Prashanth L.A.

Contact Information
Department of Computer Science and Engineering,
Indian Institute of Technology Madras,
Chennai 600036
email: prashla@cse.iitm.ac.in
www: http://www.cse.iitm.ac.in/~prashla
tel: +91-44-22574377

Research Interests

Theory: Reinforcement learning, simulation optimization, multi-armed bandits

Applications: Vehicular traffic contol, recommendation systems, service systems, wireless networks, portfolio optimization

Education

March 2013, Ph.D. in Computer Science and Automation, Indian Institute of Science (IISc)
Dissertation Topic: "Resource Allocation for Sequential Decision Making under Uncertainty: Studies in Vehicular Traffic Control, Service Systems, Sensor Networks and Mechanism Design"
Advisor: Prof. Shalabh Bhatnagar

August 2008, M.Sc. (Engg) in Computer Science and Automation, IISc
Dissertation Topic: "OFDM-MAC algorithms and their impact on TCP performance in next generation mobile networks"
Advisor: Prof. K. Gopinath

May 2002, B.E. in Computer Engineering, National Institute of Technology, Surathkal

Professional Experience

Aug 2024 - present, Visiting Associate Professor
At: Centre for Machine Intelligence and Data Sciences (C-MInDS), Indian Institute of Technology Bombay, Mumbai

Nov 2022 - Jul 2024, Associate Professor
At: Department of Computer Science and Engineering, Indian Institute of Technology Madras, Chennai

Mar 2017 - Oct 2022, Assistant Professor
At: Department of Computer Science and Engineering, Indian Institute of Technology Madras, Chennai

April 2015 - Feb 2017, Postdoctoral Researcher
At: Institute for Systems Research, University of Maryland, College Park MD

November 2014 - April 2015, Research Associate
At: Computer Science and Automation, Indian Institute of Science (IISc)

November 2012 - October 2014, Postdoctoral Researcher
At: SEQUEL project, INRIA Lille - Nord Europe

July 2002 - May 2009, Senior Software Systems Engineer

At: Texas Instruments (India) Pvt. Ltd (On leave of absence from Feb 2008 - May 2009)

May 2009 - May 2011, Project Associate

For: Dept. of Information Technology (India) project on wireless sensor networking for industrial automation

May 2011 - August 2011, Summer Researcher

At: IBM Research Labs, Bangalore, INDIA

Awards

IEEE ITSS Best Ph.D. Dissertation 2014 - Third Prize: awarded by IEEE Intelligent Transportation Systems Society (ITSS)

IBM PhD Fellowship, 2012

Publications

Books/Book Chapters

B1 **Prashanth.L.A.** and Michael Fu, Risk-Sensitive Reinforcement Learning via Policy Gradient Search, *Foundations and Trends in Machine Learning*, Vol. 15: No. 5, pp 537-693, 2022.

B2 S.Bhatnagar, H.L.Prasad and **Prashanth.L.A.**, Stochastic Recursive Algorithms for Optimization: Simultaneous Perturbation Methods, *Lecture Notes in Control and Information Sciences Series*, Vol. 434, Springer, ISBN 978-1-4471-4284-3, Edition: 2013, 302 pages.

B3 S. Bhatnagar, V. Borkar and **Prashanth.L.A.**, Adaptive Feature Pursuit: Online Adaptation of Features in Reinforcement Learning, *Reinforcement Learning and Approximate Dynamic Programming for Feedback Control*, by F. Lewis and D. Liu (eds.), IEEE Press Computational Intelligence Series, pp. 517-534, 2012, **Invited article**.

Journals

J1 Vishwajit Hegde, Arvind S. Menon, **Prashanth L.A.**, Krishna Jagannathan, Online Estimation and Optimization of Utility-Based Shortfall Risk, Mathematics of Operations Research (to appear), 2024.

J2 Akash Mondal, **Prashanth L.A.**, Shalabh Bhatnagar, Truncated Cauchy random perturbations for smoothed functional-based stochastic optimization, Automatica, vol. 162, 2024.

