Laboratory for Information and Decision Systems

The Laboratory for Information and Decision Systems (LIDS) is an interdepartmental laboratory staffed by faculty, research scientists, and graduate students from several departments and centers across MIT. LIDS provides an intellectually cohesive and collaborative environment that fosters high-quality, forward-looking research, and instills in our students the disciplinary depth and interdisciplinary understanding required of research and engineering leaders of today and tomorrow. In 2019, the lab celebrated its 80th anniversary.

Participants and Collaborations

The faculty members within LIDS are principally drawn from the departments of Electrical Engineering and Computer Science (EECS), Aeronautics and Astronautics (AeroAstro), and Civil and Environmental Engineering (CEE). However, LIDS has long been interdisciplinary, and recent research foci, combined with the pervasiveness of the analytical methodologies advanced by LIDS researchers, has broadened our collaborative scope even further. Some of the many entities at MIT with which LIDS has a strong relationship include: the Computer Science and Artificial Intelligence Laboratory, the Research Laboratory of Electronics (RLE), the Operations Research Center (ORC), the Departments of Brain and Cognitive Sciences, Economics, Mathematics, Mechanical Engineering (MechE), and the Sloan School of Management.

LIDS is also a key component of the Institute for Data, Systems, and Society (IDSS), which in addition to advancing disciplinary methodologies in statistics, data science, and information and decision systems, nurtures cross-cutting connections, especially with the social sciences, in order to address complex societal challenges. LIDS faculty play a pivotal role in defining the IDSS intellectual agenda—leading efforts in statistics and around flagship projects (e.g., in finance, autonomy, and smart cities), designing new academic programs (such as those in statistics and social and engineering systems), and being heavily involved in the search for new faculty in the fields of networks and in statistics. In particular, LIDS faculty have played a central role in launching the IDSS MicroMasters in statistics and data science.

LIDS researchers continue to have great success in obtaining funding for our broad and deep research agenda, and continue to develop relationships with industrial organizations and national laboratories, including Lincoln Laboratory, NASA Jet Propulsion Laboratory, Ford Motor Company, MIT-IBM Watson AI Lab, IBM, SES S.A., Liberty Mutual Insurance Company, Lockheed Martin Corporation, and Accenture. Also, thanks to a rich history of research excellence and leadership, LIDS remains a magnet for the very best, attracting not only outstanding students, but also a continuous stream of world-leading researchers as visitors and collaborators.

LIDS has also been strengthened significantly this academic year, with Assistant Professor Cathy Wu, a new CEE hire, as well as Associate Professor Saurabh Amin (CEE) joining the lab. Professor Wu focuses on understanding and shaping the impact of autonomy on society, from a dynamical systems perspective. Research directions include sample-efficient reinforcement learning, coping with distribution shift, bridging machine learning and automation science, and applying automation science in the context of mobility. Professor Amin focuses on the design and implementation of network inspection and control algorithms for improving the resilience of large-scale critical infrastructures, such as transportation systems and water and energy distribution networks, against cyber-physical security attacks and natural events. In addition, a new EECS hire, Assistant Professor Ashia Wilson, will be joining the lab in January 2021. Her research activities center on optimization, with an emphasis on the context of learning algorithms, and broadened to related areas, such as statistical inference and fairness.

Additionally, LIDS has recently redefined the role of LIDS affiliate members, fostering stronger ties with a number of MIT faculty who, although not LIDS principal investigators (PIs), are intellectually aligned with the lab's research agenda and are engaged in substantial interactions. The current list of affiliate members is composed of Professors Anant Agarwal (EECS), Hamsa Balakrishnan (AeroAstro), Tamara Broderick (EECS), David Gamarnik (Sloan), Song Han (EECS), Stefanie Jegelka (EECS), and Youssef Marzouk (AeroAstro), as well as Senior Research Scientist Anuradha Annaswamy (MechE).

The already strong core of existing LIDS PIs, together with the extended community of affiliate members, have turned LIDS into a preeminent entity—both within MIT and more broadly in the academic world—in the fields of data science and the foundations of machine learning. At the same time, traditional LIDS core areas (communications, information theory, networks, optimization, control, and autonomy) remain active and strong.

Joining the Stephen A. Schwarzman College of Computing

LIDS officially joined the newly formed Schwarzman College of Computing as one of its research units. The intellectual mission of LIDS, with its emphasis on information, data, and decision making, falls squarely within the center of gravity at the college, and LIDS faculty are actively engaged in cross-cutting initiatives taking place under the umbrella of the college.

LIDS Intellectual Vision

The mission of LIDS is to develop and apply rigorous approaches and tools for modeling, analysis, design, and optimization of physical or artificial systems that process information and rely on information for decision making. Research in LIDS encompasses the development of new analytical methodologies, as well as the adaptation and application of advanced methods to specific contexts and application domains.

Many of the important recent technological advances involve systems that collect, exchange, and process data and information. Information is then exploited to make decisions, including statistical decisions, resource allocation decisions, or real-time control decisions. This schema captures much of what is happening in an impressive range of fields, such as robotics, autonomous systems, intelligent systems, machine learning, life science informatics, computer networks, societal infrastructures, electric power systems, and more. Advances in all of these domains are made in several labs and departments, and domain expertise is typically critical. At the same time the information-to-decisions viewpoint, often associated with the legacy of Norbert Wiener and Claude Shannon, rests on an intellectual core and on fundamental methodologies that can be applied across disciplines and domains. The objective of LIDS is to serve as a focal point for this intellectual core, while advancing work in select application domains that—in a virtuous cycle—also provide inspiration for further methodological research.

To achieve these aims, LIDS research is underpinned by the following:

- A set of core mathematical disciplines, including probability and statistics; dynamical systems; optimization and decision theory
- A set of core engineering disciplines, including inference, statistical data processing, data science, and machine learning; transmission of information; networks; systems and control
- A set of broad challenges in traditional and emerging applications of critical societal importance

The simultaneous efforts along each of these dimensions within the same lab leads to strong synergies: work in the mathematical disciplines leads to new methodologies that advance core engineering disciplines and interdisciplinary applied investigations; conversely, work on new interdisciplinary challenges provides the inspiration and direction for fundamental disciplinary research, as well as the charting of emerging new disciplines.

Research Areas

The lab's multiple research strands are usually cross-cutting and cannot be neatly organized into categories. Nevertheless, they can be broadly classified in terms of the following core areas.

Statistical Inference and Machine Learning

This area deals with complex systems, phenomena, and data that are subject to uncertainty and statistical variability. It also includes the development of large-scale data processing software systems. Research ranges from the development of basic theory, methodologies, algorithms, and computational infrastructures, to adaptations of this work for challenging applications in a broad array of fields. Typical applications involve causal inference in experimental design, social data processing, and e-commerce, as well as image processing, computer vision, and automation of data engineering. Other current topics include reinforcement learning and online optimization, recommendation systems, graphical models, large-scale software systems for data engineering, medical image processing, causal inference in genetics, and high-dimensional statistics.

Optimization

This area aims to develop analytical and computational methods for solving optimization problems in engineering, data science, and operations research, with applications in communication networks, control theory, power systems, machine learning, and computer-aided manufacturing. In addition to linear, nonlinear, dynamic, convex, and network programming, methods that exploit the algebraic structure of large-scale problems, as well as simulation-based methods are also studied.

Systems Theory, Control, and Autonomy

This area deals with all aspects of system identification, inference, estimation, control, and learning for feedback systems. Theoretical research includes quantification of fundamental capabilities and limitations of feedback systems, development of practical methods and algorithms for decision making under uncertainty, robot sensing and perception, inference and control over networks, as well as architecting and coordinating autonomy-enabled infrastructures for transportation, energy, and beyond.

