle access to a function f and a proximity parameter epsilon, and it has to distinguish, with high probability, functions that have the property from functions that differ on at least an epsilon fraction of values from every function with the property. The Lipschitz property was first studied by Jha and Raskhodnikova (FOCS'11) who motivated it by applications to data privacy and program verification. They presented efficient testers for the Lipschitz property of functions on the domains $\{0, 1\}^d$ and $\{1, ..., n\}$. Our tester for functions on the more general domain $\{1, ..., n\}^d$ runs in time $O(d^1.5n(\log n))$.

The main tool in the analysis of our tester is a smoothing procedure that makes a function Lipschitz by modifying it at a few points. Its analysis is already nontrivial for the 1-dimensional version, which we call Bubble Smooth, in analogy to Bubble Sort. In one time step, Bubble Smooth modifies two consecutive values that violate the Lipschitz property, namely, differ by more than 1. It decreases the larger and increases the smaller by 1, i.e., it transfers a unit from the larger to the smaller. We introduce the transfer graph to keep track of the transfers, and use it to show that the L1 distance between f and Bubble-Smooth[f] is at most twice the L1 distance from f to the nearest Lipschitz function.

Limitations of Local Filters of Lipschitz and Monotone Functions

Pranjal Awasthi, Madhav Jha, Marco Molinaro and Sofya Raskhodnikova

We study local filters for two properties of functions of the form $f : 0, 1^d$ to R: the Lipschitz property and monotonicity. A local filter with additive error a is a randomized algorithm that is given black-box access to a function f and a query point x in the domain of f. Its output is a value $F(x)$, such that (i) the reconstructed function $F(x)$ satisfies the property (in our case, is Lipschitz or monotone) and (ii) if the input function f satisfies the property then for every point x in the domain, with high constant probability the reconstructed value $F(x)$ differs from $f(x)$ by at most a. Local filters were introduced by Saks and Seshadhri (SICOMP 2010). The relaxed definition we study is due to Bhattacharyya et al. (RANDOM 2010), except that we further relax it by allowing additive error. Local filters for Lipschitz and monotone functions have several important applications to areas such as data privacy.

We show that every local filter for Lipschitz (resp., monotone) functions runs in time exponential in the dimension d in the worst case. This holds even for filters with significant additive error, as well as for both previously studied definitions. Prior lower bounds (for local filters with no additive error, i.e., with $a = 0$) applied only to more restrictive notions of filters, e.g., nonadaptive filters that are required to specify all their lookups in advance, before obtaining values of f on any points. To prove our lower bounds, we construct families of hard functions, and show that lookups of a local filter on these functions are captured by a combinatorial object that we call a c-connector. Then we present a lower bound on the maximum outdegree of a c-connector, and show that it implies the desired bounds on the running time

of local filters. Our lower bounds, in particular, imply the same bound on the running time for a class of privacy mechanisms.

Optimal Mechanisms for Selling Information

Moshe Babaioff, Robert Kleinberg and Renato Paes Leme

The buying and selling of information is taking place at a scale unprecedented in the history of commerce, thanks to the formation of online marketplaces for user data.Data providing agencies sell user information to advertisers to allow them to match ads to viewers more effectively. In this paper we study the design of optimal mechanisms for a monopolistic data provider to sell information to a buyer, in a model where both parties have (possibly correlated) private signals about a state of the world, and the buyer uses information learned from the seller, along with his own signal, to choose an action (e.g., displaying an ad) whose payoff depends on the state of the world.

We provide sufficient conditions under which there is a simple one-round protocol (i.e. a protocol where the buyer and seller each sends a single message, and there is a single money transfer) achieving optimal revenue. In these cases we present a polynomial-time algorithm that computes the optimal mechanism. Intriguingly, we show that multiple rounds of partial information disclosure (interleaved by payment to the seller) are sometimes necessary to achieve optimal revenue if the buyer is allowed to abort his interaction with the seller prematurely.