J3 **Prashanth L.A.** and Sanjay P. Bhat, A Wasserstein distance approach for concentration of empirical risk estimates, Journal of Machine Learning Research, vol. 23, no. 238, pp. 1-61, 2022.

J4 Nirav Bhavsar and **Prashanth L.A.**, Non-asymptotic bounds for stochastic optimization with biased noisy gradient oracles, IEEE Transactions on Automatic Control, vol. 68, issue 3, pp. 1628 - 1641, 2023.

J5 N. Vijayan and **Prashanth L.A.**, Smoothed functional-based gradient algorithms for off-policy reinforcement learning: A non-asymptotic viewpoint. Systems & Control Letters, vol. 155, 2021.

J6 **Prashanth L.A.**, Nathaniel Korda and Remi Munos, Concentration bounds for temporal difference learning with linear function approximation: The case of batch data and uniform sampling, Machine learning, vol. 110, issue 3, pp. 559-618, 2021.

J7 **Prashanth L.A.**, S. Bhatnagar, N. Bhavsar, Michael Fu and Steve Marcus, Random directions stochastic approximation with deterministic perturbations, IEEE Transactions on Automatic Control, vol. 65, no. 6, pp. 2450-2465, 2020.

J8 Ravi Kumar Kolla, **Prashanth L.A.**, Sanjay P. Bhat, Krishna Jagannathan, Concentration bounds for empirical conditional value-at-risk: The unbounded case, Operations Research Letters, vol. 47, issue 1, pp. 16-20, 2019.

J9 Jie Cheng, **Prashanth L.A.**, Michael Fu, Steve Marcus and Csaba Szepesvari, Stochastic optimization in a cumulative prospect theory framework, IEEE Transactions on Automatic Control, vol. 63, no. 9, pp. 2867-2882, 2018.

J10 **Prashanth L.A.**, S.Bhatnagar, Michael Fu and Steve Marcus, Adaptive system optimization using (simultaneous) random directions stochastic approximation, *IEEE Transactions on Automatic Control*, vol. 62, issue 5, pp. 2223-2238, 2017.

J11 **Prashanth L.A.** and Mohammad Ghavamzadeh, Variance-Constrained Actor-Critic Algorithms for Discounted and Average Reward MDPs, Machine Learning, vol. 105, no. 3, pp. 367-417, 2016.

J12 **Prashanth L.A.**,H.L.Prasad, S.Bhatnagar and P. Chandra, A constrained optimization perspective on actor critic algorithms and application to network routing, *Systems & Control Letters*, vol. 92, pp. 46-51, 2016.

J13 **Prashanth L.A.**, H.L.Prasad, N.Desai and S.Bhatnagar, Simultaneous Perturbation Methods for Adaptive Labor Staffing in Service Systems, *Simulation*, vol. 91, issue 5, pp. 432-455, 2015.

J14 S.Bhatnagar and **Prashanth L.A.**, Simultaneous Perturbation Newton Algorithms for Simulation Optimization, *Journal of Optimization Theory and Applications,* vol. 164, issue. 2, pp. 621-643, 2015.

J15 **Prashanth L.A.**, A.Chatterjee and S.Bhatnagar, Two Timescale Convergent Q-learning for Sleep–Scheduling in Wireless Sensor Networks, *Wireless Networks*, vol. 20, issue. 8, pp. 2589-2604, 2014.

J16 H.L.Prasad, **Prashanth L.A.**, S.Bhatnagar and N.Desai, Adaptive Smoothed Functional Algorithms for Optimal Staffing Levels in Service Systems, *Service Science (INFORMS)*, vol. 5, no. 1, pp. 29-55, 2013.

J17 **Prashanth L.A.** and S.Bhatnagar, Threshold Tuning using Stochastic Optimization for Graded Signal Control, *IEEE Transactions on Vehicular Technology*, vol. 61, no. 9, pp. 3865-3880, 2012.

J18 **Prashanth L.A.** and S.Bhatnagar, Reinforcement learning with function approximation for traffic signal control, *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 2, pp. 412-421, 2011.