Networks

This area includes communications, information theory, and networking, with applications to wireless systems, optical networks, and data centers. Research in this area includes the development of fundamental limits on communications systems, the design of optimal resource allocation schemes for wireless networks, and the design of optimal architectures and control algorithms for data centers and cloud networks. Additional recent directions include the analysis of social networks and of agent interactions in networked systems, with applications ranging from the analysis of data generated by large-scale social networks to the study of dynamics and risk in large interconnected financial, transportation, and power systems.

Moreover, the availability of increasingly capable sensing, communication, and computation enables the collection and transfer of large amounts of data pertaining to complex and heterogeneous interconnected systems, and opens up many new avenues for methodological research in all of the above areas, with some ubiquitous themes, such as data fusion, distributed learning, and decision making, as well as issues of scalability, robustness, and performance limits.

Some particular areas of significant recent activity that we wish to highlight are the following (see also the Faculty Activities section below for more details):

- Biological systems and biomedical data analysis
- High-performance, unmanned autonomous systems
- Energy systems analysis, economics, and design
- Human-level perception for robotics platforms
- Machine learning for recommendation systems and social media
- Network scheduling and routing
- High-dimensional inference in graphical models
- Networking and information transmission in the context of the internet of things
- Social network analysis and characterization
- Network navigation and localization
- Transportation network analysis, control, and design
- Ultra-wideband and other emerging communications technologies

Furthermore, the recognition that research within traditional boundaries in information and decision sciences is not adequate to address many of the emerging societal challenges has motivated LIDS research to branch out to areas at the intersection of several disciplines. As a result, LIDS is now engaged in several research thrusts that cut across disciplinary boundaries and involve considerable interaction and collaboration with colleagues in other MIT units and in other disciplines, such as:

- Foundations of network science, including network dynamics, control, and efficient algorithms
- Foundational research in game theory and mechanism design involving the study of new equilibrium notions and dynamics in games, and the design of efficiently computable incentive methods for large-scale, networked, dynamic environments
- New frameworks for modelling and understanding systemic risk
- Fundamental issues in cyber-physical systems, including architectural design, security and privacy, cross-layer algorithms, and tools for analysis, verification, and performance guarantees
- Development of scalable and efficient inference algorithms for problems involving big data, including basic research on graphical models—a general and rich framework for high dimensional inference
- Development of causal inference methods for gene regulation and early disease diagnostics

Research Highlights

Detailed information on the research carried out by each LIDS PI is provided in a subsequent section.

Some recent highlights include the following:

Several LIDS members have responded to the COVID-19 pandemic through engagement in relevant research, either individually, or through IDSS's Isolat initiative.

Professor Dimitri Bertsekas published Rollout, Policy Iteration, and Distributed Reinforcement Learning with Athena Scientific.

Professor Guy Bresler and his collaborators have made major new advances in their theory on the (average case) computational difficulty of high-dimensional statistical problems.

Professor Luca Carlone and his group have developed certifiable perception algorithms and have developed a first spatial perception engine that builds a dynamic scene graph from sensor data.

Professor Munther Dahleh and his group have created a new framework and mechanism for buying and selling training data for machine learning algorithms.

Marija Ilic led a Lincoln Laboratory team that completed a major study on sustainable and resilient electricity services in Puerto Rico.

Professor Ali Jadbabaie and his group developed a new framework for the study of misinformation spreading in social networks.

Professor Eytan Modiano and his group developed new mechanisms for optimizing information freshness in wireless networks.

Professor Alexandre Rakhlin and his group made significant advances in our understanding of the performance of overparameterized statistical learning algorithms.

Professor Devavrat Shah and his group developed new tensor estimation methods for enabling causal inference through machine learning.

Professor Suvrit Sra and his group made significant advances in our understanding of stochastic gradient descent algorithms used in neural network training.

Professor Caroline Uhler applied causal inference methods developed in her research to tackle the problem of causal structure discovery in gene regulation.

Kalyan Veeramachaneni and his group introduced a new framework, named MLBazaar, for developing automated machine learning software systems.

Professor Moe Win and his group have worked in the area of quantum information science and developed remote entanglement distribution protocols.

Professor Cathy Wu and her students demonstrated the effectiveness of transfer learning across different traffic control problems.

Faculty Activities

The activities listed below are organized in terms of individual faculty. Nevertheless, many of the major research activities not only cut across the disciplines, applications, and emerging areas mentioned previously, but are also collaborative with others within LIDS and elsewhere at MIT.

Saurabh Amin

The Resilient Infrastructure Networks group led by Associate Professor Saurabh Amin is working in three main areas: resilient network control; information systems in strategic environments; and, optimal resource allocation. He seeks to develop practically relevant methods and tools for improving the performance of large-scale critical infrastructure systems in the face of disruptions—both stochastic and adversarial. His work is grounded in the domains of highway transportation, electric power distribution, and urban water and gas networks. The methods and tools ae drawn from stochastic control, applied game theory, and optimization in networks.

The research on resilient network control aims to design operational strategies that maintain system performance under disturbances caused by random incidents or adversarial attacks. We ask questions about problems in transportation system operations, such as effects of capacity perturbations on stability of traffic queues and system throughput, operation of autonomous vehicle platoons in mixed traffic, and incident-aware ramp metering. The work on information systems in strategic environments examines the effects of heterogeneities in information access and accuracy on the equilibrium behavior of user populations, and consequently, the impacts to network-to-network efficiency. Recent work in Bayesian routing games systematically analyzes the relative value of heterogeneous traffic information systems for traveler populations who have incomplete and asymmetric information about the network state and other travelers. In optimal resource allocation, we design algorithms for strategic positioning of sensing resources in large-scale pipeline networks (e.g., water and natural gas) to obtain detection guarantees against multiple strategic failure events. We also study the value of proactive defense investments for securing multiple infrastructure facilities with different levels of criticality as the technological costs of attack and defense vary. To improve the resilience of electricity distribution networks against correlated failures (particularly, remote attacks and storm-induced disruptions), we investigate the use of microgrid operations and allocation/dispatch of low-inertia distributed energy resources.

Dimitri Bertsekas

Professor Bertsekas, McAfee Professor of Engineering, performs research on problems of sequential decision making under uncertainty, which are pervasive in communication networks, manufacturing systems, and logistics, and in the control of nonlinear dynamical systems. In theory, such problems can be addressed with dynamic programming techniques. In practice, only problems with a moderately sized state space can be handled. This research effort deals with the application of neural networks and other approximation and interpolation methodologies to overcome the curse of dimensionality of real-world stochastic control problems. Recent effort has focused on the new book *Rollout, Policy Iteration, and Distributed Reinforcement*, which aims to clarify the uses of distributed computation in reinforcement learning in the context of the rollout and policy iteration methods and incorporates Professor Bertsekas's latest research.

Robert C. Berwick

Professor Berwick's research during the past year continued to systematically evaluate the currently popular deep learning natural language processing systems by comparison with the capabilities of human systems. These systems do not perform as well as people, and we discovered that the claimed abilities of such networks to mirror human language performance have been greatly overestimated.

To remedy this, Professor Berwick's research group has computer implemented the first system, called MGSExplorer, that axiomatizes modern linguistic theory, using so-called satisfiability modulo theories (modern solvers for satisfiability problems like the Traveling Salesperson problem). By solving for syntax this constraint system can effectively find solutions or parses to sentences, given the axioms. Because the system is set up as a set of logical constraints, it can use optimized solvers for these constraints. Further, because such logical constraints are declarative and run in either direction, by pairing sentences with sketchy versions of "who did what to whom" (event representations that we may assume children have access to), then this system can actually acquire the necessary information to develop the dictionary for an English grammar. Such a system does readily learn how to parse "what to do" correctly. It can even go on to handle complex examples in Dutch that no other computational natural language system has been able to solve.

