ISSN 2186-7437

NII Shonan Meeting Report

No. 221

Algorithmic tools and their applications in emerging models of computation

Guy Even
Robert Krauthgamer
Gregory Schwartzman

March 24 – 28, 2025

National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda-Ku, Tokyo, Japan

Description of the Meeting

The algorithms community has made significant strides in the development and advancement of various algorithmic tools. These tools have proven to be essential in solving complex computational problems efficiently and effectively in specific settings and computational models. We held a Shonan meeting that brought together experts from diverse disciplines, aiming to distill these algorithmic tools and expand their applications to various emerging computational models, including parallel computing, distributed computing, data-stream algorithms and sampling algorithms.

Algorithmic Tools

1. **Interior Point Methods (IPM):** Interior point methods have revolutionized the field of convex optimization by providing efficient solutions to linear and nonlinear programming problems. These methods, combined with the Laplacian Paradigm and the development of special data structures, have led to dramatic improvements in classical combinatorial problems such as maximum flow and min-cost flow. During the meeting, we explored the applicability of interior point methods in emerging models of computation and discussed their potential in solving optimization problems in these models.

2. **Random Projections:** Random projections have gained attention as a very effective method to reduce the data's dimensionality while preserving its structure. We investigated the applications of random projections in emerging models of computation, including their impact on data-analysis tasks such as clustering, classification, and nearest neighbor search.

3. **Graph Decompositions:** Graph decomposition techniques have played a major role in the design of parallel, distributed, online, and dynamic algorithms. In recent years, expander decompositions have been key to the development of new algorithms for Laplacian solvers, for dynamic graph problems, and for routing. We explored the potential of decomposition methods in emerging models of computation and discussed their efficiency and scalability in these models.

4. **Smoothed Analysis:** Smoothed analysis is a framework that combines worst-case analysis with random perturbations to provide a more realistic understanding of algorithmic performance. We investigated the applicability of smoothed analysis in emerging models of computation and discussed how it can help analyze the robustness and stability of algorithms in these models.

5. **Computing with Advice (from Machine Learning):** The remarkable recent success of machine learning, including deep learning, has spurred interest in algorithms that can leverage advice provided by machine-learning models. We explored the possibilities of computing with machine-learning advice in streaming, parallel, distributed, and quantum computing and discussed how these models can enhance the efficiency and effectiveness of algorithms in various applications.

Computational Models

The meeting also addressed several computational models:

1. **Distributed Computing:** In distributed computing, every node in a network has a local input and communicates via network links in synchronous rounds. The goal is to compute local outputs that satisfy a common global objective. For example, computing a minimum spanning tree in the network requires that each node knows only which of its incident links belong to the computed tree. The goal in distributed algorithms is to design algorithms with a minimum number of communication rounds.

2. **Data Structures:** Data is organized in a way that facilitates efficient operations, such as insertions, deletions, and queries. The challenge is twofold: to develop data structures that require as little space as possible and to support operations as quickly as possible. Locality is captured by the fact that memory is accessed in blocks, and subsequent accesses are faster if the data is in a cache. The challenge is addressed via algorithms providing upper bounds and lower bounds on space usage and running time.

3. **Data-Stream Model:** In the data-stream model, an algorithm can only read the input sequentially and has bounded space, making it impractical to store much of the input. This model is especially relevant in communication networks, where routers must monitor massive amounts of information with limited memory. The challenge is to design algorithms that deduce interesting properties of the input despite space limitations.

4. **Parallel Computing:** Parallel computing leverages multiple processors or cores to execute computations simultaneously, significantly speeding up execution time. By dividing the workload across multiple processors, many tasks are executed with improved performance, which is especially important given the slowdown in Moore's law. Efficient communication and synchronization mechanisms are crucial for optimal performance.

The meeting brought together leading researchers from the aforementioned fields, who shared their recent results, insights, and vision. New ideas and collaborations emerged from the discussions dedicated to the topic "algorithmic tools and their applications to emerging models of computation." Participants from leading institutions in Israel, Japan, Europe, and the USA enriched the dialogue and helped strengthen the academic connection between these international research communities – especially between Japanese and non-Japanese researchers.

We believe that the diverse mix of researchers had a very positive impact on these fields, as ideas and techniques from one subfield were successfully applied to related subfields.