Node-weighted Network Design in Planar and Minor-closed Families of Graphs

Chandra Chekuri, Alina Ene and Ali Vakilian

We consider *node-weighted* network design in planar and minor-closed families of graphs. In particular we focus on the survivable network design problem (SNDP). The input consists of a node-weighted undirected graph $G = (V, E)$ and an integer connectivity requirements $r(uv)$ for each pair of nodes uv. The goal is to find a minium node-weighted subgraph H of G such that for each pair uv H contains $r(uv)$ disjoint paths between u and v. Three versions of the problem are edge-connectivity (EC-SNDP), element-connectivity SNDP (Elem-SNDP) and vertex-connectivity SNDP (VC-SNDP) depending on whether the path are required to be edge, element or vertex disjoint. Our main results are an $O(k)$-approximation algorithms for EC-SNDP and Elem-SNDP when the graph is planar; here $k = \max_{uv} r(uv)$ is the maximum requirement. This improves the $O(k \log n)$-approximation known for node-weighted EC-SNDP and Elem-SNDP in general graphs [**?**]. Our results are inpsired by and generalize the results and ideas from the work of Demaine, Hajiaghayi and Klein [**?**] who considered node-weighted Steiner tree and Steiner forest problems in planar graphs and gave constant factor approximations for them, thereby overcoming the $\Omega(\log n)$-hardness that holds for general graphs.

Streaming Balanced Graph Partitioning for Random Graphs

Isabelle Stanton

A major systems problem is distributing graph data across a cluster of machines. While most graph computation systems currently use a random balanced cut (which nearly maximizes the communication volume), many support custom partitionings. Given the size of the data, we study streaming heuristics for distributing a graph onto a cluster of machines during the loading phase. We give a lower bound of o(n) for the approximation of the general problem but experimental results are surprisingly positive. We explain this performance by analyzing the best heuristic on a random graph model with a random stream ordering.

On Bitcoin and Red Balloons

Moshe Babaioff, Shahar Dobzinski, Sigal Oren and Aviv Zohar

Many large decentralized systems rely on information propagation to ensure their proper function. However, it is common that only participants that are aware of the information can compete for some reward, and thus informed participants have an incentive not to propagate information to others. One recent scenario in which such tension arises is the 2009 DARPA Network Challenge (finding red balloons). We focus on another prominent scenario: Bitcoin, a decentralized electronic currency system.

Bitcoin represents a radical new approach to monetary systems. It has been getting a large amount of public attention over the last year, both in policy discussions and in the popular press. Its cryptographic fundamentals have largely held up even as its usage has become increasingly widespread. We find, however, that it exhibits a fundamental problem of a different nature, based on how its incentives are structured. We propose a modification to the protocol that can eliminate this problem. Bitcoin relies on a peer-to-peer network to track transactions that are performed with the currency. For this purpose, every transaction a node learns about should be transmitted to its neighbors in the network. As the protocol is currently defined and implemented, it does not provide an incentive for nodes to broadcast transactions they are aware of. In fact, it provides a strong incentive not to do so. Our solution is to augment the protocol with a scheme that rewards information propagation. Since clones are easy to create in the Bitcoin system, an important feature of our scheme is Sybil-proofness. We show that our proposed scheme succeeds in setting the correct incentives, that it is Sybil-proof, and that it requires only a small payment overhead, all this is achieved with iterated elimination of dominated strategies. We complement this result by showing that there are no reward schemes in which information propagation and no self-cloning is a dominant strategy.

Matrix Lie Algebra Isomorphism

Joshua A. Grochow

We study the problem of Matrix Isomorphism of Matrix Lie Algebras. Lie algebras arise centrally in areas as diverse as differential equations, particle physics, group theory, and the Mulmuley–Sohoni Geometric Complexity Theory program. A matrix Lie algebra is a set L of matrices that is closed under linear combinations and the operation $[A, B] := AB - BA$. Two matrix Lie algebras are matrix isomorphic if there is an invertible change-of-basis matrix M such that $L = ML'M^{-1}$, that is, $L = \{MAM^{-1} : A \in L'\}$.