Proceedings of International Conferences

C1 Sumedh Gupte, **Prashanth L.A.**, Sanjay Bhat, Optimization of utility-based shortfall risk: A non-asymptotic viewpoint, IEEE Conference on Decision and Control (CDC), 2024 (accepted).

C2 Shubhada Agrawal, **Prashanth L.A.**, Siva Theja Maguluri, Policy Evaluation for Variance in Risk-sensitive Average Reward Reinforcement Learning, *41st International Conference on Machine Learning,* PMLR 235:471-502, 2024.

C3 Gugan Thoppe, **Prashanth L.A.**, Sanjay Bhat, Risk Estimation in a Markov Cost Process: Lower and Upper Bounds *41st International Conference on Machine Learning,* PMLR 235:48124-48138, 2024.

C4 Mizhaan Prajit Maniyar, **Prashanth L.A.**, Akash Mondal, Shalabh Bhatnagar, A Cubic-regularized Policy Newton Algorithm for Reinforcement Learning, *27th International Conference on Artificial Intelligence and Statistics (AISTATS),* PMLR 238:4708-4716, 2024.

C5 N. Vijayan and **Prashanth L.A.**, A policy gradient approach for optimization of smooth risk measures, *Proceedings of the Thirty-Ninth Conference on Uncertainty in Artificial Intelligence (UAI)*, PMLR 216:2168-2178, 2023.

C6 S.Bhatnagar and **Prashanth L.A.**, Generalized Simultaneous Perturbation Stochastic Approximation with Reduced Estimator Bias, *57th Annual Conference on Information Sciences and Systems (CISS)*, pp. 1-6, doi: 10.1109/CISS56502.2023.10089720, 2023.

C7 Gandharv Patil, **Prashanth L.A.**, Dheeraj Nagaraj, Doina Precup, Finite time analysis of temporal difference learning with linear function approximation: Tail averaging and regularisation, *26th International Conference on Artificial Intelligence and Statistics (AISTATS)*, PMLR 206:5438-5448, 2023.

C8 Vincent Y. F. Tan, **Prashanth L.A.**, and Krishna Jagannathan, A Survey of Risk-Aware Multi-Armed Bandits, International Joint Conference on Artificial Intelligence (IJCAI) (Survey track), pp. 5623-5629, 2022.

C9 Ajay Pandey, **Prashanth L.A.**, and Sanjay Bhat, Estimation of Spectral Risk Measures, The Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI), pp. 12166-12173, 2021.

C10 **Prashanth L.A.**, Krishna Jagannathan and Ravi Kumar Kolla, Concentration bounds for CVaR estimation: The cases of light-tailed and heavy-tailed distributions, *37th International Conference on Machine Learning (ICML)*, PMLR 119:5577-5586, 2020.

C11 Sanjay P. Bhat and **Prashanth L.A.**, Concentration of risk measures: A Wasserstein distance approach, *33rd Conference on Neural Information Processing Systems*, pp. 11762-11771, 2019.

C12 V. P. Boda and **Prashanth L.A.**, Correlated bandits: or How to minimize mean-squared error online, *36th International Conference on Machine Learning (ICML)*, PMLR 97:686-694, 2019.

C13 Aditya Gopalan, **Prashanth L.A.**, Michael Fu and Steve Marcus, Weighted bandits or: How bandits learn distorted values that are not expected, *31st AAAI Conference on Artificial Intelligence (AAAI)*, pp. 1941-1947, 2017.

C14 D. Sai Koti Reddy, **Prashanth L.A.** and S. Bhatnagar, Improved Hessian estimation for adaptive random directions stochastic approximation, *IEEE Conference on Decision and Control (CDC)*, pp. 3682-3687, 2016.

C15 **Prashanth L.A.**, Cheng Jie, Michael Fu, Steve Marcus and Csaba Szepesvári, Cumulative Prospect Theory Meets Reinforcement Learning: Prediction and Control, *33rd International Conference on Machine Learning (ICML)*, pp. 1406-1415, 2016.