Meeting Schedule

Check-in day (March 23rd)

- Welcome banquet

Day 1 (March 24th)

- Invited talk by Peter Macgregor
- Lecture session 1
- Group photo shooting
- Lecture session 2
- Free discussion

Day 2 (March 25th)

- Invited talk by Arie Matsliah
- Lecture session 1
- Lecture session 2
- Free discussion

Day 3 (March 26th)

- Invited talk by Sayan Bhattacharya
- Lecture session 1
- Excursion
- Main banquet

Day 4 (March 27th)

- Invited talk by Nikhil Bansal
- Lecture session 1
- Lecture session 2
- Free discussion

Day 5 (March 28th)

- Invited talk by Jelani Nelson
- Lecture session
- Wrap up

Overview of Talks

Modern Spectral Clustering: Recent Progress and Future Directions

Peter Macgregor

Spectral clustering has a long history as a theoretically grounded, general-purpose clustering algorithm. While it performs well in practice, its adoption has been limited by computational expense and interpretability challenges. Recent work has developed modified spectral clustering algorithms–using techniques such as sparsification, dimensionality reduction, and random sampling–that improve running time while preserving performance guarantees. This talk presented these advancements and argued that it is time to challenge k-means' dominance.

Vizing's Theorem in Near-Linear Time

Sayan Bhattacharya

A textbook theorem by Vizing from the 1960s states that any graph with maximum degree Δ admits an edge coloring using only $(\Delta+1)$ colors. This talk presented the first algorithm achieving such a $(\Delta+1)$-edge coloring in near-linear time, specifically in $O(m \log \Delta)$ time, thereby matching longstanding bounds for special cases.

Connectomics: A Playground for Graph Algorithms

Arie Matsliah

Representing the brain as a graph–where vertices are neurons and edges are synaptic connections–introduces a host of computational challenges. This talk explored the enormous scale of the newly published fly connectome and discussed open problems in graph clustering, embedding, sorting, and matching, as well as the biological significance of these challenges.

The Power of Two Choices: Beyond Greedy Strategies

Nikhil Bansal

The power of two choices paradigm improves load balancing by selecting the less loaded option among two random choices. This talk analyzed situations where a purely greedy strategy may be suboptimal–especially when deletions or selection constraints come into play–and introduced new strategies that approach optimal performance.

New local differentially private protocols for frequency and mean estimation

Jelani Nelson

Consider the following examples of distributed applications: a texting app wants to train ML models for autocomplete based on text history residing on-device across millions of devices, or the developers of some other app want to understand common app settings by their users. In both cases, and many others, a third party wants to understand something in the aggregate about a large distributed database but under the constraint that each individual record requires some guarantee of privacy. Protocols satisfying so-called local differential privacy have become the gold standard for guaranteeing privacy in such situations, and in this talk I will discuss new such protocols for two of the most common problems that require solutions in this framework: frequency estimation, and mean estimation.

Based on joint works with subsets of Hilal Asi, Vitaly Feldman, Huy Le Nguyen, and Kunal Talwar.

Self-Stabilizing Algorithms for Mobile Agents

Yuichi Sudo

Recent studies have shown that self-stabilizing algorithms for mobile agents require techniques distinct from those used in standard distributed systems. This talk discussed these differences and presented open problems in the context of self-stabilizing graph exploration.

MPC: On Two Problems Resisting Round Compression

Slobodan Mitrović

Round compression is an approach for partitioning graphs in the Massively Parallel Computation (MPC) setting, proving effective for tasks such as maximal matching and vertex cover. This talk revisited round compression, highlighting its strengths as well as its limitations when applied to problems like set cover and matching in 3-rank hypergraphs.

Data Dependent Framework for Data Structures

Alex Andoni

This talk introduced a framework for designing efficient data structures for high-dimensional pattern-matching problems. By leveraging efficient communication protocols under product distributions, the method achieves improvements over traditional, data-oblivious techniques such as JL dimension reduction.

Coreset Spectral Clustering

Ben Jourdan

Understanding community structure in large graphs is crucial in modern data analysis. This talk exploited the connection between the Normalised Cut and Kernel k-means problems, applying coreset techniques to spectral clustering. The presentation detailed theoretical results and outlined promising directions for future research.

The Surprising Graph Behind k-Nearest Neighbors and Contrastive Learning

Orr Fisher

This talk delved into contrastive learning, where comparisons among triplets of data points determine relative similarity. It explored the dimensionality required to satisfy contrastive constraints and to preserve k-nearest neighbor relationships, combining combinatorial and algebraic techniques to analyze the underlying constraint graph.