We show that certain cases of Matrix Isomorphism of Lie Algebras are equivalent to Graph Isomorphism. On the other hand, we give polynomial-time algorithms for other cases of Matrix Isomorphism of Lie Algebras, which allow us to essentially derandomize a recent result of Kayal on affine equivalence of polynomials. These hardness results and algorithms give a fairly complete picture of the complexity of this problem for the broad and broadly studied classes of abelian, semisimple, and completely reducible matrix Lie algebras.

Online Mixed Packing and Covering

Umang Bhaskar and Lisa Fleischer

In many problems, the inputs to the problem arrive over time, and must be dealt with irrevocably when they arrive. Such problems are online problems. A common method of solving online problems is to first solve the corresponding linear program online, and then round the fractional solution obtained. We give algorithms for solving mixed packing and covering linear programs, when the covering constraints arrive online. No prior sublinear competitive algorithms are known for this problem. We give the first such — a polylogarithmic-competitive algorithm for mixed packing and covering online. We also show a nearly tight lower bound.

We apply our techniques to solve two online fixed-charge problems with congestion. These problems are motivated by applications in machine scheduling and facility location. The linear program for these problems is more complicated than mixed packing and covering, and presents unique challenges.

We show that our techniques combined with a randomized rounding procedure give polylogarithmic-competitive integral solutions. These problems generalize online set-cover, for which there is a polylogarithmic lower bound. Hence, our results are close to tight.

Privacy in Sparse, High-dimensional Learning Problems

Daniel Kifer, Adam Smith and Abhradeep Guha Thakurta

We consider differentially private algorithms for convex empirical risk minimization. Differential privacy (Dwork et al., 2006) is a recently introduced notion of privacy which guarantees that an algorithm's output does not depend on the data of any individual in the data set. This is crucial in fields, such as genomics, collaborative filtering, and economics, that handle sensitive data. Our motivation is the design of private algorithms for sparse learning problems, in which one aims to find solutions (e.g., regression parameters) with few non-zero coefficients. To this end:

(a) We significantly extend the analysis of the "objective perturbation" algorithm of (Chaudhuri et al., 2010) for convex ERM problems. A principle tool in our analysis is a new nontrivial limit theorem for differential privacy. (b) We give the first private algorithms for sparse regression problems in high-dimensional settings, where p is much larger than n. We analyze their performance for linear regression: under standard assumptions on the data, our algorithms have vanishing empirical risk for $n = poly(s, \log p)$ when there exists a good regression vector with s nonzero entries.

Tight Bounds on Proper Equivalence Query Learning of DNF

Lisa Hellerstein, Devorah Kletenik, Linda Sellie and Rocco Servedio

We prove a new structural lemma for partial Boolean functions f, which we call the seed lemma for DNF. Using the lemma, we give the first subexponential algorithm for proper learning of poly(n)-term DNF in Angluin's Equivalence Query (EQ) model. The algorithm has time and query complexity $2^{(\tilde{O}\sqrt{n})}$, which is optimal.

We also give a new result on certificates for DNF-size, a simple algorithm for properly PAC-learning DNF, and new results on EQ-learning $\log n$-term DNF and decision trees.

Approximately Revenue-Maximizing Auctions for Deliberative Agents

L. Elisa Celis, Anna R. Karlin, Kevin Leyton-Brown, C. Thach Nguyen and David R. M. Thompson

In many real-world auctions, a bidder may not know her exact value for an item, but can work to reduce her uncertainty. Relatively little is known about such deliberative environments. In this poster, we present a new approach that allows us to leverage classical revenue-maximization results in deliberative environments. In particular, we use Myerson's optimal algorithm to construct the first non-trivial (i.e., dependent on deliberation costs) upper bound on revenue in deliberative auctions. This bound allows us to apply existing approximation algorithms in the classical environment to a deliberative environment. In addition, we show that many deliberative environments, revenue-maximizing dominant strategy mechanisms are characterized by sequential posted price auctions. Please see `http://tinyurl.com/delib-agents` for a preliminary version of this work.