C16 Xiaowei Hu, **Prashanth L.A.**, András György and Csaba Szepesvári, (Bandit) Convex Optimization with Biased Noisy Gradient Oracles, *19th International Conference on Artificial Intelligence and Statistics (AISTATS)*, pp. 819-828, 2016.

C17 Nathaniel Korda and **Prashanth L.A.**, On TD(0) with function approximation: Concentration bounds and a centered variant with exponential convergence, *32nd International Conference on Machine Learning (ICML)*, pp. 626-634, 2015.

C18 H.L.Prasad, **Prashanth L.A.** and S.Bhatnagar, Two Timescale Algorithms for Learning Nash Equilibria in General-Sum Stochastic Games, *14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pp. 1371-1379, 2015.

C19 Nathaniel Korda, **Prashanth L.A.** and Remi Munos, Fast gradient descent for drifting least squares regression, with application to bandits, *29th AAAI Conference on Artificial Intelligence (AAAI)*, pp. 2708-1714, 2015.

C20 Raphael Fonteneau and **Prashanth L.A.**, Simultaneous Perturbation Algorithms for Batch Off-Policy Search, *53rd IEEE Conference on Decision and Control (CDC)*, pp. 2622-2627, 2014.

C21 **Prashanth L.A.**, Policy Gradients for CVaR-Constrained MDPs, *25th International Conference on Algorithmic Learning Theory (ALT)*, pp. 155-169, 2014.

C22 **Prashanth L.A.**, Nathaniel Korda and Remi Munos, Fast LSTD using stochastic approximation: Finite time analysis and application to traffic control, *7th European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD)*, pp. 66-81, 2014.

C23 **Prashanth L.A.**, A. Chatterjee and S.Bhatnagar, Adaptive sleep-wake control using reinforcement learning in sensor networks, *6th International IEEE Conference on Communication Systems and Networks (COMSNETS)*, pp. 1-8, 2014.

C24 **Prashanth L.A.** and Mohammad Ghavamzadeh, Actor-Critic Algorithms for Risk-Sensitive MDPs, *27th Annual Conference on Advances in Neural Information Processing Systems (NIPS)*, **Full oral presentation (%1.4 acceptance - 20 out of 1420 submissions)**, pp. 252-260, 2013.

C25 **Prashanth L.A.**, H.L.Prasad, N.Desai and S.Bhatnagar, Mechanisms for Hostile Agents with Capacity Constraints, *12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pp. 659–666, 2013.

C26 **Prashanth L.A.**, H.L.Prasad, N.Desai, S.Bhatnagar and G.Dasgupta, Stochastic optimization for adaptive labor staffing in service systems, *9th International Conference on Service Oriented Computing (ICSOC)*, pp. 487–494, 2011.

C27 **Prashanth L.A.** and S.Bhatnagar, Reinforcement Learning with Average Cost for Adaptive Control of Traffic Lights at Intersections, *14th International IEEE Conference on Intelligent Transportation Systems*, pp. 1640–1645, 2011, **Invited article**.

C28 **Prashanth.L.A.**, K. Gopinath, OFDM-MAC algorithms and their impact on TCP performance in next generation mobile networks, *3rd International IEEE Conference on COMmunication System softWAre and MiddlewaRE (COMSWARE)*, pp. 133-140, 2008.

C29 **Prashanth L.A.**, Sajal Kumar Das, K Gopinath, MAC design for heterogeneous application support in OFDM based wireless systems, *5th IEEE Consumer Communications and Networking Conference (CCNC)* (short paper) , pp. 412-413, 2008.

Workshops

W1 Xiaowei Hu, **Prashanth L.A.**, András György and Csaba Szepesvári, (Bandit) Convex Optimization with Biased Noisy Gradient Oracles, *8th NIPS Workshop on Optimization for Machine Learning*, 2015.

W2 **Prashanth L.A.** and Mohammad Ghavamzadeh, SPSA based Actor-Critic Algorithm for Risk Sensitive Control, *11th European Workshop on Reinforcement Learning (EWRL)*, 2013.

