SUBHOJYOTI MUKHERJEE

RISE lab
Department of Computer Science & Engineering
Indian Institute of Technology Madras
Chennai, India 600036

Phone: +91 97486 83510
Email: subho@cse.iitm.ac.in,
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

Indian Institute of Technology Madras, Chennai, India   January 2015–present
M.S, Computer Science & Engineering
Guides: Dr. Balaraman Ravindran and Dr. Nandan Sudarsanam
CGPA: 8.40/10

Meghnad Saha Institute of Technology, Kolkata, India   2009–2013
Bachelor of Technology, Computer Science & Engineering
CGPA: 8.42/10

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.

2. **Subhojyoti Mukherjee**, K.P. Naveen, Nandan Sudarsanam, and Balaraman Ravindran, "*Efficient UCBV: An Almost Optimal Algorithm using Variance Estimates*", *To appear in Proceedings of the Thirty-Second Association for the Advancement of Artificial Intelligence **(AAAI-18)***, main conference track. Accepted for oral presentation.

3. **Subhojyoti Mukherjee**, and Odalric-Ambrym-Maillard, "*Improved Changepoint Detection in Piece-wise Stochastic Bandits*", *Under Review in Proceedings of the Thirty-Fifth International Conference on Machine Learning **(ICML-18)***, main conference track.

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.

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   **Thresholding Bandits with Augmented UCB**
PROJECTS   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 Piece-wise Stochastic Bandits
We proposed two UCB algorithms, UCB-CPD and ImpCPD for the piecewise stochastic environment. Both these algorithms are adaptive and try to locate the change-points and restart, thereby performing better than the passive algorithms. Theoretically, ImpCPD achieves an order optimal regret bound of $O\left(G\sqrt{T}\right)$, 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 in all the considered environments.

Generalized Latent Bandits
We study the problem of recommending the best items to users who are coming sequentially. The learner has access to very less prior information about the users and it has to adapt quickly to the user preferences and suggest the best item to each user. Furthermore, we consider the setting where users are grouped into clusters and within each cluster the users have the same choice of the best item, even though their quality of preference may be different for the best item. These

clusters along with the choice of the best item for each user are unknown to the learner. Also, we assume that each user has a single best item preference. This complex problem can be conceptualized as a low rank stochastic bandit problem where the goal of the learner is to minimize the cumulative regret by quickly identifying the best item for each user.

RESEARCH INTERNSHIP

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.

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.

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, Adobe Research, San Jose, USA

TEACHING EXPERIENCE

Teaching Assistant, IIT Madras   January 2015–present
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.

Workshop 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	

Awards

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
Associate 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@adobe.com
BigData Experience Lab
Adobe Research, San Jose, CA, USA