Tatonnement in Ongoing Markets of Complementary Goods

Yun Kuen Cheung, Richard Cole and Ashish Rastogi

This paper continues the study,initiated by Cole and Fleischer, of the behavior of a tatonnement price update rule in Ongoing Fisher Markets. The prior work showed fast convergence toward an equilibrium when the goods satisfied the weak gross substitutes property and had bounded demand and income elasticities. The current work shows that fast convergence also occurs for the following types of markets

when using the same simple price update rule: all pairs of goods are complements to each other, and the demand and income elasticities are suitably bounded. In particular, these conditions hold when all buyers in the market are equipped with CES utilities, where all the parameters ρ, one per buyer, satisfy $-1 < \rho \leq 0$. We also extend the above result to markets in which mixture of complements and substitutes occur, which include characterizing a class of nested CES utilities for which fast convergence holds.

An interesting technical contribution, which may be of independent interest, is an amortized analysis for handling asynchronous events in settings in which there are a mix of continuous changes and discrete events.

Finding Fair and Fast Resource Allocations

Junghwan Shin and Sanjiv Kapoor

We consider routing problems that determine bounded delay paths. The *min-max delay routing problem* is the problem of finding flows on a set of paths $\mathcal{P}_{s_i,t_i}$ from source s_i to sink t_i, $\forall i$, which together carry r_i units of flow and minimize the maximum delay, i.e. find $\mathcal{P}_{s_i,t_i}$ and an assignment of flow $f_{P_j^i}, \forall P_j^i \in \mathcal{P}_{s_i,t_i}$ s.t. $\sum_{P \in \mathcal{P}_{s_i,t_i}} f_P = r_i$ with the following objective: $\min \max_{i, P \in \mathcal{P}_{s_i,t_i}} \Phi(f(P))$. We define the related problem of *min-avg-max delay routing* which requires determining a set of paths to optimize the objective function: $\min \sum_i r_i \max_{P \in \mathcal{P}_{s_i,t_i}} \Phi(f(P))$. Routing flows is a special case of the general resource allocation problem, and we extend the definition to the case of resource allocation problem. We show that the resource allocation problem for homogenous functions has a *fair* min-avg-max allocation. That is, for convex homogeneous cost functions, an allocation x, that is an optimal solution to the min-avg-max cost allocation problem is also fair for each player. The solution is also a social optimal solution. Furthermore, given a multi-commodity network with convex homogeneous delay function, we show a convex program that determines a social optimum flow and provides a solution to the min-avg-max delay problem that is fair. We also devise polynomial sized linear programs to determine fair min-max flows in the case when delay functions are linear. We contrast with previous work by Dafermos and Sparrow [Dafermos and Sparrow 1969] that studies the relationship between flow that are at Nash Equilibrium and flows that optimize a social welfare function. The study of Nash equilibrium in the context was initialized by Pigou [Pigou 1943], and furthered by Wardrop [Wardrop 1952] and by Dafermos and Sparrow amongst others. Our results extend the work of Dafermos and Sparrow, who show that a flow is at Wardrop equilibrium if and only if it has minimal total cost when cost functions are of the form $c_e(f_e) = a_e x_e^b$ for a fixed b. However, our results apply to the general class of homogenous functions which are not considered by Dafermos and Sparrow, e.g. $c_e(f_e) = \sqrt{\sum_i (a_e^i f_e^i)^2}$. We also consider the resource allocation problem when the required resource sets are expressible as a matroid. We describe a combinatorial algorithm to achieve an ϵ-approximation for the min-avg-max cost matroid resource allocation problem.

Bicriteria Approximation for the Reordering Buffer Problem

Siddharth Barman, Shuchi Chawla and Seeun William Umboh

In the reordering buffer problem (RBP), a server is asked to process a sequence of requests lying in a metric space. In order to serve a request the server must move to the corresponding point in the metric. The requests can be processed slightly out of order; in particular, the server has a buffer of capacity k, and can "store" up to k requests in this buffer as it reads in the sequence of requests. The goal is to reorder the requests in such a manner that the buffer constraint is satisfied and the total travel cost of the server is minimized. The RBP arises in many applications that require scheduling with a limited buffer capacity, such as scheduling a disk arm in storage systems, switching colors in paint shops of a car manufacturing plant, and rendering 3D images in computer graphics.