In a second thread that parallels this research, Professor Berwick's students systematically analyzed the formal properties of many deep net systems, demonstrating that they are not, in fact, physically realizable in many cases: if they do not have infinite precision arithmetic, then they actually amount to finite-state automata. While such automata plainly can have valuable engineering applications, they fail to approach the capacities that children attain.

Last year, Robert C. Berwick and Edward Stabler's edited collection *Minimalist Parsing*, was published by Oxford University Press.

Audun Botterud

Audun Botterud has been a principal research scientist in LIDS since September 2016. He is also a research affiliate in MIT's Center for Energy and Environmental Policy Research and he holds a co-appointment at Argonne National Laboratory. The main goal of Botterud's research is to improve the understanding of the complex interactions between engineering, economics, and policy in electricity markets, and ultimately enable the transition toward a cost efficient and reliable low-carbon energy system. Toward this end, he uses analytical methods from operations research and decision sciences combined with fundamental principles of electrical power engineering and energy economics.

Electricity markets with renewable energy

In the past year we have continued to investigate electricity market design questions that arise from more renewables on the grid. In particular, we have analyzed the interaction between the capacity values of energy storage and variable renewable energy, such as wind and solar. We have also proposed ways to speed up the solution of large-scale capacity expansion problems, and analyzed how uncertainty impacts policy incentives in electricity markets. Furthermore, we have investigated electricity market design with increasing shares of renewable energy and energy storage. We are also working on data-driven models for power system analytics, deriving theoretical bounds on prediction errors and speeding up cascading failure searches with graph convolutional networks.

Decarbonization of energy systems

We have completed a project on the potential transition to a low-carbon energy system in Italy through electrification. In an extension of the Electrify Italy project, we have investigated the importance of considering multiple years of renewable energy resource data in capacity expansion planning. We are also completing a study on the future role of renewable energy and natural gas in India. As an important measure toward energy system decarbonization, we have also published on optimization across different energy infrastructure systems (electricity, district heating, natural gas, hydrogen) through development of novel formulations and decomposition schemes for computational tractability of complex scheduling and planning problems. Finally, we are currently looking at the importance of considering reactive power and system inertia constraints in planning models for decarbonized power systems with high shares of renewable energy.

Energy storage analytics

In the last years, we have worked on factoring in empirical laboratory test results in model development and parameter tuning for power grid optimization models. We are also looking at the role of energy storage in distributed and isolated power systems, investigating how different battery technologies may contribute to meet cost, emissions, and reliability targets, considering the impact of battery degradation. We are also analyzing the optimal sizing and location of batteries for off-shore wind developments. Finally, we are developing improved analytical and computational algorithms to solve the multiperiod stochastic battery scheduling problem, demonstrating impressive computation performance in two new conference papers.

Guy Bresler

Associate Professor Bresler works in information theory, statistics, and applied probability. Specifically, his research aims to understand the relationship between combinatorial structure and computational tractability of high-dimensional inference in graphical and other statistical models.

In joint work with student Matthew Brennan, Professor Bresler has continued their line of work investigating the interplay between computational and statistical limits in statistical inference problems. Prior work characterizing computational limits by relating inference problems typically start with the planted clique problem, where a fully connected subgraph (clique) is hidden (planted) within a random graph, and map an instance to the problem of interest. While this approach has been modestly successful, virtually all reductions maintained the structure of the hidden signal, mapping to problems having the form sparse sub-matrix plus noise which is the same as that of planted clique itself. The insight in this work is that a slight generalization of the planted clique conjecture—secret leakage planted clique (PC-SL), wherein a small amount of information about the hidden clique is revealed—gives rise to a variety of new average-case reduction techniques, yielding a web of reductions relating statistical problems with very different structure. Based on generalizations of the planted clique conjecture to specific forms of PC-SL, the paper deduces tight statistical-computational tradeoffs for a diverse range of problems, including robust sparse mean estimation, mixtures of sparse linear regressions, robust sparse linear regression, tensor Principal Component Analysis (PCA), variants of dense k-block stochastic block models, negatively correlated sparse PCA, semirandom planted dense subgraph, detection in hidden partition models and a universality principle for learning sparse mixtures. This gives the first reduction-based evidence supporting a number of statistical-computational gaps observed in the literature. The paper introduces a slew of new techniques for average-case reductions that are expected to be useful also for relating other problems. The work raises a number of open problems and suggests that previous technical obstacles to average-case reductions may have arisen because planted clique is not the right starting point. An expanded set of hardness assumptions, such as PC-SL, may be a key first step toward a more complete theory of reductions among statistical problems.

In joint work with student Dheeraj Nagaraj, Professor Bresler studied the representation power of feedforward neural networks, writing two papers on the topic. The first paper develops a corrective mechanism for neural network approximation: the total available nonlinear units are divided into multiple groups and the first group approximates the function under consideration, the second approximates the error in approximation produced by the first group and corrects it, the third group approximates the error produced by the first and second groups together and so on. This technique yields several new representation and learning results for neural networks. The second paper proves sharp, dimension-free representation results for neural networks with D ReLU layers under square loss for a class of functions $G(D)$ defined in the paper.

Luca Carlone

Certifiable perception algorithms for high-integrity autonomous systems

Perception algorithms are increasingly used in safety-critical and high-integrity applications, including intelligent transportation, search-and-rescue, and space robotics, where algorithmic failures may have dire consequences and may put human lives at risk. My work has focused on the design of geometric perception algorithms that work in challenging, real-world conditions and provide formal performance guarantees. My major contribution has been to propose the notion—and establish the foundations—of certifiable perception algorithms, which includes algorithms for geometric perception that are able to formally assess the quality of their estimates and certify their correctness in the presence of noise and outliers. I have provided the first examples of certifiable algorithms for object localization in 3D point clouds and images.

Spatial AI

Autonomous operation in complex and dynamic environments requires robots to build and maintain a multifaceted model of the environment, including both a 3D geometric model (useful for navigation) and a semantic model (useful to characterize the presence/class of objects in the environment). My work has focused on the design of new representations of the environment and on real-time algorithms to infer these representations from sensor data. I proposed the notion of 3D dynamic scene graph, which extends typical maps used in robotics to capture geometry and semantics at multiple levels of abstraction. Dynamic scene graphs are directed graphs where nodes represent spatial concepts (e.g., objects, rooms, robots, humans) and edges represent pairwise spatio-temporal relations (e.g., a person is sitting on a chair). Their hierarchical and time-varying nature makes them an ideal representation to bridge low-level motion planning with high-level decision making in dynamic environments. However, building these representations from sensor data is challenging. My key contribution has been to provide the first spatial perception engine that builds a dynamic scene graph from sensor data.

Munther Dahleh

Professor Dahleh, director of the Institute for Data, Systems, and Society and the William A. Coolidge Professor of EECS, focused his research on advancing problems in machine learning for controlled dynamical systems with application neural learning and control.

"A Marketplace for Data: An Algorithmic Solution"

In this work, we aim to design a data marketplace; a robust real-time matching mechanism to efficiently buy and sell training data for machine learning tasks. As our main contributions we propose a mathematical model for a two-sided data market and formally define the key associated challenges; and, construct algorithms for such a market to function and analyze how they meet the challenges defined. We highlight two technical contributions, a new notion of fairness required for cooperative games with freely replicable goods; and, a truthful, zero regret mechanism to auction a class of combinatorial goods based on utilizing Myerson's payment function and the multiplicative weights algorithm.

"On Enhancing Resilience to Cascading Failures via Post-Disturbance Tweaking of Line Reactances"

The primary goal of this paper is to develop an optimization framework for studying the efficacy of transmission line reactance tweaking as a mechanism for post-disturbance control in a transmission network. Our numerical case study suggests that post-disturbance reactance tweaking, even on only a small number of lines, can be effective in reducing the amount of load shed in some scenarios in the tested system.

