Subhojyoti Mukherjee

Wisconsin Institute of Discovery
University of Wisconsin-Madison
Madison, WI 53715

Phone: +1 669 208 8939
Email: smukherjee27@wisc.edu, subhojyotimukherjee22@gmail.com
Website: https://subhojyoti.github.io/

Research Interests

Machine learning, Reinforcement learning, Online Learning, Multi-armed bandits, Applied Probability, Optimization.

Education

University of Wisconsin-Madison, Madison, USA — Fall 2019 – current
Ph.D., Electrical & Computer Engineering
Adviser: Dr. Robert Nowak
(Transferred out from UMass Amherst)

University of Massachusetts, Amherst, USA — 2018 – 2019
Ph.D., Computer Science

Indian Institute of Technology Madras, India — 2015–2018
M.S (Research), Computer Science
Advisers: Dr. Balaraman Ravindran and Dr. Nandan Sudarsanam

West Bengal University of Technology, Kolkata, India — 2009–2013
Bachelor of Technology, Computer Science & Engineering

Publications

1. **Subhojyoti Mukherjee**, and Odalric-Ambrym-Maillard, "*Distribution-dependent and Time-uniform Bounds for Piecewise i.i.d Bandits*", *Accepted in Thirty-sixth International Conference on Machine Learning* ***(ICML-19)***, Workshop on Reinforcement Learning for Real Life 2019 track [Poster]. [Paper]
2. **Subhojyoti Mukherjee**, K.P. Naveen, Nandan Sudarsanam, and Balaraman Ravindran, "*Efficient UCBV: An Almost Optimal Algorithm using Variance Estimates*", *Proceedings of the Thirty-Second Association for the Advancement of Artificial Intelligence* ***(AAAI-18)***, main conference track [Oral].[Paper]
3. **Subhojyoti Mukherjee**, K.P. Naveen, Nandan Sudarsanam, and Balaraman Ravindran, "*Thresholding Bandits with Augmented UCB*", *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence* ***(IJCAI-17)***, main conference track [Oral + Poster]. [Paper]

Research Internships

1. **CMU, ECE Dept., Pittsburgh, USA:** From 10th June, 2019 to 16th August, 2019. Host Dr. Gauri Joshi.
2. **Adobe Research, San Jose, USA:** From 22nd January, 2018 to 20th April, 2018. Host Dr. Branislav Kveton.
3. **INRIA, SequeL Lab, Lille, France:** Frrom 1st September, 2017 to 28th November, 2017. Host Dr. Odalric Maillard.

Master's Thesis

Finite-time Analysis of Frequentist Strategies for Multi-armed Bandits
This thesis studies the following topics in the area of Reinforcement Learning: Multi-armed bandits in stationary distribution with the goal of cumulative regret minimization and Thresholding bandits in pure exploration setting. The common

underlying theme is the study of bandit theory and its application in various types of environments. In the first part of the thesis, we study the classic multi-armed bandit problem with a stationary distribution, one of the first settings studied by the bandit community and which successively gave rise to several new directions in bandit theory. We propose a novel algorithm in this setting and compare both theoretically and empirically its performance against the available algorithms. Our proposed algorithm termed as Efficient-UCB-Variance (EUCBV) is the first arm-elimination algorithm which uses variance estimation to eliminate arms as well as achieve an order optimal regret bound. Empirically, we show that EUCBV outperforms most of the state-of-the-art algorithms in the considered environments. In the next part, we study a specific type of stochastic multi-armed bandit setup called the thresholding bandit problem and discuss its usage, available state-of-the-art algorithms on this setting and our solution to this problem. We propose the Augmented-UCB (AugUCB) algorithm which again uses variance and mean estimation along with arm elimination technique to conduct exploration. We give theoretical guarantees on the expected loss of our algorithm and also analyze its performance against state-of-the-art algorithms in numerical simulations in multiple synthetic environments. [Thesis]

Research Projects

Thresholding Bandits with Augmented UCB
Proposed the Augmented-UCB (AugUCB) algorithm for a fixed-budget version of the thresholding bandit problem (TBP), where the objective is to identify a set of arms whose quality is above a threshold. A key feature of AugUCB is that it uses both mean and variance estimates to eliminate arms that have been sufficiently explored. This is the first algorithm to employ such an approach for the considered TBP setting. [Paper]

Efficient UCBV: An Almost Optimal Algorithm using Variance Estimates
Proposed a novel algorithm for the stochastic multi-armed bandit (MAB) problem. Our proposed Efficient UCB Variance method, referred to as EUCBV is an arm elimination algorithm based on UCB-Improved and UCBV strategy which takes into account the empirical variance of the arms and along with aggressive exploration factors eliminate sub-optimal arms. Through a theoretical analysis, we establish that EUCBV achieves a better gap-dependent regret upper bound than UCB-Improved, MOSS, UCB1, and UCBV algorithms. EUCBV enjoys an order optimal gap-independent regret bound same as that of OCUCB and MOSS, and better than UCB-Improved, UCB1 and UCBV. [Paper]