W3 **Prashanth L.A.** and S.Bhatnagar, Control of traffic lights at junctions using reinforcement learning, *Computer Aided Transportation Planning and Traffic Engineering*, 2009.

Preprints

P1 Ayon Ghosh, **Prashanth L.A.** and Krishna Jagannathan, Concentration Bounds for Optimized Certainty Equivalent Risk Estimation, arxiv preprint arXiv:2405.20933, Under review in a top-ML conference, 2024.

P2 Arpan Mukherjee, Meltem Tatli, **Prashanth L. A.**, Karthikeyan Shanmugam, Ali Tajer, Risk-sensitive Bandits: Arm Mixture Optimality and Regret-efficient Algorithms, Under review in a top-ML conference, 2024.

P3 Soumen Pachal, S.Bhatnagar and **Prashanth L.A.**, Generalized Simultaneous Perturbation Stochastic Approximation with Reduced Estimator Bias, arxiv preprint arXiv:2405.20933, Under review in IEEE Transactions on Automatic Control, 2024.

P4 N. Vijayan and **Prashanth L.A.**, Policy gradients for distortion risk measures, arxiv preprint arXiv:2107.04422, 2024.

P5 Ravi Kumar Kolla, Prashanth L.A., Aditya Gopalan, Krishna Jagannathan, Michael Fu, Steve Marcus, Bandit algorithms to emulate human decision making using probabilistic distortions, arxiv preprint arXiv:1611.10283, 2023.

P6 Xiaowei Hu, **Prashanth L.A.**, András György and Csaba Szepesvári, (Bandit) Convex Optimization with Biased Noisy Gradient Oracles, arxiv preprint arXiv:1609.07087, 2020.

P7 **Prashanth L.A.**, H.L.Prasad, and S.Bhatnagar, Actor-Critic Algorithms for Learning Nash Equilibria in N-player General-Sum Games, arxiv preprint arXiv:1401.2086, 2015.

Tutorials

1. *Risk-sensitive reinforcement learning via policy gradient search*, co-presented with Michael Fu at AAAI Conference on Artificial Intelligence (AAAI), 2023.

2. *Risk-Aware Multi-armed Bandits,* co-presented with Krishna Jagannathan at IEEE International Conference on Signal Processing and Communications (SPCOM), 2022.

3. *Reinforcement learning,* ACM India Summer School on theoretical and algorithmic aspects on Machine Learning, 2019.

4. *Simultaneous perturbation methods for simulation optimization*, co-presented with Shalabh Bhatnagar at Indian Control Conference, 2018.

5. *Simultaneous perturbation methods for stochastic non-convex optimization*, ACM MobiHoc, 2017.

Invited Talks

1. *Finite time analysis of temporal difference learning with linear function approximation: Tail averaging and regularisation,* Data science: Probabilistic and optimization methods, discussion meeting at International Centre for Theoretical Sciences, Bengaluru, Jul 2023.

2. *Finite time analysis of temporal difference learning with linear function approximation: Tail averaging and regularisation*, Networks Seminar Series at Indian Institute of Science, Feb 2023.

3. *A Wasserstein distance approach for concentration of empirical risk estimates*, Information Theory and Data Science Workshop at National University of Singapore, Jan 2023.

4. *Concentration of risk measures: A Wasserstein distance approach*, IITB Workshop on Stochastic Models, Feb 2022.

5. *Concentration bounds for temporal difference learning with linear function approximation: The case of batch data and uniform sampling,* IISc workshop on Deep Reinforcement Learning, Oct 2021.

6. *Concentration of risk measures: A Wasserstein distance approach,* University of Maryland College Park, Dec 2019.

7. *Cumulative prospect theory meets bandits and reinforcement learning,* Department of Computer Science and Automation, Indian Institute of Science, Bangalore, June 2017.

8. *Concentration bounds for TD(0) with function approximation*, Communication, Control and Signal Processing Seminar, University of Maryland - College Park, USA, 2015.

9. *Cumulative Prospect Theory Meets Reinforcement Learning: Estimation and Control*, AI seminar, University of Alberta, Edmonton, Canada, 2015.