A model approximation approach for learning noisy stochastic jump LTI systems

This work addresses two fundamental problems in the context of jump linear systems (JLS). The first problem is concerned with characterizing the minimal state space dimension solely from input–output pairs and without any knowledge of mode switches. The second problem is concerned with characterizing the number of discrete modes of the JLS. For the first problem, we develop a linear system theory-based approach and construct an appropriate Hankel-like matrix. The rank of this matrix gives us the state space dimension. For the second problem we show that minimal number of modes corresponds to the minimal rank of a positive semidefinite matrix obtained by linearly transforming an output tensor.

"Performance Limitations in Sensorimotor Control: Trade-Offs Between Neural Computation and Accuracy in Tracking Fast Movements"

We use feedback control principles to quantify performance limitations of the sensorimotor control system (SCS) to track fast periodic movements. We show that linear models of the SCS fail to predict known undesirable phenomena, including skipped cycles, overshoot and undershoot, produced when tracking signals in the fast regime, while nonlinear pulsatile control models can predict such undesirable phenomena; and, tools from nonlinear control theory allow us to characterize fundamental limitations in this fast regime. Using a validated and tractable nonlinear model of the SCS, we derive an analytical upper bound on frequencies that the SCS model can reliably track before producing such undesirable phenomena as a function of the neurons' biophysical constraints and muscle dynamics.

Jonathan How

Professor How leads research efforts focused on the control of multiple autonomous agents, with an emphasis on distributed decision making with uncertainty; path planning, activity, and task assignment; mission planning for unmanned aerial vehicles and unmanned ground vehicles; sensor network design; and robust, adaptive, and nonlinear control. Professor How is also the principal investigator of the Aerospace Control Laboratory (ACL). A key aspect of ACL is Real-time indoor Autonomous

Vehicle test ENvironment (RAVEN), a unique experimental facility that uses a motion capture system to enable rapid prototyping of aerobatic flight controllers for helicopters and aircraft, and robust coordination algorithms for multiple vehicles; and a ground projection system that enables real-time animation of the planning environment, beliefs, uncertainties, intentions of the vehicles, predicted behaviors (e.g., trajectories), and confidence intervals of the learning algorithms.

Marija Ilic

Marija Ilic (professor emerita, Carnegie Mellon University) is a senior research scientist at LIDS and a permanent senior staff member at Lincoln Laboratory. She is also affiliated with MIT's Institute for Data, Systems, and Society. She continues to proactively work toward building a collaboration among the MIT-Lincoln Laboratory Group 73, LIDS, IDSS, and the MIT Energy Initiative (MITEI). She has done this through organizing and participating in several meetings with current and potential industry sponsors and by presenting her work and vision at workshops concerning electric energy systems. She has played a key role in the vision and growth of MIT's Electric Power Systems Center (an MITEI low-carbon energy center). Ilic is leading projects in major research areas, such as modeling and control for changing electric energy systems, with emphasis on microgrids, resiliency, and electricity markets. Several of her graduate student advisees have completed their degree work in this period (two PhD, and three MEng). She will have one more PhD student complete her studies in July 2020. Of note, three of these graduate students (two PhD, one MEng) are women.

Described below are highlights of the work this year that revolve around the novel modeling for simulation and control of electric energy systems pursued under her leadership in the Electric Energy Systems Group.

Ilic, together with two of her PhD students, (Xia Miao and Rupamathi Jaddivada) had a US patent issued on May 19, 2020, titled, Plug-and-Play Reconfigurable Electric Power Microgrids, US Patent 10656609B2.

Ilic led an MIT Lincoln Lab team on a Department of Homeland Security/Research Division funded study of future architectures for sustainable and resilient electricity service in Puerto Rico. It was shown that the Dynamic Monitoring and Decision Systems framework, as the basis for such architectures, can significantly enable utilization of clean solar power during normal operations and support gradual degradation of service during extreme events, instead of experiencing major, wide-spread blackouts.

The two graduate students that have successfully completed their PhD degrees are Xia Miao and Ana Jevtic. Xia Miao's thesis introduced the notion of autonomous electrical energy systems, in particular as related to the use of power-electronically-switched nonlinear control as enablers of autonomy. Ana Jevtic's thesis was devoted to conceptualizing cybersecure methods for electric power systems.

Rupamathi Jaddivada's PhD thesis (which will be completed in July 2020) proposes a unified approach to modeling generalized reactive power dynamics in the changing electric power industry with many intermittent clean resources.

In fall 2019, Ilic offered a new course, 6.247/6.082 Principles for Modeling, Simulation, and Control for Electric Energy Systems. This course introduced new multilayered modeling in transformed state space, which makes it possible to pose and discuss technology-agnostic solutions to future sustainable and resilient energy systems across multiple disciplines.

Ali Jadbabaie

During AY2020, PI Ali Jadbabaie finished a multi-investigator project sponsored by Defense Advanced Research Project Agency (DARPA) on scalable, nonconvex optimization, as part of the DARPA Lagrange Program. This work led to a series of papers and oral spotlight presentations in top machine learning venues, including AI STATS 2019, NeurIPS 2019, ICLR 2019, and ICML 2020. In addition, the work led to a new Office of Naval Research (ONR) program, jointly with LIDS PI Alexandre Rakhlin.

Furthermore, together with LIDS PI Pablo Parrilo, Professor Jadbabaie co-organized a second workshop, Learning for Dynamics and Control, which was held virtually on Zoom and was attended by more than 560 people worldwide.

In another work, PI Jadbabaie together with LIDS research scientist Amir Ajorlou and Social and Engineering Systems PhD student Chin-Chia Hsu submitted a new paper on misinformation spreading in social networks to *Review of Economic Studies*. LIDS PIs Jadbabaie and Asuman Ozdaglar were recipients of an IBM grant for research at the interface of machine learning and optimization theory.

Furthermore, since April 2020, PI Jadbabaie and his students have been active members of the IDSS Isolat group on modelling, intervention, control, and policy evaluation for the COVID-19 pandemic. In particular, SES PhD student Arnab Sarker together with PIs Jadbabaie and Devavrat Shah have created a prediction model for spread of COVID-19. Furthermore, SES PhD student Paolo Bertolotti's work on pool-testing is being used by the NIH RadX team for developing recommendation of testing protocols for COVID-19.

Patrick Jaillet

Dugald C. Jackson Professor and Co-director of the Operations Research Center Professor Jaillet's research focuses on online optimization and learning; machine learning; and decision making under uncertainty. Examples include online and dynamic versions of assignment/matching problems, secretary problems, routing problems, and their generalizations. The research deals with provable results (algorithmic design and analysis) on how to solve such problems under uncertainty, with or without explicit stochastic modeling of the uncertainty. Methodological tools include those from online optimization (competitive analysis), stochastic optimization (robust analysis), online learning (min/max regret analysis, Bayesian updates), reinforcement learning, and game theoretic concepts (price of anarchy) and their integrations.

Motivating applications arise from modern transportation sharing systems, dynamic resource allocation problems in various applications arising from the digital economy (search engines and online auctions), health care (kidney exchange programs), and social interactions (job search, house exchanges).

His research group at MIT this past academic year has included six ORC PhD students (Phillip Chodrow, Samuel Gilmour, Victor Gonzales, Jason Liang, Julia Romanski, and Sohil Shah), and two ORC SM students (Moise Blanchard and Matthew Yuan). His research group in Singapore has included two Singapore-MIT Alliance for Research and Technology (SMART) postdocs (Chungling Luo and Tien Mai), three National University of Singapore (NUS) PhD students (Zhongxiang Dai, Chi Lam, and Haibin Yu), one Nanyang Technological University PhD student (Anatoliy Prokhorchuk) one Singapore Management University PhD student (Meghna Lowalekar), and one Singapore University of Technology and Design PhD student (Gary Goh).

