

# Prashanth L.A.

## Contact Information

Institute For Systems Research,  
A.V. Williams Bldg (rm 3103),  
University of Maryland  
College Park, MD 20742  
email: prashla@umd.edu  
www: <http://www.isr.umd.edu/~prashla>  
tel: +13014051241

## Research Interests

**Theory:** Reinforcement Learning, Stochastic Optimization, Multi-armed Bandits

**Applications:** Road Traffic Control, Recommendation Systems, Service Systems, Wireless Networks

## 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

**April 2015 - present, 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 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.
- B2 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 **Prashanth L.A.**, S.Bhatnagar, Michael Fu and Steve Marcus, Adaptive system optimization using (simultaneous) random directions stochastic approximation, *IEEE Transactions on Automatic Control (To Appear)*, 2016.
- J2 **Prashanth L.A.** and Mohammad Ghavamzadeh, Variance-Constrained Actor-Critic Algorithms for Discounted and Average Reward MDPs, *Machine Learning Journal*, 2016.
- J3 **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*, pp. 1-8, 2016.
- J4 **Prashanth L.A.**, H.L.Prasad, N.Desai and S.Bhatnagar, Simultaneous Perturbation Methods for Adaptive Labor Staffing in Service Systems, *Simulation*, DOI: 10.1177/0037549715581198, pp. 1-24, 2015.
- J5 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.
- J6 **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.
- J7 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.
- J8 **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.
- J9 **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 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) (Accepted)*, 2017.
- C2 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) (Accepted)*, 2016.
- C3 **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)*, 2016.
- C4 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)*, 2016.
- C5 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.
- C6 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.
- C7 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.
- C8 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.
- C9 **Prashanth L.A.**, Policy Gradients for CVaR-Constrained MDPs, *25th International Conference on Algorithmic Learning Theory (ALT)*, pp. 155-169, 2014.
- C10 **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.
- C11 **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.
- C12 **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.
- C13 **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.
- C14 **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.
- C15 **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**.

- C16 **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.
- C17 **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) 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 Xiaowei Hu, **Prashanth L.A.**, András György and Csaba Szepesvári, (Bandit) Convex Optimization with Biased Noisy Gradient Oracles, *Under review in Journal of Machine Learning Research (JMLR)*, arxiv preprint arXiv:1609.07087, 2016.
- P2 **Prashanth L.A.**, Nathaniel Korda and Remi Munos, Stochastic approximation for speeding up LSTD (and LSPI), *Revised version under review in Machine learning Journal (MLJ)*, 2015.
- P3 **Prashanth L.A.**, H.L.Prasad, and S.Bhatnagar, Actor-Critic Algorithms for Learning Nash Equilibria in N-player General-Sum Games, *Revision under preparation*, arxiv preprint arXiv:1401.2086, 2015.

## Invited Talks

1. *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.
2. *Actor-critic algorithms for risk-sensitive MDPs*, French Meeting on Planning, Decision Making and Learning, Liege, Belgium, 2014.
3. *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.
4. *Stochastic approximation for speeding up LSTD/LSPI (and least squares regression/LinUCB)*, Department seminar, Computer Science and Automation, Indian Institute of Science, India, 2014.
5. *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.
6. *Cumulative Prospect Theory Meets Reinforcement Learning: Estimation and Control*, AI seminar, University of Alberta, Edmonton, Canada, 2015.
7. *Concentration bounds for TD(0) with function approximation*, Communication, Control and Signal Processing Seminar, University of Maryland - College Park, USA, 2015.

## Professional Service

**Conference reviewer:** NIPS 2016, COLT 2016, CDC 2016, AAAI 2016, NIPS 2015, ICML 2015, IJCAI 2015, WSC 2015, ICML 2014.

**Reviewer for journals:** Systems & Control Letters, IEEE Transactions on Vehicular Technology, IEEE Transactions on Intelligent Transportation Systems, Operations Research, IEEE Transactions on Parallel and Distributed Systems.

## Teaching Assistance

### Dept. of Computer Science and Automation, Indian Institute of Science

Shared responsibility for lectures, exams, homework assignments, and grading for the following courses:

- Topics in Stochastic Approximation Algorithms, Aug-Dec 2011.
- Computer Communication Networks, Jan-Apr 2011, Jan-Apr 2012.
- Topics in Stochastic Control and Reinforcement Learning, Aug-Dec 2010.
- Linear Algebra, Aug-Dec 2010.

## Courses taken

### Ph.D. CGPA: 6.8/8.0

- Analysis-I
- Analysis-II (Measure Theory)
- Probability Theory
- Linear Algebra
- Convex Optimization
- Topics in Stochastic Approximation Algorithms
- Computer Communication Networks

### M.Sc. CGPA: 7.0/8.0

- Discrete Structures
- Formal Methods in Computer Science
- Modeling and Simulation
- Operating Systems

## Computer Skills

- Languages: C, C++, Java, Perl, Python
- Applications: Traffic simulation software, simulation of service systems and wireless networks
- Algorithms: Experience programming stochastic optimization, bandit and reinforcement learning algorithms
- Operating Systems: Unix/Linux, Windows

## Interests and Activities

- Lawn tennis, chess and literature
- Enabling education to the underprivileged: member and former trustee of IndiaSudar ([www.indiasudar.org](http://www.indiasudar.org))

# References

Shalabh Bhatnagar  
Professor  
Computer Science and Automation Dept.  
Indian Institute of Science  
Bangalore-560012, INDIA  
+91 80 2293-2987  
shalabh@csa.iisc.ernet.in

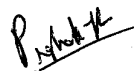
Rémi Munos  
Senior Researcher  
SequeL team  
INRIA Lille - Nord Europe  
59650 Villeneuve dAscq, France  
+33 (0)3 59 57 79 06  
remi.munos@inria.fr

Steve Marcus  
Professor, Electrical & Computer Engineering Dept.  
& Institute for Systems Research  
University of Maryland  
College Park, MD 20742  
+1 301 405 4252  
marcus@umd.edu

Michael Fu  
Professor, Robert H. Smith School of Business  
& Institute for Systems Research  
University of Maryland  
College Park, MD 20742  
+1 301 405 2241  
mfu@rhsmith.umd.edu

Vivek Shripad Borkar  
Institute Chair Professor  
Department of Electrical Engineering  
Indian Institute of Technology Bombay  
Mumbai 400076, India  
+91 22 2576-9405  
borkar@ee.iitb.ac.in

I hereby certify that the information above is true and accurate.



(Prashanth.L.A.)