10. *On the convergence rate of TD(0) with function approximation: Non-asymptotic bounds in online and batch settings*, Recent Advances in Reinforcement Learning Workshop, Indian Institute Of Technology, Madras, India, 2015.

11. *Stochastic approximation for speeding up LSTD/LSPI (and least squares regression/LinUCB)*, Department seminar, Computer Science and Automation, Indian Institute of Science, India, 2014.

12. *Fast gradient descent for drifting least squares regression with applications to news-recommendation systems*, Large scale Online Learning and Decision Making Workshop, Cumberland Lodge, Windsor, UK, 2014.

13. *Actor-critic algorithms for risk-sensitive MDPs*, French Meeting on Planning, Decision Making and Learning, Liege, Belgium, 2014.

14. *Online gradient descent for LS regression: Non-asymptotic bounds and application to bandits*, Large scale Online Learning and Decision Making Workshop, Cumberland Lodge, Windsor, UK, 2013.

Projects

DST-Early Career Research Award for a project entitled *Simulation-based Optimization in a Cumulative Prospect Theory Framework*, 2018-2021.

Mentoring

Ph.D.
- Nithia V (CS17D003): Graduated.
- Sumedh Gupte (CS21D014): Joined Jul 2021; external candidate working with TCS Research; Co-guide: Sanjay Bhat; Status: comprehensive exam and research proposal seminar completed.
- Soumen Pachal (CS22D009): Joined Jul 2022; external candidate working with TCS Research; Co-guide: Avinash Achar; Status: comprehensive exam and research proposal seminar completed.

M.S.
- Nirav Bhavsar (CS17S016): Joined Jul 2017, Status: Graduated. Winner of the 'Biswajit Sain MS Thesis Award 2021'.
- Ajay Pandey (CS17S011): Joined Jul 2017, Status: Graduated.
- Dipayan Sen (CS18S012): Joined Jul 2018, Status: In progress.
- Udit Narayan (CS23S038): Joined Jul 2023. Status: In progress.

Professional Service

Senior program committee member: AAAI 2022, AAAI 2021.

Conference reviewer: CDC 2024, AISTATS 2024, AISTATS 2023, ACC 2022, ICML 2022, ICML 2021, AISTATS 2021, NeurIPS 2020, AAAI 2020, ICML 2020, NIPS 2019, CDC 2019, ICML 2019, NIPS 2018, ECML 2018, ICML 2018, COLT 2018, IJCAI 2018, AAAI 2018, NIPS 2017, ECML 2017, NIPS 2016, COLT 2016, CDC 2016, AAAI 2016, NIPS 2015, ICML 2015, IJCAI 2015, WSC 2015, ICML 2014.

Reviewer for journals: IEEE Transactions on Automatic Control, Systems & Control Letters, Mathematics of Operations Research, Journal of Optimization Theory and Applications, Stochastic Systems, IEEE Transactions on Information Theory, Statistics and Probability Letters, IEEE Transactions on Vehicular Technology, IEEE Transactions on Intelligent Transportation Systems, Operations Research, IEEE Transactions on Parallel and Distributed Systems.

Teaching

At IITB:

Introduction to machine learning (DS 303): Jul-Nov 2024.

At IITM:

Linear algebra and random processes (CS6015): Jul-Nov 2017, Jul-Nov 2019.

Multi-armed bandits (CS6046): Jan-May, 2018, Jan-May 2019.

Reinforcement learning (CS6700): Jul-Nov, 2018, Jan-May 2021, Jul-Nov, 2021.

Pattern recognition and machine learning (CS5691): Jan-May 2019, Jan-May 2020.

Stochastic optimization (CS6515): Jan-May 2023.

Stochastic modeling and the theory of queues (EE6150): Sep-Dec 2020.

Topics in reinforcement learning (CS7011): Jan-May 2022.

Object oriented algorithms, implementation and analysis lab (CS2810): Jan-May 2022.

CS2700 (Programming and data structures): Jul-Nov 2022.

CS3500 (Operating systems): Jul-Nov 2023.