Dynamic Graph Partitioning: Offline and Online

Marcin Bienkowski

Dynamic graph partitioning involves adapting clusters as the edge set of a graph changes over time. This talk surveyed recent advances in both dynamic bisection and dynamic linear arrangement, discussing the trade-offs between the cost of serving edges and the cost of re-clustering in offline and online settings.

Neuronal Cell Type Detection from Connectivity

Gregory Schwartzman

The accurate identification of neuronal cell types is fundamental to understanding neural circuit function. Traditional methods rely on morphological analysis, which compares neuronal shapes to classify cell types. However, these approaches can be time-consuming and may lack robustness, especially when neurons exhibit diverse morphologies. Additionally, accurately mapping mirror cells across hemispheres remains a challenge. In this study, we present an automated method for cell type detection based solely on neuronal connectivity patterns.

Competitive Vertex Recoloring

Boaz Patt-Shamir

Motivated by the challenge of job placement in physical machines, this talk introduced the problem of online vertex recoloring. Here, vertices (representing jobs) are re-colored in response to dynamically arriving edges, with the aim of minimizing the competitive ratio between the online algorithm's cost and that of an optimal offline solution.

Cut Sparsification and Succinct Representation of Submodular Hypergraphs

Robert Krauthgamer

Cut sparsification seeks to approximate every cut of a hypergraph within a $1 \pm \epsilon$ factor using far fewer hyperedges. This talk presented new bounds for sparsifying submodular hypergraphs–including the first polynomial bound for all submodular functions–and discussed methods for achieving succinct representations of reweighted sub-hypergraphs.

Graph Coloring Across Models

Magnus Halldorsson

This talk explored the convergence of ideas in randomized graph coloring across various computational models, including dynamic, streaming, MPC, distributed, and local algorithms. Several favorite open problems were highlighted, illustrating the rich interplay between these diverse approaches.

On Sparse Partitions

Arnold Filtser

A (σ, τ, Δ)-sparse partition of a metric space guarantees that each cluster has a diameter at most Δ and that every ball of radius Δ/σ intersects only a limited number of clusters. This talk examined the construction of such partitions and their diverse applications in algorithm design.

Algorithmic Stability and Connections to Other Computational Models

Yuichi Yoshida

I introduce the concept of algorithmic sensitivity for graph problems, which quantifies how the output changes–in terms of Hamming distance–when a single edge is deleted from the input graph. Beyond its intrinsic interest, I highlight connections between sensitivity and a range of topics, including statistical physics, differential privacy, probabilistically checkable proofs (PCP), and distributed algorithms. I conclude with a discussion of several open problems in this emerging area.

Clustering Oracles

Michael Kapralov

Clustering oracles are compact data structures that provide consistent query access to a graph clustering. This talk reviewed recent progress on designing clustering oracles with small space and fast query time, their applications to problems such as MAX-CUT, and discussed several promising open research directions.

Approximately Counting and Sampling Hamiltonian Motifs on Sublinear Time

Talya Eden

Counting small subgraphs (motifs) in large graphs is a fundamental challenge in graph analysis. This talk introduced a novel algorithm for approximately counting and nearly uniformly sampling Hamiltonian motifs in sublinear time, marking an important step toward bridging the gap between standard and augmented query models. The work is in collaboration with Reut Levi, Dana Ron, and Ronitt Rubinfeld.

Streaming Graph Algorithms in the Massively Parallel Computation Model

Artur Czumaj

We discuss a recently introduced model of graph algorithms in the streaming setting on massive distributed and parallel systems inspired by practical data processing systems. The objective is to design algorithms that can efficiently process evolving graphs via large batches of edge insertions and deletions using as little memory as possible. We will survey the key challenges and key advances in such setting, focusing on the nowadays canonical model for the study of theoretical algorithms for massive networks, the Massively Parallel Computation (MPC) model. We design MPC algorithms that efficiently process evolving graphs: in a constant number of rounds they can handle large batches of edge updates for problems such as connectivity, minimum spanning forest, and approximate matching while adhering to the most restrictive memory regime, in which the local memory per machine is strongly sublinear in the number of vertices and the total memory is sublinear in the graph sizes. Joint work with Gopinath Mishra and Anish Mukherjee.

Advances in All-Pairs Max-Flow

Ohad Trabelsi

The All-Pairs Max-Flow problem has attracted considerable attention over the past decade, particularly in the undirected, edge-capacitated setting, culminating in a near-linear time algorithm for the Gomory-Hu problem. However, the problem remains wide open in many other settings. In this talk, I will survey known results in these different settings and highlight some of the most intriguing open problems in the area.