Distribution-dependent and Time-uniform Bounds for Piecewise i.i.d Bandits
We consider the setup of stochastic multi-armed bandits in the case when reward distributions are piecewise i.i.d. and bounded with unknown changepoints. We focus on the case when changes happen simultaneously on all arms, and in stark contrast with the existing literature, we target gap-dependent (as opposed to only gap-independent) regret bounds involving the magnitude of changes and optimality-gaps. Under a slightly stronger set of assumptions, we show that as long as the compounded delayed detection for each changepoint is bounded there is no need for extra exploration to actively detect changepoints. We introduce two adaptations of UCB-strategies that employ scan-statistics in order to actively detect the changepoints, without knowing in advance the changepoints and also the mean before and after any change. Our first method UCBLaplace-CPD does not know the number of changepoints G or time horizon T and achieves the first time-uniform

concentration bound for this setting using the Laplace method of integration. The second strategy ImpCPD makes use of the knowledge of T to achieve the order optimal regret bound of thereby closing an important gap with respect to the lower bound in a specific setting. Our theoretical findings are supported by numerical experiments on synthetic and real-life datasets. [Paper]

Non-Stochastic Low Rank Bandit
We study the problem of learning the maximum entry of a low-rank non-negative matrix, from sequential observations. In this setting, the learner chooses tuples of rows and columns at every round and observes the product of their values. The main challenge in this setting is that the learner does not observe the individual latent values of rows and columns as its feedback. Diverging from previous works we assume that the preference matrix is non-stochastic and hence our setting is more general in nature. Existing methods for solving similar problems rely on UCB-type algorithms based on constructing conservative confidence intervals with the strong assumption that underlying distributions are stochastic. We depart from this standard approach and consider the case when the best row and column pair can be learned jointly with help of two separate bandit algorithms working individually on rows and columns. We propose a simple and computationally efficient algorithm that implements this procedure, which we call Low Rank Bandit, and prove a sub-linear bound on its n-step regret in the rank-1 special case. We evaluate the algorithm empirically on several synthetic and real-world datasets. In all experiments, we outperform existing state-of-the-art algorithms. [Paper]

Collaborators

1. Dr. Robert Nowak, ECE Department, UW-Madison
2. Dr. Balaraman Ravindran, CSE Department, IIT Madras
3. Dr. Nandan Sudarsanam, Department of Management Science, IIT Madras
4. Dr. K.P. Naveen, Deprtment of Electrical Engineering, IIT Tirupati
5. Dr. Odalric-Ambrym Maillard, INRIA, SequeL Lab, Lille, France
6. Dr. Branislav Kveton, Google Research, Mountain View, USA
7. Dr. Gauri Joshi, ECE Department, CMU, Pittsburgh

Teaching Experience

Teaching Assistant, UMass Amherst 2018–current
Assisted in preparing and conducting lab assignments and class tutorials for the following courses:
Natural Language Processing - Prof. Mohit Iyyer
Design of Algorithms - Prof. Daniel Sheldon
Teaching Assistant, IIT Madras 2015–2018
Assisted in preparing and conducting lab assignments and class tutorials for the following courses:
Introduction to Programming - Prof. Raghavendra Rao B. V.
Reinforcement Learning(twice) - Prof. Balaraman Ravindran
Compiler Design - Prof. Rupesh Nasre

Work Experience

Tata Consultancy Services Ltd., Kolkata, India March 2014–December 2014
Assistant System Engineer Trainee
Software development and test engineer in Digital Enterprise Service and Solution.

Professional Activities

Reviewer

1. Assisted Dr. Balaraman Ravindran in reviewing for IJCAI 2017.
2. Assisted Dr. Branislav Kveton in reviewing for ICML 2018.

Volunteer

1. Assisted Dr. Balaraman Ravindran in conducting the *"Recent Advances in Reinforcement Learning, 2015"* workshop held at IIT Madras. Some of the key speakers include, Dr. Richard Sutton, Dr. Csaba Szepesvari, Dr. Sridhar Mahadevan, and Dr. Satindar Singh.

Relevant Coursework [more information]

Introduction to Machine Learning	Reinforcement Learning
Natural Language Processing	Linear Algebra and Random Processes
Multi-variate Data Analysis	Data Analysis for Research
Artificial Intelligence	Design and Analysis of Algorithms

Award Grants and Fellowship

1. University of Wisconsin-Madison Chancellor's Opportunity Fellowship 2019-2020.
2. IIT Madras student travel grant of USD 2300.
3. Google travel grant of USD 1700.
4. AAAI grant of USD 500.
5. Microsoft travel grant of USD 1435.
6. University of Wisconsin-Madison ECE Welcome Award of USD 3000.

Other Achievements

Scored 318/340 in Graduate Record Examinations **(GRE)** 2018.
Scored 111/120 in Test of English as a Foreign Language **(TOEFL)** 2017.
Ranked 1150/155190 candidates in Graduate Aptitude Test in Engineering **(GATE)** 2014.
Secured 98.93 percentile in Common Admission Test **(CAT)** 2014 among 196988 candidates.

References

Dr. Balaraman Ravindran
Professor
`ravi@cse.iitm.ac.in`
Department of Computer Science & Engg.
Indian Institute of Technology Madras

Dr. Nandan Sudarsanam
Assistant Professor
`nandan@iitm.ac.in`
Department of Management Studies
Indian Institute of Technology Madras

Dr. K.P. Naveen
Assistant Professor
`naveenkp@iittp.ac.in`
Department of Electrical Engg.
Indian Institute of Technology Tirupati

Dr. Odalric Maillard
INRIA Researcher (CR1)
`odalricambrym.maillard @ inria.fr`
SequeL Team
INRIA Lille, France

Dr. Branislav Kveton
Machine Learning Scientist
`kveton@google.com`
Google Research, Mountain View, CA, USA