SUBHOJYOTI MUKHERJEE

ALL Lab & Bio-NLP Group
College of Information and Computer Sciences
University of Massachusetts Amherst
Amherst, MA 01002

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

Research Interests Machine learning, Reinforcement learning, Online Learning, Multi-armed bandits, Recommender Systems, Applied Probability, Optimization.

Education

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

Indian Institute of Technology Madras, India — 2015–2018
M.S, Computer Science
Advisers: Dr. Balaraman Ravindran and Dr. Nandan Sudarsanam
CGPA: 8.4/10

Meghnad Saha Institute of Technology, Kolkata, India — 2009–2013
Bachelor of Technology, Computer Science & Engineering
CGPA: 8.4/10

Publications

1. **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. [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. Accepted for oral presentation.[Paper]
3. **Subhojyoti Mukherjee**, and Odalric-Ambrym-Maillard, "*Improved Changepoint Detection in Piece-wise Stochastic Bandits*", *Under Review in Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics **(AISTATS-19)***, main conference track. [Paper]

Research Internships

Adobe Research, San Jose: Research internship under Dr. Branislav Kveton in the Adobe Research, San Jose, USA from 22nd January, 2018 to 20th April, 2018 for a period of 3 months.

INRIA, SequeL Lab: Research internship under Dr. Odalric Maillard in the INRIA Sequel Lab, Lille, France from 1st September, 2017 to 28th November, 2017 for a period of 3 months.

Master's Thesis 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]

B.Tech Project

This project studies the area of Sentiment Analysis in Natural Language Processing. Identifying the sentiment of a movie review or a product review from user comments forms a vital form of feedback in recommender systems. The learning algorithm can use this feedback to understand the recent trends and then suggest an interesting item to a user that will generate its interest. We develop an algorithm that takes input a recent trending topic in the internet which then crawls the Twitter in identifying the sentiments of the user regarding the topic from their associated tweets and then outputs whether the general sentiment is positive, negative or neutral regarding the topic. The algorithm uses bag-of-words model where it uses several existing dictionaries to store the sentiment of words before-hand to output the general sentiment regarding the topic.

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.

Efficient UCBV: An Almost Optimal Algorithm using Variance Estimates

Presented 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.

Improved Changepoint Detection in 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. with unknown changepoints. Out of generality, we assume the reward distributions to be bounded and thus do not restrict to spe-

cific parametric exponential families. Due to the regret minimization objective, we study the change of mean, in the context when not only the change times are unknown, but also the mean before and after any change. 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. We introduce two simple adaptations of UCB-strategies that employ scan-statistics in order to actively detect the changepoints, without knowing in advance the number of changepoints G. We also derive gap-independent regret bounds. The first strategy UCB-CPD does not know the time horizon T and achieve a $O(\sqrt{GT}\log T)$ regret bound, while the second strategy ImpCPD makes use of the knowledge of T to remove the $\log T$ dependency thereby closing an important gap with respect to the lower bound. Empirically, ImpCPD outperforms most of the passive and adaptive algorithms except the oracle-based algorithms that have access to the exact changepoints in all the considered environments.

Latent Ranked Bandit
We study the problem of learning personalized ranked lists of diverse items for multiple users, from sequential observations of user preferences. The user-item preference matrix is non-negative and low-rank. Existing methods for solving similar problems are based on reconstructing the preference matrix from its noisy observations using matrix factorization techniques, and typically require strong assumptions on the reconstructed matrix. We depart from this standard approach and consider a family of low-rank matrices, where the set of most preferred items of all users is small and can be learned efficiently. Then we learn to present this set to each user in a personalized manner, in the order of the descending preferences of the user. We propose a computationally efficient algorithm that implements this procedure, and prove a sublinear bound on its n-step regret. We evaluate the algorithm empirically on several synthetic and real-world datasets. In all experiments, we outperform existing state-of-the-art algorithms.

Collaborators

1. Dr. Balaraman Ravindran, CSE Department, IIT Madras
2. Dr. Nandan Sudarsanam, Department of Management Science, IIT Madras
3. Dr. K.P. Naveen, Deprtment of Electrical Engineering, IIT Tiruapti
4. Dr. Odalric-Ambrym Maillard, INRIA, SequeL Lab, Lille, France
5. Dr. Branislav Kveton, Google Research, Mountain View, USA
6. Dr. Anup Rao, Adobe Research, San Jose, USA

Teaching Experience

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 - 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

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

Fundamentals of Experimentation for Management

Award and Grants

1. Our paper titled "Thresholding Bandits with Augmented UCB" was awarded IIT Madras student travel grant of USD 2300.
2. Our paper titled "Efficient UCBV: An Almost Optimal Algorithm using Variance Estimates" was awarded Google travel grant of USD 1700 and AAAI grant of USD 500.

Other Achievements

Scored 314/340 in Graduate Record Examinations **(GRE)** 2017.
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