

Electrical Computer Engineering & Ph.D. Applicant
 University of Wisconsin Madison Subhojyoti Mukherjee (subho@cs.umass.edu) [CV]

1 Completed/Ongoing Research

1. Variance aware Bandits: An important problem which arises in mobile channel recommendation and online item recommendation where there is a separation between testing and deployment phases is how to employ these bandits in a pure-exploration setting. In the pure-exploration [Thresholding Bandits](#) (Mukherjee, Subhojyoti et al., 2017) we proposed an algorithm which finds the best set of actions with qualities above a particular threshold. The challenge lies in the observation that this best set can be very large and moreover that the candidate set of actions can be exponentially large. Our novel algorithm solves this problem by estimating the mean and the *variance* (hence being variance-aware) of the actions to determine their qualities, thereby significantly reducing the exploration. We also provide theoretical bounds for the maximum possible loss suffered by our method. This work was published in **International Joint Conference on Artificial Intelligence 2017**.

2. Changepoint detection and piecewise i.i.d Bandits: Applying Bandit strategies in environments where the quality of all the actions change abruptly and simultaneously has proved to be quite challenging. This is quite common in *medical domain* where the behavior of drug-resistant bacteria may change abruptly prompting a different response to all treatments (actions). We proposed two novel algorithms called [UCB-CPD](#) and [ImpCPD](#) for this setting with interesting theoretical guarantees and closed some gaps in the existing bounds. This work was jointly done with [Dr. Odalric Maillard](#) while I was an intern at INRIA, SequeL Lab, Lille and is currently under review in **International Conference on Artificial Intelligence and Statistics 2019**.

These works and more fundamental insights into several aspects of Bandits, RL and Learning theory have been captured in my [MS thesis](#) which was accepted by my technical committee at IIT Madras without any major revision.

2 Future Directions

"Make Algorithms Trustworthy Again (#MATA)"

There are several domains where the need of the hour is to provide safe and "trustworthy" algorithms that can predict correct actions with high confidence. Working on theoretical research in RL setting now allows me to expand my research frontier in a variety of new but related research areas including but not limited to statistical learning theory and optimization and to explore problems that bridge the gap between learning theory and empiricism. A few of the interesting research problems that I intend to work are,

1. Off-policy Reinforcement Learning: Often the complete knowledge of the world is not available to the learner forcing it to adopt different strategies. Examples include maze navigation by a robot maze using faulty sensors, medical treatments with only a partial knowledge about the true state of the patient (as it requires information to an atomic level), and user preference recommender systems without cumbersome exploration strategies of complete user logs. Partial observability is a broad area of research in RL, and providing theoretical guarantees is hard due to the high variance in the performance of various strategies given only a limited state-space exposure. I can draw from my previous experience as this problem boils down to intelligent exploration for the reduction of performance variance and to provide theoretical guarantees.

2. Safe Reinforcement Learning: I am interested in safe RL which also poses some of the same challenges as above. For example, in the medical domain, not only do we have to work in a partially observable environment but we should also give confidence guarantees for the proposed action (medical treatment) since this gives confidence to the medical practitioner in a high-risk environment where faulty actions might cause death. This is again directly related to my previous experience as this requires a sound knowledge of concentration inequalities to give the confidence on an action taken. Safe RL is of extreme importance in areas of robotics as well, for example in autonomous driving, and motion planning and control.

3. High Dimensional Reinforcement Learning: There has been several recent successes of RL in large state spaces such as the game Go and autonomous driving, wherein the success has come by employing powerful nonlinear function approximation techniques like Neural Networks. But sound theoretical guarantees for RL only exist for small state spaces using linear function approximation techniques. I am interested in bridging this gap between the theoretical and empirical performances of RL as this will directly benefit some of the challenges mentioned before, and I can draw my knowledge from concentration measure theory and RL to contribute to this problem.

4. Optimal Design & Active Learning: Many supervised learning tasks operate on limited labeled data, and hence, the learning models have high uncertainty in their prediction. Also in certain tasks, it is costly to run experiments, and hence, the learner needs to carefully pick subsets of data to minimize model uncertainty. There are many ways to address this problem including active learning, online learning, and bandit feedback. This is a combinatorial optimization problem having many open-ended questions and requiring strong theoretical guarantees, to which I can contribute.

3 Program and Faculty of Interest

I believe that a doctoral degree from the Department of Electrical & Computer Engineering will propel me a long way in my goal to become a professor. Its broad reach and in-depth courses on a variety of subjects ranging from practical to theoretical applications will help me in my endeavor. Even though my previous experience has mainly centered around different areas of Computer Science, working on learning theory has really made me appreciate the broad scope of information theory. Especially, the area I want to venture into during my doctoral studies requires strong fundamentals of information theory (such as stochastic random processes and concentration measure theory), and it is the main motivation in applying to this department. There are several professors at University of Wisconsin Madison such as **Dr. Robert Nowak**, **Dr. Karl Rohe**, and **Dr. Laurent Lessard** whose publications/projects are especially appealing to me. All of these faculties work in the areas that I know of and I am interested to work with them for my doctoral research. Also, I will be grateful if the admission committee can find a suitable fit for me in the department based on my research experience and interests.

References

- Mukherjee, Subhojyoti;** Kolar Purushothama, N.; Sudarsanam, N.; and Ravindran, B. 2017. Thresholding bandits with augmented ucb. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, 2515–2521.
- Mukherjee, Subhojyoti;** Kolar Purushothama, N.; Sudarsanam, N.; and Ravindran, B. 2018. Efficient-UCBV: An almost optimal algorithm using variance estimates. In *Proceedings of the 32nd Association for the Advancement of Artificial Intelligence*.