Funded research programs over this past academic year came from Office of Naval Research (Online Optimization and Learning in a Complex Environment); Air Force Office of Scientific Research (Building Attack Resilience into Complex Networks: Deterrence, Inspection, and Recovery); and from SMART (Future Mobility).

Sertac Karaman

Associate Professor Sertac Karaman carries out research in the areas of control theory, optimization, formal methods, stochastic processes, and applied probability with applications to robotics, mobile sensor networks, cyber-physical systems, and dynamic data-driven application systems.

Sanjoy Mitter

Professor Sanjoy Mitter has continued his research on the role of free energy estimation in Bayesian estimation. He has shown how this leads to an information-theoretic interpretation of nonlinear filtering relating Shannon information and Fisher information. This work is now being extended to distributed Kalman filtering.

This is joint work with Professor Moe Win and his graduate student Zhenyu Liu.

Eytan Modiano

Professor Eytan Modiano, associate director of LIDS, leads the Communications and Networking Research Group (CNRG), consisting of eight graduate students, and one postdoc. The primary goal of CNRG is the design of architectures for communication networks that are cost effective, scalable, and meet emerging needs for high data rate and reliable communications. In recent years the group has focused on robust network designs, wireless networks, data center networks, and interdependent cyber-physical networks.

Application domains, such as autonomous vehicles, command and control systems, virtual reality, and sensor networks, heavily rely upon the distribution of time-critical information. Over the past few years, CNRG has been developing network algorithms for optimizing information freshness in wireless networks. This past year, the group authored a number of papers on fundamental limits on information freshness and network control algorithms for optimizing freshness in wireless networks. Moreover, recently the group developed a wireless networking test bed consisting of fully programmable radios. This test bed is being used to experiment and validate network control algorithms for optimizing information freshness.

In recent years, the group started exploring the use of reinforcement learning (RL) in network control algorithms. In particular, the need for RL arises in networks with a mix of controllable and uncontrollable nodes, where the actions of the uncontrollable nodes can only be observed via feedback. RL can thus be used to make decisions at the controllable nodes, based on such feedback, in order to optimize overall network performance. To that end, they have been exploring various model-based reinforcement learning schemes that can be applied to network control and optimization.

CNRG's research crosses disciplinary boundaries by combining techniques from network optimization, queueing theory, graph theory, network protocols and algorithms, hardware design, and physical layer communications.

Asuman Ozdaglar

Professor Asuman Ozdaglar is the head of EECS as well as the deputy dean of Academics for the Schwarzman College of Computing. Her research group focuses on modelling, analysis, and optimization of large-scale dynamic multiagent networked systems. The research draws on advances in game theory, optimization theory, dynamical systems, and stochastic network analysis. It focuses on both investigating substantive problems in these areas and developing new mathematical tools and algorithms for the analysis and optimization of the systems and for processing large-scale data.

Pablo A. Parrilo

Professor Pablo Parrilo's research group is focused on mathematical optimization, systems theory, and control, with emphasis on development and application of computational tools based on convex optimization and algorithmic algebra.

Yury Polyanskiy

Associate Professor Yury Polyanskiy conducts research in the areas of mathematics of information (information theory), coding theory, and theory of random processes. His current work focuses on nonasymptotic characterization of the performance limits of communication systems, information-theoretic methods in statistics and machine learning, fault-tolerant circuits and probabilistic methods in combinatorics.

"Sampling of the Wiener Process for Remote Estimation Over a Channel with Random Delay"

Professor Polyanskiy and colleagues consider a problem of sampling a Wiener process, with samples forwarded to a remote estimator over a channel that is modeled as a queue. The estimator reconstructs an estimate of the real-time signal value from causally received samples. They study the optimal online sampling strategy that minimizes the mean square estimation error subject to a sampling rate constraint. They prove that the optimal sampling strategy is a threshold policy, and find the optimal threshold. Their work reveals an interesting connection between the age of information and remote estimation error. Their comparisons show that the estimation error achieved by the optimal sampling policy can be much smaller than those of age-optimal sampling, zero-wait sampling, and periodic sampling.

"Improved Log-Sobolev Inequalities, Hypercontractivity and Uncertainty Principle on the Hypercube"

Professor Polyanskiy and colleagues develop a new class of log-Sobolev inequalities (LSIs) that provide a non-linear comparison between the entropy and the Dirichlet form. For the hypercube, these LSIs imply a new version of the hypercontractivity for functions of small support. As a consequence, they derive a sharp form of the uncertainty principle for the hypercube: a function whose energy is concentrated on a set of small size, and whose Fourier energy is concentrated on a small Hamming ball must be zero.

"Broadcasting on Random Directed Acyclic Graphs"

In this paper, we treat the case of randomly constructed Directed Acyclic Graphs (DAGs), for which we show that broadcasting is only possible if the Biological Safety Cabinets noise level δ is below a certain (degree and function dependent) critical threshold. In particular, we show that the first r layers of such DAGs can be generated in either deterministic quasipolynomial time or randomized polylogarithmic time in r. These results portray a doubly-exponential advantage for storing a bit in bounded degree DAGs compared to trees, where d = 1 but layer sizes need to grow exponentially with depth in order for broadcasting to be possible.

Alexander Rakhlin

Professor Alexander Rakhlin has continued the investigation of statistical properties of learning procedures that interpolate (memorize) the training data. This study is motivated by the good out-of-sample performance of over-parametrized neural networks that achieve zero training error. The joint work with Tengyuan Liang and Xiyu Zhai analyzed the behavior of the risk for kernel ridgeless regression for various scaling regimes of sample size and dimension. The work pointed to a new, surprising multidescent phenomenon, where the out-of-sample performance has alternating regions of increasing and decreasing risk as sample size increases.

The work with Gil Kur and Yuval Dagan resolved (in the affirmative) long-standing open problems about optimality of least squares in convex regression and optimality of maximum likelihood estimation for log-concave distributions in dimensions above three. In contrast, our follow-up work with Gil Kur and Aditya Guntuboyina established that least squares are not optimal for the problem of recovering of convex sets from noisy support function measurements in dimensions above five.

The work with Adam Block and co-authors has focused on convergence of a generative modeling method that first estimates the score function of the distribution using denoising auto-encoders or denoising score matching and then employs Langevin diffusion for sampling. Given the recent practical success of such generative methods, we investigated the mixing rates of the Langevin diffusion under the manifold hypothesis and proposed new multiscale methods.

Another line of work focused on sequential decision making and reinforcement learning. In particular, in the work with Dylan Foster, we established the first optimal and efficient reduction from online regression to contextual bandits. When coupled with an optimal online regression method, the algorithm balances exploration and exploitation through a certain inverse gap weighting distribution. Current focus is on applying similar techniques to more general reinforcement learning settings.

Mardavij Roozbehani

Principal Research Scientist Mardavij Roozbehani led several research efforts on mathematical modeling, optimization, and control for cyber-physical networked systems. The application areas of focus include energy networks, transportation networks, financial networks, and digital or precision farming. The main theme in Roozbehani's research project are concentrated around models and methodologies for quantifying robustness, efficiency, and risk in energy networks, and understanding the trade-offs that the system architecture induces between these metrics.

Online learning and digital farming

The long-term goal of this new line of research is to empower farmers, particularly in rural areas in Africa and developing countries to improve their production through investments in technologies. The idea is to establish mechanisms to leverage the potential increase (through technology) in productivity and profits to provide an insurance for shocks. Our approach is based on a data sharing platform that incentivizes investors to participate and rewards farmers for it.