We study the offline version of RBP and develop bicriteria approximations. When the underlying metric is a tree, we obtain a solution of cost no more that $9OPT$ using a buffer of capacity $4k + 1$ where OPT is

the cost of an optimal solution with buffer capacity k. Constant factor approximations were known previously only for the uniform metric (Avigdor-Elgrabli et al., 2012). Here we present the first constant-factor approximation for tree metrics. Via tree embeddings, this implies an $O(\log n)$ approximation to cost and $O(1)$ approximation to buffer size for general metrics, improving upon a $O(\log^2 k \log n)$ approximation by Englert al. (2007).

Finding Overlapping Communities in Social Networks: Toward a Rigorous Approach

Sanjeev Arora, Rong Ge, Sushant Sachdeva and Grant Schoenebeck

A community in a social network is usually understood to be a group of nodes more densely connected with each other than with the rest of the network. This is an important concept in most domains where networks arise: social, technological, biological, etc. For many years algorithms for finding communites implicitly assumed communities are nonoverlapping (leading to use of clustering-based approaches) but there is increasing interest in finding overlapping communities. A barrier to finding communities is that the solution concept is often defined in terms of an NP-complete problem such as Clique or Hierarchical Clustering.

This work seeks to initiate a rigorous approach to the problem of finding overlapping communities, where "rigorous" means that we clearly state the following: (a) the object sought by our algorithm (b) the assumptions about the underlying network (c) the (worst-case) running time. Our assumptions about the network lie between worst-case and average-case. An averagecase analysis would require a precise probabilistic model of the network, on which there is currently no consensus. However, some plausible assumptions about network parameters can be gleaned from a long body of work in the sociology community spanning five decades focusing on the study of individual communities and ego-centric networks (in graph theoretic terms, this is the subgraph induced on a node s neighborhood). Thus our assumptions are somewhat "local" in nature. Nevertheless they suffice to permit a rigorous analysis of running time of algorithms that recover global structure. Our algorithms use random sampling similar to that in property testing and algorithms for dense graphs. We note however that our networks are not necessarily dense graphs, not even in local neighborhoods. Our algorithms explore a local-global relationship between ego-centric and socio-centric networks that we hope will provide a fruitful framework for future work both in computer science and sociology.

Feasibility and Completeness of Cryptographic Tasks in the Quantum World

Jonathan Katz, Fang Song, Hong-Sheng Zhou and Vassilis Zikas

Cryptographic feasibility results can change drastically in going from the classical to the quantum world; for example, there exist quantum protocols for unconditionally secure key exchange, whereas classical protocols for such task (with information-theoretic security) are impossible. With this in mind, we study feasibility of quantum protocols for universally composable, two-party secure computation in both the information-theoretic and computational settings. We show that with respect to computational security feasibility results carry over unchanged from the classical to the quantum world: a functionality can be realized (without setup) against quantum adversaries iff it can be realized against classical adversaries; and a functionality is complete (i.e., can be used to realize arbitrary other functionalities) in the quantum world iff it is complete in the classical world. Along the way, we also prove the analogue of the Canetti-Fischlin result for quantum protocols: namely, that there exist functionalities that cannot be realized without some additional trusted setup.

In contrast, the situation in the information-theoretic setting is more complex and, in particular, there are functionalities that are complete in the quantum world but not in the classical case, e.g., commitment. We identify a few more complete functionalities in the quantum world, and simplify the classical landscape of two-party functionalities in the IT setting to only three families: feasible, complete and XOR-like.

Approximating the Exponential, the Lanczos Method and an $\tilde{O}(m)$-Time Spectral Algorithm for Balanced Separator

Lorenzo Orecchia

We give a novel spectral approximation algorithm for the balanced edge-separator problem that, given a graph G, a constant balance $b \in (0, 1/2]$, and a parameter γ, either finds an $\Omega(b)$-balanced cut of conductance $O(\sqrt{(\gamma)})$ in G, or outputs a certificate that all b-balanced cuts in G have conductance at least γ, and runs in time $\tilde{O}(m)$. This settles the question of designing asymptotically optimal spectral algorithms for balanced edge-separator. Our algorithm relies on a variant of the heat kernel random walk. The poster will emphasize the connections between random walks and primal-dual algorithms for solving semidefinite-programming formulations of the graph partitioning problem.