On the Power of Graphical Reconfigurable Circuits

Yuval Emek

In this talk, we introduce the graphical reconfigurable circuits (GRC) model as an abstraction for distributed graph algorithms whose communication scheme is based on local mechanisms that collectively construct long-range reconfigurable channels (this is an extension to general graphs of a distributed computational model recently introduced by Feldmann et al. (JCB 2022) for hexagonal grids). The crux of the GRC model lies in its modest assumptions: (1) the individual nodes are computationally weak, with state space bounded independently of any global graph parameter; and (2) the reconfigurable communication channels are highly restrictive, only carrying information-less signals (a.k.a. beeps). Despite these modest assumptions, we prove that GRC algorithms can solve many important distributed tasks efficiently, i.e., in polylogarithmic time. On the negative side, we establish various runtime lower bounds, proving that for other tasks, GRC algorithms (if they exist) are doomed to be slow.

The talk is based on a recent joint work with Yuval Gil and Noga Harlev (DISC 2024), which belongs to a more general research agenda dedicated to uniform distributed computational models.

Locally Uniform Hashing

Ioana Bercea

Hashing is a commonly used technique for getting improved algorithms. Most analyses, however, assume access to fully-random hash functions, which is unrealistic in terms of the resources available to the algorithm. A fundamental line of work has thus been to design realistic hash functions (with small space and fast evaluation time) that make the algorithm perform almost as if it used fully-random hash functions. In this talk, we will review one such hash function, called a tornado tabulation hash function, that provides state-of-the-art randomness guarantees for several fundamental algorithms. In particular, we will define and discuss one key property that it exhibits, that of local uniformity. only used technique for getting improved algorithms. Most analyses, however, assume access to fully-random hash functions, which is unrealistic in terms of the resources available to the algorithm. A fundamental line of work has thus been to design realistic hash functions (with small space and fast evaluation time) that make the algorithm perform almost as if it used fully-random hash functions. In this talk, we will review one such hash function, called a tornado tabulation hash function, that provides state-of-the-art randomness guarantees for several fundamental algorithms. In particular, we will define and discuss one key property that it exhibits, that of local uniformity.

Seffi Naor

Non Standard Caching Models

Caching is a focal point within the field of online algorithms and many modern techniques in online algorithms have emerged via the study of caching algorithms. I will consider in this talk three non-standard caching model that are both interesting theoretically and well-grounded in practice: Writeback-aware caching, Block-aware caching, Non-linear caching. From each model we will learn an interesting lesson regarding formulations, designing competitive online algorithms, and rounding fractional solutions into randomized (integral) algorithms.

Yannic Maus

Sparsification for communication-efficient distributed symmetry-breaking

We present a sparsification method for designing communication-efficient distributed algorithms, particularly for graph coloring. This method builds on the constructive Lovász Local Lemma, a central topic in ongoing research.

Bernhard Haeupler

Length-Constrained Expanders and Graph Decompositions

Graph decompositions are ubiquitous algorithmic tools with many applications such as graph algorithms in various settings (sequential, parallel, distributed, dynamic, or streaming). They also give rise to simpler structures for approximating graphs such as spanners, hop-sets and shortcuts, oblivious routing, metric embeddings, vertex sparsifiers. Length-constrained expander decompositions are a recent versatile and powerful and generalization of both - low-diameter decomposition, which captures l_1-quantities like lengths and costs, and - expander decomposition, which captures l_infinity-quantities like flows and congestion.

Length-constrained expander decompositions significantly extend the range of applications of graph decomposition techniques and have already led to multiple breakthrough results.

This talk will give an intuitive introduction to various aspects of length-constrained expanders and their applications.

Discussion and Outcomes

The meeting provided a dynamic platform for discussing the synergy between algorithmic tools and emerging computational models. Over the course of the seminar, numerous cross-cutting themes and challenges emerged, giving rise to fruitful discussions and future research directions. Below we summarize the key outcomes and insights from these exchanges.

Algorithmic Techniques in New Contexts

A central point of discussion was how classical algorithmic tools can be adapted and extended for modern computational environments. Graph decomposition techniques, including expander and low-diameter decompositions, were a prominent topic. Several talks, notably by Bernhard Haeupler, highlighted their versatility across models. The concept of length-constrained expander decompositions stood out as a promising unifying structure that might enable a new generation of hybrid algorithms.

Random projections and dimensionality reduction techniques were also extensively discussed, particularly in the context of data streams and machine learning. There is growing interest in understanding how these techniques can be used to preserve privacy (via differential privacy) while still allowing meaningful downstream computations in federated learning environments.