In parallel efforts we are exploring application of online learning, in particular reinforcement learning, to optimize farming actions such as irrigation and soil fertilization based on the state of soil and crops.

Real-time demand response optimization in electricity markets

In this line of research, we studied demand response procurement for residential electricity markets. The combination of (weakly) positive incentives and private information makes it challenging to determine the optimal demand response incentive. Typical practice is to use an estimate of the customer's counterfactual consumption as a baseline and to pay customers based on the difference between their final consumption and that baseline. Our research showed that the baseline should not necessarily be an estimate of customers' counterfactual consumption. We utilized tools from online learning to explore customer-specific cost functions and iteratively choose better demand response baselines.

Capacity-constrained real-time scheduling algorithms

In this research we studied the problem of scheduling of sporadic preemptible tasks. We developed a computationally efficient and compact characterization of all the scheduling decisions at a given time that maintain feasibility of the overall scheduling problem. This result allows us to optimize the scheduling decision based on exogenous optimality criteria that become available only in real time. Potential applications include electric vehicle charging, cloud computing, and web advertising.

Devavrat Shah

Professor Devavrat Shah and his research group are currently involved in developing theoretical foundations and algorithmic solutions for questions arising in the context of social data processing and decision making. Social data is the data generated by members of societies, through use of all sorts of modern services, which includes e-commerce portals, media, polls, utilities, and so on. Social data provides us with a granular look into the inner workings of our societies. On the positive side, it can help in designing better policies, improve social living by enabling efficient labor markets, and help businesses operate efficiently. On the flip side, it can be used to manipulate people online, corrupt their opinions, spread conspiracy theories and misinformation, and potentially affect the outcomes of otherwise democratic elections. If indeed we believe election outcomes were manipulated, it is essential to precisely identify the causes and mechanisms so that such actions can be prevented, and appropriate punitive actions can be taken.

Intellectually, this requires developing robust statistical models that capture the universal aspects of social behavior, ability to do causal inference with extremely limited information in the presence of a large number of potential causes, and algorithms that can scale with the amount of data while extracting meaningful information in the high-dimensional setting—it is like estimating biases of millions of heterogeneous coins simultaneously when we have access to at most one or few tosses per coin!

In summary, addressing the above-mentioned challenges requires fundamental progress at the interface of statistics, machine learning, computation, and social sciences. Professor Shah's group has been working to address these challenges by making progress on canonical questions in social data processing over the past decade or so, including learning choice, recommendations, ranking, crowdsourcing, causal inference, and reinforcement learning.

The ongoing research activities in Professor Shah's group involve developing methods for performing causal inferences using observational as well as experimental studies. Specifically, they have brought a novel lens from recent developments in machine learning to causal inference through tensor estimation to develop such methods. As a consequence, they provide what if analysis tools for policy evaluation, including the COVID-19 pandemic.

In addition to developing the methods, Professor Shah's group is developing software systems with extremely easy to use interfaces to provide broader access. These activities are synergetic to other ongoing activities in terms of data efficient methods for reinforcement learning, time series analysis, and socially responsible machine learning.

Suvrit Sra

Suvrit Sra joined EECS as an assistant professor in January 2018, and was promoted to associate professor in 2019; he continues as a LIDS PI, where he has been since January 16, 2015. His research interests lie in machine learning and optimization, and pure and applied mathematics.

Machine learning and optimization

Professor Sra's primary research is in optimization for machine learning. Over AY2020, he continued his work on developing theoretical foundations in nonconvex optimization, especially where geometry plays a role as well as for the central optimization problems in the field of deep neural networks. One main result from his group during this time was on establishing the memorization power of neural networks, which shows that already small-to-medium size neural networks are sufficiently powerful to perfectly memorize input data. Another fundamental theoretical result from his group was to discover and develop a theory of accelerated gradient methods for Riemannian manifolds, bringing to fruition a multiyear-long search for such algorithms, and thereby extending famous accelerated gradient results of Yurii Nesterov to the richer and continually expanding area of Riemannian optimization. Professor Sra and collaborators have also worked on developing foundational theory for machine learning in a weakly-supervised setting, a setting that occurs widely in applications, and is a part of his project National Science Foundation (NSF) TRIPODS + X.

During this period Professor Sra also continued to serve as area chair for several leading conferences in machine learning (including NeurIPS, ICML, and others). He also received an NSF CAREER Award during this period for supporting his theoretical work.

John Tsitsiklis

Professor John Tsitsiklis (Clarence J. LeBel Professor of Electrical Engineering, LIDS director, and IDSS associate director) has been primarily focused on the administrative aspects of LIDS, as well as on an online class offered through the IDSS MicroMasters in statistics and data science.

On the research side, he is interested in system modeling, analysis, optimization, and control in possibly stochastic and dynamic environments. Recent activities include the following:

Inference in complex dynamical networks

Joint work with two IDSS postdocs (Santiago Segarra and Michael Schaub) has addressed the problem of learning a statistical model of a network of interacting agents who are connected according to a (hidden and unknown) community structure (stochastic block model). Inference is to take place without knowledge of the network's edges, but based solely on nodal observations of a certain process. More concretely, we focused on observations that consist of snapshots of a diffusive process that evolves over the unknown network. Using tools from random matrix theory, we developed spectral algorithms that provably solve the problem of inferring both a partition of the nodes into blocks, as well as the parameters of the underlying stochastic block model.

Stochastic scheduling in the presence in heavy-tailed streams

From prior work in an earlier LIDS/MIT doctoral thesis, it is known that if a stochastic network is operated according to a popular scheduling policy (max-weight), and if one of the traffic streams is heavy-tailed, there may result excessive delays in other (light-tailed) streams. However, exact conditions for such a phenomenon to take place had been left as an open problem. In recent work with Arsalan Sharifnassab, a former visiting student, we have now developed tight necessary and sufficient conditions for such propagation of large delays, from one stream to another.

Caroline Uhler

Henry L. and Grace Doherty Associate Professor Caroline Uhler carries out research in the areas of machine learning and statistics with applications to genomics. In particular, her current research is centered around three interconnected pillars, namely causal inference, gene regulation, and early disease diagnostics.

"Causal Inference and Gene Regulation"

A central problem in biology and for biomedical discovery is the inference of gene regulatory networks—causal networks that allow predicting the effect of any intervention in the system. While large-scale interventional data is nowadays available in genomics, the problem of causal structure discovery in this area is challenging in practice for various reasons, including: (1) the presence of latent (unmeasured) confounders, (2) off-target intervention effects (knock-outs may also target other genes with similar sequences), (3) measurement error in the data collection process (single-cell RNA-seq data is highly zero-inflated), and (4) unknown disease subtypes and hence data coming from a mixture of causal models.

Professor Uhler and her collaborators have tackled these problems. In particular, they showed how to extend their permutation-based causal structure discovery framework to learn the intervention targets simultaneously while greedily optimizing the graph structure to solve (2) and how to extend their framework from permutations to posets to be able to incorporate bidirected edges and hence latent confounders to solve (1). In addition, they showed how higher-order moments can be used to learn causal structure among the latent causal variables based on the observed variables with measurement error to solve (3), and identified the causal structures that can be identified when given data from a mixture of causal graphs to solve (4). All algorithms are implemented in the group's causal DAG python library and have been validated using gene expression data and genomic interventions.