Our algorithm also requires a subroutine to compute $\exp(-L)v$ where L is the Laplacian of a graph related to G and v is a vector. Algorithms for computing this matrix-exponential-vector product efficiently comprise our next set of results. We give a new algorithm which computes a good approximation to $\exp(-A)v$ for a class of PSD matrices A and a given vector u, in time roughly $\tilde{O}(m_A)$, where m_A is the number of non-zero entries of A. This uses, in a non-trivial way, the result of Spielman and Teng on inverting SDD matrices in $\tilde{O}(m_A)$ time. The poster will highlight the ideas from Approximation Theory and Iterative Methods that allow us to achieve this result, together with some further applications to graph partitioning.

Testing and learning submodular functions

Sofya Raskhodnikova and Grigory Yaroslavtsev

We present algorithms with polynomial query complexity for learning and testing submodular functions with integral range of bounded size. Our work answers one of the open problems posed by Seshadhri and Vondrak (ICS 2011), who gave the first algorithm for testing submodular functions with subexponential running time.

A Near-Linear Time ϵ-Approximation Algorithm for Geometric Bipartite Matching

Pankaj Agarwal and R. Sharathkumar

For point sets $A, B \subset R^d, |A| = |B| = n$, and for a parameter $\epsilon > 0$, we present an algorithm that computes, in $O(n\text{poly}(\log n, 1/\epsilon))$ time, an ϵ-approximate perfect matching of A and B with high probability; the previously best known algorithm takes $\Omega(n^{3/2}))$ time. We approximate the L_p-norm using a distance function, $d(\cdot, \cdot)$, based on a randomly shifted quad-tree. The algorithm iteratively generates an approximate minimum-cost augmenting path under $d(\cdot, \cdot)$ in time proportional to the length of the path. We show that the total length of the augmenting paths generated by the algorithm is $O((n/\epsilon)\log n)$, implying a near-linear running time of our algorithm.

Cutting Spending by adding options: Getting out of an Obligation by Providing a Menu of Cheaper Alternatives

Vincent Conitzer, Janardhan Kulkarni, Kamesh Munagala and Xioming Xu

Graph maintenance problems and churn complexity for distributed overlays

Lucas T. Cook

The study of distributed algorithms in dynamic networks has become popular, typically focusing on edge dynamics controlled by an unknown adversary. For distributed overlay networks – in which processes build a subgraph on top of a previously known topology – node dynamics are the core issue in the problem of characterizing *churn*, the process of nodes joining and leaving the network. We present

a family of online *graph maintenance problems* for distributed overlay networks, and introduce the *churn complexity* as the maximum rate of node churn at which these problems are solvable by any distributed protocol. Using known bounds from centralized complexity and lower bounds for distributed coloring, we are able to bound the churn rate for a broad class of practical overlay protocols that resemble the distributed hash table Chord. For the graph maintenance problems that simulate routing overlays, we conjecture that the churn complexity is mostly independent of the routing procedure used.

Analyzing Graph Connectivity via Random Linear Projections

Kook Jin Ahn, Sudipto Guha and Andrew McGregor

In this poster, we present a sequence of algorithmic results for analyzing graph connectivity via random linear projections or "sketches". We start with results for evaluating basic connectivity and k-connectivity and then use these primitives to construct combinatorial sparsifiers that allow every cut to be approximated up to a factor $(1+\epsilon)$. Our results have numerous applications including single-pass, semi-streaming algorithms for constructing sparsifiers in fully-dynamic graph streams where edges can be added and deleted in the underlying graph.

The results presented have appeared in SODA 2012 and PODS 2012.