Emerging Challenges and Research Directions

One of the strongest recurring themes was the increased role of data-driven advice in traditional algorithmic pipelines. Talks on learning-augmented algorithms, particularly those using ML models to offer "advice" in streaming and parallel models, inspired debate on theoretical frameworks for measuring such algorithms' robustness and adaptivity. Discussions focused on understanding when and how learning-based advice can be trusted and the cost of incorrect advice in worst-case scenarios.

Another direction of interest involved algorithmic stability and sensitivity, as introduced by Yuichi Yoshida. The idea that small changes in input — such as deletion of a single edge — could drastically alter the algorithm's output sparked broader conversations about resilience, fault tolerance, and the interplay between sensitivity and computational models. These insights could have practical ramifications in designing distributed algorithms for unreliable networks or adversarial settings.

The application of graph algorithms to neuroscience, such as in the talk by Arie Matsliah, also prompted intriguing discussions about the interdisciplinary potential of algorithms. Actually, there was some much interest in the topic, that we scheduled an additional session on the topic. With biological data becoming increasingly available in graph-like forms, many participants expressed interest in extending classical graph mining techniques — like clustering, motif detection, and embedding — to domains where correctness and interpretability are as important as efficiency.

Collaborations and Outcomes

The format of the meeting, with its emphasis on extended discussions, fostered collaborations between researchers from different subfields. Several participants reported initiating joint work or planning follow-up visits to explore new problems identified during the seminar. For example, the intersection of streaming algorithms and MPC received heightened attention, with multiple groups discussing possible unification of techniques and memory-efficient protocols for evolving graphs.

There was also enthusiasm for establishing benchmarks and shared datasets to evaluate the effectiveness of algorithmic tools across models. Such benchmarking could facilitate the comparison of, for instance, dynamic spectral clustering when implemented in both streaming and parallel settings.

Conclusion

Overall, the seminar surfaced key insights into how classical algorithmic tools must evolve in response to the realities of emerging computational models. From theoretical investigations into sensitivity and advice-based computation to practical applications in privacy, neuroscience, and large-scale systems, the discussions highlighted the richness of the field and the necessity of collaboration across domains. This seminar not only charted immediate research directions but also laid the groundwork for a broader agenda centered on scalable, adaptable, and robust algorithm design.

List of Participants

- Prof. Guy Even, Tel Aviv University
- Prof. Robert Krauthgamer, Weizmann institute of science
- Prof. Gregory Schwartzman, JAIST
- Prof. Talya Eden, Bar-Ilan University
- Prof. Yuichi Sudo, Hosei University
- Prof. Yuval Emek, Technion – Israel Institute of Technology
- Prof. Peter Macgregor, University of St Andrews
- Prof. Magnus M Halldorsson, Reykjavik University
- Prof. Yuichi Yoshida, National Institute of Informatics
- Mr. Benjamin Jourdan, University of Edinburgh
- Prof. Artur Czumaj, University of Warwick
- Prof. Jelani Nelson, University of California, Berkeley
- Prof. Boaz Patt-Shamir, Tel Aviv University
- Prof. Yannic Maus, Graz University of Technology
- Prof. Joseph Seffi Naor, Technion – Israel Institute of Technology
- Prof. Arnold Filtser, Bar Ilan University
- Prof. Slobodan Mitrovic, University of California, Davis
- Prof. Nikhil Bansal, University of Michigan
- Dr. Orr Fischer, Bar-Ilan University
- Prof. Michael Kapralov, École Polytechnique Fédérale de Lausanne (EPFL)
- Prof. Ioana Oriana Bercea, KTH Royal Institute of Technology
- Prof. Marcin Bienkowski, University of Wrocław
- Prof. Dennis Olivetti, Gran Sasso Science Institute
- Prof. Alkida Balliu, Gran Sasso Science Institute
- Prof. Sayan Bhattacharya, University of Warwick
- Prof. Sebastian Brandt, CISPA Helmholtz Center for Information Security
- Dr. Arie Matsliah, The Princeton Neuroscience Institute
- Prof. Bernhard Haeupler, ETH Zurich & Carnegie Mellon University & INSAIT
- Prof. Satoru Iwata, University of Tokyo

- Prof. Taisuke Izumi, Osaka University
- Dr. Francesco d'Amore, Gran Sasso Science Institute
- Prof. Alexandr Andoni, Columbia University
- Prof. Ohad Trabelsi, Toyota Technological Institute