"Predicting Cell Lineages Using Autoencoders and Optimal Transport"

Many key biological processes, such as development and disease progression, require analyzing lineages of cells backward as well as forward in time. However, current single-cell experiments tend to be destructive to cells, so that a single lineage can only be measured at one point in time. Karren Yang, a PhD student in Uhler's group, developed a computational framework consisting of an autoencoder to first embed the single-cell images into a joint coordinate system, followed by optimal transport to compute the most plausible path for a population of cells to transition from one time point to the next, thereby predicting the lineage of cells from a single snapshot in time based on measurements of other cells at other time points. In collaboration with G. V. Shivashankar's lab, an experimental cell biology group at NUS and now at ETH Zurich, we validated our computational pipeline on imaging data taken during the activation of fibroblasts by tumor cells in engineered 3D tissues as well as in breast cancer cell lines and human tissue samples, showing that we could identify early disease biomarkers.

Kalyan Veeramachaneni

Kalyan Veeramachaneni joined LIDS in 2016 as a principal research scientist. His research group focuses on building algorithms that emulate human thought processes in order to effectively marshal the increasing amounts of data generated by the world today. The team develops automation technologies for data science, works to enable seamless interaction between people and analytical systems, and builds algorithms, systems, and open-source software now deployed for applications in the financial, medical, and education sectors.

MLBazaar: A bazaar of machine learning components for application development

Veeramachaneni's team's past work in the automation of machine learning model development (AutoML) makes them well-positioned to address this problem. While AutoML was originally meant to provide automated search for machine learning components, it evolved into a systematization of end-to-end model development, treating the process as a computational graph with abstract data processing blocks.

In 2019, Veeramachaneni and team introduced the Machine Learning (ML) Bazaar, a new framework for developing machine learning and automated machine learning software systems. They introduced ML primitives—which abstract away glue code, data flow, and data storage to represent ML steps as directed acyclic graphs—a unified API, and specifications for importing data processing and ML components from different software libraries. These pipelines are further paired with a hierarchy of AutoML strategies, including Bayesian optimization and bandit learning.

Multiple primitives, pipelines, and search strategies were then used to create a general-purpose, multitask, end-to-end AutoML bazaar that provides solutions for a variety of data modalities (e.g., image, text, graph, tabular, relational) and problem types (e.g., classification, regression, anomaly detection, graph matching).

Can neural networks compete with graphical models for generative modeling?

Tabular data is collected regularly in our society through both online and physical systems. Statisticians and the machine learning community have worked for decades to find effective modeling techniques for tabular data, but as data collection has proliferated, the types of variables and how they are distributed across possible values has diversified as well, challenging these techniques.

A new neural network technique called generative adversarial networks (GANs) had shown promise in modeling complex data and nonlinear correlations between variables. Veeramachaneni and PhD student Lei Xu, in collaboration with Professor Alfredo Cuesta-Infante from the University of Ray Juan Carlos, and PhD student Maria Skoularidou from the University of Cambridge, developed a conditional GAN (called CTGAN), which addressed the aforementioned complexities using several techniques. In their experiments, they outperformed the classical techniques in seven out of eight complex datasets. While others had attempted to use GANs to model tabular data, none had been able to outperform a technique from 1968.

Moe Win

The Wireless Information and Network Sciences Laboratory, led by Professor Moe Win, is involved in multidisciplinary research that encompasses developing fundamental theories, designing practical algorithms, and conducting network experiments for a broad range of real-world problems.

To advocate for outreach and diversity, the group is committed to attracting—and has a strong track record of actively recruiting—graduate and undergraduate students from underrepresented minority groups and giving them exposure to theoretical and experimental research at all levels.

Current research topics being investigated by Professor Win and his group include network localization and navigation, multi-object tracking, network interference exploitation, intrinsic wireless secrecy, adaptive diversity techniques, ultrawide bandwidth systems, and quantum information science. Details of a few specific projects are given below.

"Network Localization and Navigation"

The group has made notable contributions in the field of network localization and navigation from three aspects: joint localization and synchronization; localization based on channel state information (CSI); and, multi-object tracking (MOT). In particular, the group developed network localization and synchronization, a new paradigm that jointly considers localization and synchronization to achieve performance gains in a completely asynchronous wireless network.

The group also developed a localization algorithm based on CSI measurements. CSI measurements are available from various commercial, off-the-shelf devices such as laptops and smartphones. Therefore, the developed algorithm enables ubiquitous localization for daily use without additional hardware. By adopting a proposed message-passing methodology on a network factor graph, the algorithm can be implemented efficiently and is applicable to a large network. The group is currently extending the framework to incorporate both spatial and temporal cooperation, and to build a real-world localization system based on the developed algorithm.

In addition, the group has continued the development of a paradigm for scalable MOT and network operation for MOT applications. In particular, the group proposed Bayesian inference algorithms for localizing and tracking extended objects in the presence of data association uncertainty. These algorithms are based on message-passing methodology on a network factor graph. In addition to tracking algorithms, the group has proposed network operation algorithms for sensor management. More specifically, the group devised a sensor selection algorithm for multi-sensor MOT. The group is also continuing its work on MOT by performing experiments using emerging millimeter-wave radar technology.

Quantum information science

The entanglement properties of quantum states serve as a valuable resource for many important quantum operations. In this context, the group has devoted its efforts to establishing entangled qubit pairs between agents that are far apart. The group has characterized the memory size used in recurrence quantum entanglement distillation algorithms and developed a scheduling policy to achieve optimal memory efficiency. Moreover, the group has designed remote entanglement distribution (RED) protocols that maximize the entanglement distribution rate. The new concept of enodes has been introduced to represent the entangled qubit pairs in a network. This concept enables the design of optimal RED protocols based on linear programming. The group has also investigated queuing delay in quantum networks.

Cathy Wu

Cathy Wu started in 2019 as an assistant professor in CEE, a member of IDSS, and as a PI in LIDS. As a Gilbert W. Winslow (1937) Career Development Assistant Professor working at the intersection of machine learning, control theory, and transportation, her goal is to develop algorithms that increase the capacity for society to make good decisions. In her first year, Professor Wu recruited three PhD students and bootstrapped a new lab through educational, research, and service activities.

Education

In accordance with MIT's mission to bridge computing with other disciplines, Professor Wu revamped a core undergraduate course, 1.041 Transportation Systems Modeling; she introduced hands-on computational modules and a new unit on reinforcement learning. She also developed a new graduate course, 6.246/6.247 Advanced Topics in Control, which seeks to bridge theoretical and practical perspectives on the topic. Due to the interdisciplinary nature of Professor Wu's work, her students' participation in these courses was a valuable part of jumpstarting their research training.

Research

This year, the group studied several approaches to address scalability challenges faced by reinforcement learning algorithms in transportation. In one example, in the context of employing RL for learning congestion management strategies with automated vehicles, her student Zhongxia "Zee" Yan demonstrated the surprising effectiveness of transfer learning across different traffic control problems, including those with different network topologies and different fractions of automated vehicles (publication in preparation). In order to bridge theory and practice, the group has begun a research collaboration with the National Center for High-Performance Computing (NCHC) in Taiwan and Taichung City, Taiwan. The goal is to develop scalable, deep RL methods to quantify and realize the public health benefits of coordinated traffic signal control, where about one-third of air pollution results from transportation.

Events and Communications

LIDS continues to organize its signature events: the broadly attended LIDS Seminar Series, and the LIDS Student Conference, a student-run conference that celebrated its 25th year in spring 2020, which provides an interactive forum for students to discuss their research, and features several distinguished plenary speakers each year.

LIDS faculty also remain involved in the organization of major workshops and conferences, such as this year's Learning for Dynamics and Control (L4DC) workshop, now in its second year and co-organized by Professors Ali Jadbabaie and Pablo Parrilo, which brought together experts in machine learning and AI together with researchers in control theory and robotics.

In addition, LIDS faculty have been key contributors to the organization of various events under the umbrella of IDSS and the Statistics and Data Science Center, which sits within IDSS. These include the annual Women in Data Science conference (jointly presented by Harvard and MIT, and co-organized by Professor Caroline Uhler); the annual Statistics and Data Science Day, a community-building event for those interested in statistics at MIT; and, the weekly LIDS and Stats Teas, which provide students and postdocs an opportunity to give brief research presentations to the community in an informal setting.