Structure from Local Optima: Factoring Distributions and Learning Subspace Juntas

Santosh Vempala and Ying Xiao

Independent Component Analysis (ICA), a well-known model in statistics, assumes that data is generated by applying an a ne transformation to a product distribution, and aims to recover the orthogonal basis corresponding to the product distribution. We consider a generalization of ICA, wherein the data is generated as an a ne transformation applied to a product of distributions on two orthogonal subspaces, and the goal is to recover the two component subspaces. Our main result, extending the work of Frieze, Jerrum and Kannan, is an algorithm for generalized ICA that uses local optima of high moments and recovers the component subspaces; when one of the components is on a k-dimensional relevant subspace and satises some mild assumptions while the other is noise, modeled as an $n-k$-dimensional Gaussian, then the complexity of the algorithm is $T(k;)+\text{poly}(n)$ where T depends only on the k-dimensional distribution. We apply this result to learning a k-subspace junta, i.e., an unknown $\{0,1\}$ function in R^n determined by an unknown k-dimensional subspace. This is a common generalization of the problems of learning an unknown k-junta in R^n and of learning an intersection of k halfspaces in R^n, two important problems in learning theory. Our main tools are the use of local optima to recover global structure, a gradient-based algorithm for optimization over tensors, and an approximate version of the Schwartz-Zippel polynomial identity test. Together they signicantly extend the analysis of ICA and tensor-PCA, and the class of k-dimensional labeling functions that can be learned efficiently.

Limits of Random Oracle in Secure Computation

Mohammad Mahmoody, Hemanta Maji and Manoj Prabhakaran

The seminal result of Impagliazzo and Rudich (STOC 1989) gave a black-box separation between one-way functions and public-key encryption: informally, a public-key encryption scheme cannot be constructed using one-way functions as the sole source of computational hardness. In addition, this implied a black-box separation between one-way functions and protocols for certain Secure Function Evaluation (SFE) functionalities (in particular, Oblivious Transfer). Surprisingly, however, *since then there has been no further progress in separating one-way functions and SFE functionalities* (though several other black-box separation results were shown). In this work, we present the complete picture for deterministic 2-party SFE functionalities. We show that one-way functions are black-box separated from *all such SFE functionalities*, except the ones which have unconditionally secure protocols (and hence do not rely on any computational hardness), when secure computation against semi-honest adversaries is considered. In the case of

security against active adversaries, a black-box one-way function is indeed useful for SFE, but we show that it is useful only as much as access to an ideal commitment functionality is useful.

Technically, our main result establishes the limitations of random oracles for secure computation. We show that a two-party deterministic functionality f has a secure function evaluation protocol in the random oracle model that is (statistically) secure against semi-honest adversaries if and only if f has a protocol *in the plain model* that is (perfectly) secure against semi-honest adversaries. Further, in the setting of active adversaries, a deterministic SFE functionality f has a (UC or standalone) statistically secure protocol in the random oracle model if and only if f has a (UC or standalone) statistically secure protocol in the commitment-hybrid model.

Our proof is based on a "frontier analysis" of two-party protocols, combining it with (extensions of) the "independence learners" of Impagliazzo-Rudich/Barak-Mahmoody. We make essential use of a combinatorial property, originally discovered by Kushilevitz (FOCS 1989), of functions that have semi-honest secure protocols in the plain model (and hence our analysis applies only to functions of polynomial-sized domains, for which such a combinatorial characterization is known).

Minimax Option Pricing Meets Black-Scholes in the Limit

Jacob Abernethy, Rafael Frongillo and Andre Wibisono

Option contracts are a type of financial derivative that allow investors to hedge risk and speculate on the volatility of an asset's future market price. In short, an option has a particular payout that is based on the market price for an asset on a given date in the future. In 1973, Black and Scholes proposed a valuation model that gives a "fair price" for an option under the assumption that the price fluctuates according to geometric Brownian motion (GBM). Black and Scholes provided a continuous-time trading strategy for an investor, known as a "replication strategy" or "hedging strategy", which allows the investor to buy and sell the asset in order to "replicate" the option's payoff.

In this poster we'll look at the the design of replication strategies, and hence a pricing mechanism, for options and other derivatives that does not rely on the GBM assumption. Indeed, we shall address the following question: what can an investor achieve when the asset's price path is chosen by... an adversary? What we show is that, even under these worst-case conditions, we ultimately recover the Black Scholes pricing model but without the GBM assumption.

Index