Sequential community mode estimation*

Shubham Anand Jain, Shreyas Goenka, Divyam Bapna, Nikhil Karamchandani, Jayakrishnan Nair
Department of Electrical Engineering, IIT Bombay

1. INTRODUCTION

Several applications in online learning involve sequential sampling/polling of an underlying population. A classical learning task in this space is *online cardinality estimation*, where the goal is to estimate the size of a set by sequential sampling of elements from the set (see, for example, [2,4,7]). The key idea here is to use 'collisions,' i.e., instances where the same element is sampled more than once, to estimate the size of the set. Another recent application is *community exploration*, where the goal of the learning agent is to sample as many distinct elements as possible, given a family of sampling distributions/domains to poll from (see [3,6]).

In this paper, we focus on the related problem of *community mode estimation*. Here, the goal of the learning agent is to estimate the largest community within a population of individuals, where each individual belongs to a unique community. The agent has access to a set of sampling domains, referred to as *boxes* in this paper, which also partition the population. The agent can, at any sampling epoch, choose which box to sample from. Having chosen one such box to sample from, a random individual from this box gets revealed to the agent, along with the community that individual belongs to. After a fixed budget of samples is exhausted, the learning agent reveals its estimate of the largest community (a.k.a., the community mode) in the population. The goal of the agent is in turn to minimize the probability of mis-identifying the community mode, by optimizing (i) the policy for sequential sampling of boxes, and (ii) the decision rule that determines the agent's response as a function of all observations.

One application that motivates this formulation is election polling. In this context, communities might correspond to the party/candidate an individual votes for, while boxes might correspond, for instance, to different cities/states that individuals reside in. In this case, community mode identification corresponds to predicting the winning party/candidate. A related (and contemporary) application is the detection of the dominant strain of a virus/pathogen within a population of infected individuals. Here, communities would correspond to different strains, and boxes would correspond to different regions/jurisdictions.

The formulation we consider here has some parallels with the classical multi-armed bandit (MAB) problem [11]; specifically, the fixed budget best arm identification formulation [1]. Indeed, one may interpret communities in our formulation as arms in an MAB problem. However, there are two crucial differences between the two formulations. The first difference lies in the stochastic behavior of the reward/observation sequence. In the classical MAB problem, each pull of an arm yields an i.i.d. reward drawn from an arm specific reward distribution. However, in the community mode detection problem, the sequence of collisions (or equivalently, the evolution of the number of distinct individuals seen) does not admit an i.i.d. description. (Indeed, whether or not a certain sample from a box results in a collision depends in a non-stationary manner on the history of observations from that box.) The second difference between the two formulations lies in the extent of sampling control on part of the agent. In the MAB setting, the agent can pull any arm it chooses at any sampling epoch. However, in our formulation, the agent cannot sample directly from a community of its choice; it must instead choose a box to sample from, limiting its ability to target specific communities to explore.

2. SUMMARY OF CONTRIBUTIONS

Our contributions are summarized as follows (all details in [8]).

- We begin by considering a special case of our model where the entire population is contained within a single box; we refer to this as the *mixed community setting*. In this setting, the sampling process is not controlled, and the learning task involves only the decision rule. We show that a simple decision rule, based on counting the number of distinct individuals encountered from each community, is optimal, via comparison of an upper bound on the probability of error (misidentification of the community mode) under the proposed algorithm with an information theoretic lower bound. For this setting, we also highlight the impact of being able to identify sampled individuals (i.e., determine whether or not the sampled individual has been seen before) on the achievable performance in community mode estimation.

- Next, we consider the case where each community lies in its own box; the so-called *separated community setting*. Here, we show that the commonly used approach of detecting pairwise collisions [6] is sub-optimal. Next, a near-optimal algorithm is proposed that borrows the

*This is an extended abstract corresponding to the paper [8], which appears in the proceedings of IFIP Performance 2021.

sampling strategy of the classical *successive rejects* policies for MABs [1], but differentiates communities based on the number of distinct individuals encountered (which is different from the classical MAB setting where arms are differentiated based on their empirical average rewards).

- Next, we consider a setting that encompasses both the mixed community as well as the separated community settings; we refer to it as the *community-disjoint box setting*. Here, each community is contained within a single box (though a box might contain multiple communities). For this case, we propose novel algorithms that combine elements from the mixed and separated community settings. Finally, we show how the algorithms designed for the community-disjoint box setting can be extended to the fully general case, where communities are arbitrarily spread across boxes.

- Finally, we validate the algorithms proposed on both synthetic as well as real-world datasets.

We conclude by making a comparison between our contributions and the literature on the fixed budget MAB problem. Near optimal algorithms for the fixed budget MAB problem (see, for example, [1, 9]) follow a sampling strategy of *successive rejection* of arms, wherein the sampling budget is split across multiple phases, and at the end of each phase, a certain number of (worst performing) arms are eliminated from further consideration. Some of our algorithms for the community mode estimation problem follow a similar sampling strategy and eliminate boxes in phases; specifically, we often use the same sampling schedule as in the classical successive rejects algorithm proposed in [1]. However, the elimination criterion we use is different: it is based on the number of distinct individuals seen (so far) from each community. Given that this statistic evolves in a non-stationary Markovian fashion over time, this distinction makes our analysis more complex.

Our information theoretic lower bounds are inspired by the framework developed in [10] for the fixed budget MAB problem. However, as before, the key distinction in our proofs stems from the difference in stochastic nature of the observation process: while reward observations for each arm in the classical MAB setup are i.i.d., the number of distinct individuals seen from each community evolves as an absorbing Markov chain in the community mode estimation problem.

3. REFERENCES

[1] J.-Y. Audibert and S. Bubeck. Best Arm Identification in Multi-Armed Bandits. In *COLT - Conference on Learning Theory*, 2010.

[2] M. Bressan, E. Peserico, and L. Pretto. Simple set cardinality estimation through random sampling. *arXiv preprint arXiv:1512.07901*, 2015.

[3] S. Bubeck, D. Ernst, and A. Garivier. Optimal discovery with probabilistic expert advice: finite time analysis and macroscopic optimality. *Journal of Machine Learning Research*, 14:601–623, 2013.

[4] C. Budianu, S. Ben-David, and L. Tong. Estimation of the number of operating sensors in large-scale sensor networks with mobile access. *IEEE Transactions on Signal Processing*, 54(5):1703–1715, 2006.

[5] A. Carpentier and A. Locatelli. Tight (lower) bounds for the fixed budget best arm identification bandit problem. In *Conference on Learning Theory*, 2016.

[6] X. Chen, W. Huang, W. Chen, and J. C. Lui. Community exploration: from offline optimization to online learning. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, 2018.

[7] M. Finkelstein, H. G. Tucker, and J. A. Veeh. Confidence intervals for the number of unseen types. *Statistics & Probability Letters*, 37(4):423–430, 1998.

[8] S. A. Jain, S. Goenka, D. Bapna, N. Karamchandani, and J. Nair. Sequential community mode estimation. *Performance Evaluation*, 2021.

[9] Z. Karnin, T. Koren, and O. Somekh. Almost optimal exploration in multi-armed bandits. In *International Conference on Machine Learning*, 2013.

[10] E. Kaufmann, O. Cappé, and A. Garivier. On the Complexity of Best Arm Identification in Multi-Armed Bandit Models. *Journal of Machine Learning Research*, 17:1–42, 2016.

[11] T. Lattimore and C. Szepesvári. *Bandit algorithms.* Cambridge University Press, 2020.