Of particular note this year was our flagship event celebrating the 80th anniversary of the lab, LIDS@80: A Celebration. Held in November 2019, we welcomed over 400 people from around the world, bringing together LIDS alumni and students, members of the MIT community, friends of the lab from academia, government, and industry, and many others working in the field. Together with an impressive set of speakers, the celebration was an inspiring reminder of the lab's tremendous history and impact, as well as the many ways LIDS is poised to continue leading the field. In addition to the outstanding technical program, which highlighted the rigor and interdisciplinary nature of LIDS-style research, the event held a strong focus on the LIDS community. Appreciation of the lab's welcoming spirit, intellectual freedom, and deeply impressive faculty and alumni were common threads throughout the celebration. In particular, we took the occasion to honor our colleagues, Professors Michael Athans, Dimitri Bertsekas, Dave Forney, Robert Gallager, Sanjoy Mitter, and Alan Willsky, which gave us the opportunity to reflect on the lab's remarkable roots while looking toward an exciting future. We consider the event a great success.

The lab also continues to produce its annual community-oriented magazine, *LIDS/ALL*, which consistently receives great praise. *LIDS/ALL* features articles on important events related to LIDS as well as profiles of LIDS community members including students, faculty, alumni, and staff.

Awards

Student Enric Boix was named to the 2020 cohort of Siebel Scholars in recognition of his academic achievements, leadership, and commitment to addressing crucial global challenges.

A Lehigh University team, led by LIDS visiting scholar Alberto Lamadrid (Lehigh University) and with LIDS co-PIs Audun Botterud (principal research scientist), Marija Ilic (senior research scientist), and Patrick Jaillet, was awarded $2.5 million by Advanced Research Projects Agency-Energy of the US Department of Energy.

Guy Bresler received an NSF CAREER Award.

Team CoSTAR, co-led by Luca Carlone and including LIDS student Yun Chang, won second place in the DARPA Subterranean Challenge's Tunnel Circuit (August 2019) and first place in the Challenge's Urban Circuit (February 2020). In addition, Professor Carlone, together with LIDS postdoc Vasileios Tzoumas, and students Heng Yang and Pasquale Antonante, won Best Paper in the Robotic Vision category at the 2020 International Conference on Robotics and Automation.

Professor Carlone received a 2020 Robotics: Science and Systems Early Career Award. He also received the 2020 Vickie Kerrebrock Faculty Award recognizing faculty who have made significant contributions to building a sense of community in AeroAstro.

Jon How with LIDS student Stewart Jamieson and their collaborator won Best Paper in the Service Robotics category at the 2020 International Conference on Robotics and Automation.

Marija Ilic received the 2020 IEEE Power and Energy Society Outstanding Power Engineering Educator Award "for contributions to mentorship and education on modeling and control in power engineering."

Patrick Jaillet and collaborators were finalists in the 2019 INFORMS Data Mining Competition. Professor Jaillet also received the 2020 Burgess (1952) and Elizabeth Jamieson Prize for Excellence in Teaching from the Department of Electrical Engineering and Computer Science.

Student Peter Li, who is jointly advised by Professors Sertac Karaman (LIDS) and Vivienne Sze (RLE and Microsystems Technology Laboratory), was a 2020 Grand Finals winner of the Association for Computing Machinery's (ACM) Student Research Competition.

Students Samuel Miller, Zhenyu Liu, Zehao Yu, and George Denove, working with project advisors Professor Moe Win, Christopher C. Yu (Draper Laboratory), and Professor Andrea Conti (University of Ferrara), won first prize in the 2019 IEEE Communications Society's Communication Technology Changing the World Student Competition.

Eytan Modiano was recognized by MIT's Office of Graduate Education with a Committed to Caring award, given to faculty for providing exceptional mentoring, advising, and support to graduate students. Professor Modiano also won the 2020 IEEE INFOCOM Achievement Award for his pioneering contributions to the analysis and design of cross-layer resource allocation algorithms for wireless, optical, and satellite networks.

Pablo Parrilo, together with LIDS alums Hamza Fawzi (University of Cambridge) and James Saunderson (Monash University), received the 2020 SIAM Activity Group on Optimization Best Paper Prize. Professor Parrilo also received the 2020 Burgess (1952) and Elizabeth Jamieson Prize for Excellence in Teaching from the Department of Electrical Engineering and Computer Science.

Yury Polyanskiy received the IEEE Information Theory Society's 2020 James L. Massey Award.

Devavrat Shah, along with LIDS alum Tauhid Zaman (Yale University), received the prestigious ACM Sigmetrics Test of Time Paper Award for 2020. Notably, he was the recipient of the same award in 2019.

Student Chandler Squires received the 2020 David Adler Memorial Thesis Award for his MEng thesis. He is supervised by Professor Caroline Uhler.

Postdoc César Uribe and collaborator won Yahoo Research's 2019 Faculty and Research Engagement Program Award for their proposal, "Acceleration for Data Science and Machine Learning."

Moe Win, together with students Wenhan Dai and Tianyi Peng, received the IEEE ICNC Best Paper Award at the International Conference on Computing, Networking, and Communication, 2020. Professor Win and colleagues also received the IEEE Communications Society Best Tutorial Paper Award.

Other Honors

Professor G. David Forney was named an Eminent Member of IEEE by IEEE-HKN in their 2019 selection process.

Marija Ilic was elected a member of the Mathematical Institute of the Serbian Academy of Sciences and Arts. Ilic was also elected an ordinary member of Academia Europaea.

Ali Jadbabaie was named head of CEE.

Caroline Uhler was promoted to associate professor with tenure.

Moe Win was elected a fellow of the European Association for Signal Processing for "contributions to the theory and practice of cooperative localization."

Cathy Wu was named the Gilbert W. Winslow Career Development Assistant Professor by CEE.

Organizational Aspects

John Tsitsiklis continues to serve as LIDS director (having started in this role in April 2017). Eytan Modiano joined, effective July 1, 2017, as associate director. Professor Tsitsiklis is stepping down at the end of calendar year 2020. Sertac Karaman has been appointed as the new director; he will take office in summer 2021, with provisional leadership arrangements to be made for the interim period.

Key Statistics for AY2020

- Faculty PIs: 26
- Research staff PIs: 4
- Affiliate members: 9

- Administration, technical, and support staff: 11
- Postdocs and other research staff (non-PI): 30
- Visitors and other affiliates: 34
- Graduate students: 129
- Visiting students: 4

Overall Outlook

LIDS is a world-leading center for fundamental research in the information and decision sciences. It occupies a unique niche at the interface of theory and applications in diverse areas, and provides a central component underlying many of the recent technological advances and challenges, including in the currently vibrant fields of data science, statistics, machine learning, and intelligent systems. There are of course many activities in these domains taking place outside LIDS, including prominent applications. Within this broad range of activities, LIDS serves as a focal point in the development of the underlying fundamental methodologies, and as a meeting ground for like-minded researchers.

Besides fundamental research, LIDS is engaged in furthering collaborative efforts that balance theory and practice, for maximal impact. The umbrella provided by the College of Computing and by IDSS and the resulting opportunities for cross-cutting collaborations are very helpful in this respect.

Finally, while LIDS is a research-oriented entity, LIDS faculty maintain a leading role in curriculum innovation, thus bridging research and the classroom, in areas such as data science, control and autonomy, and networks.

John Tsitsiklis
Director
Clarence J. LeBel Professor, Department of Electrical Engineering and Computer Science

Eytan Modiano
Associate Director
Professor, Department of Aeronautics and Astronautics