Optimal Rates for Online Bayesian Persuasion (Extended Abstract)

Martino Bernasconi[1,*], Matteo Castiglioni[1], Andrea Celli[2], Alberto Marchesi[1], Francesco Trovò[1] and Nicola Gatti[1]

[1]*Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan, Italy*
[2]*Department of Computing Sciences, Università Bocconi, Milan, Italy*

Abstract

Bayesian persuasion studies how an informed sender should influence beliefs of rational receivers who take decisions through Bayesian updating of a common prior. We focus on the *online Bayesian persuasion* framework, in which the sender repeatedly faces one or more receivers with unknown and adversarially selected types. We show how to obtain a tight $\widetilde{O}(T^{1/2})$ regret bound in the case in which the sender faces a single receiver and has partial feedback, improving over the best previously-known bound of $\widetilde{O}(T^{4/5})$. We also provide the first no-regret guarantees for the multi-receiver setting with partial feedback.

Keywords

Bayesian Persuasion, Online Learning

The *Bayesian persuasion* framework, introduced by Kamenica and Gentzkow [3], is an economic model which helps to explain how individuals make decisions based on the information they receive from others, and how this information can be used to influence their behavior. The framework has found application in advertising [4, 5, 6, 7, 8], voting [9, 10, 11, 12], routing [13, 14, 15], security [16, 17], sequential decision making [18, 19, 20], and incentivized exploration in multi-armed bandits [21, 22, 23, 24, 25].

In the simplest instantiation of the model, there are a sender and a receiver with a common prior over a finite set of states of nature. The sender publicly commits to a *signaling scheme*, which is a randomized mapping from states of nature to signals being sent to the receiver. Then, the sender observes the realized state of nature, and they send a signal to the receiver following the signaling scheme. The receiver observes the signal, computes their posterior distribution over states, and selects an action maximizing their expected utility. The sender and the receiver obtain a payoff which is a function of the receiver's action, and of the realized state of nature. An optimal signaling scheme for the sender is one maximizing their expected utility.

The study of Bayesian persuasion from a computational perspective was initiated by Dughmi and Xu [26], and the original model was later extended to more complex settings such as games

IPS-RCRA-SPIRIT 2023: Italian Workshop on Planning and Scheduling, RCRA Workshop on Experimental evaluation of algorithms for solving problems with combinatorial explosion, and SPIRIT Workshop on Strategies, Prediction, Interaction, and Reasoning in Italy. November 7-9th, 2023, Rome, Italy [1]

* An extended version of this work has been published in ICML 2023 [2].

*Corresponding author.

martino.bernasconideluca@polimi.it (M. Bernasconi); matteo.castiglioni@polimi.it (M. Castiglioni); andrea.celli2@unibocconi.it (A. Celli); alberto.marchesi@polimi.it (A. Marchesi); francesco1.trovo@polimi.it (F. Trovò); nicola.gatti@polimi.it (N. Gatti)

with multiple receivers [27, 14, 28]. A key question that has emerged is whether computational techniques can be used to ease some of the assumptions made in the original model by Kamenica and Gentzkow [3]. Two main lines of research have emerged: one is aimed at developing robust algorithms that can bypass the common-prior assumption [29, 30], and the other is focused on the robustness of persuasion when the sender is unaware of the receiver's goals [31, 32, 33].

This work follows the second perspective, and studies the *online Bayesian persuasion* framework introduced by Castiglioni et al. [31, 34], where the sender repeatedly faces a receiver whose type is unknown and chosen adversarially at each round from a finite set of types.

We start by describing a general no-regret algorithm for online learning against an oblivious adversary with a *finite* number of possible loss functions. We use this algorithm to provide a tight $\tilde{O}(T^{1/2})$ regret upper bound in the setting with one receiver and partial feedback, improving over the $\tilde{O}(T^{4/5})$ rate by Castiglioni et al. [31]. This result also improves the best known bound of $\tilde{O}(T^{2/3})$ for online learning in repeated Stackelberg games provided by Balcan et al. [35]. Then, we show that our general framework can be applied to obtain the first no-regret guarantees under partial feedback in the multi-receiver setting introduced by Castiglioni et al. [32]. In particular, we provide a tight $\tilde{O}(T^{1/2})$ regret bound under the assumption the set of possible type profiles of the receivers is known beforehand by the sender. In each of these settings, our no-regret algorithms may suffer from exponential per-iteration running time, as expected from known hardness results for the online Bayesian persuasion settings [31]. In the last part of the paper, we provide the first no-regret algorithms for online Bayesian persuasion with guaranteed polynomial per-iteration running time. We do that by considering the *type reporting* framework introduced by Castiglioni et al. [36], where the sender can commit to a *menu* of signaling schemes, and then let the receivers choose their preferred signaling scheme depending on their private types. In such a setting, we provide a $O(T^{1/2})$ regret upper bound for the single-receiver setting. Moreover, by designing a general procedure based on the follow the regularized leader algorithm, we show that it is possible to achieve the same rate of convergence with polynomial-time per-iteration time complexity also in the multi-receiver setting, when receivers have binary actions and the utility of the sender is specified by a function of receivers' actions that is either supermodular or anonymous.

The main motivation for introducing the reduction from online problems with finite number of losses to online linear optimization was to solve online Bayesian persuasion problems. However, this result finds applicability in other settings such as learning in security games and bidding in combinatorial auctions.

Learning in security games [35] extends classic (one-shot) security games (see, *e.g.*, [37]) by introducing the problem of learning a no-regret strategy for the defender against a sequence of attackers that is adversarially selected. In this model, at each round t, the defender chooses a strategy $\boldsymbol{x}_t$, which is a distribution over N targets. Then, an attacker of type $d_t \in D$ best responds to such a strategy and the defender experiences a loss of $L_{d_t}(\boldsymbol{x}_t)$. Our reduction yields a $\tilde{O}(\text{poly}(D)\sqrt{T})$ regret bound under partial feedback, which improves the previously-known regret bound given by Balcan et al. [35], which is of order $O(\text{poly}(ND)T^{2/3})$.

Another application of our framework is bidding in repeated combinatorial auctions [38]. In these auctions the action space is combinatorial and, therefore, exponentially large. Our reduction to online linear optimization gives a $\tilde{O}(\text{poly}(D)\sqrt{T})$ bound for this problem under partial feedback.

Acknowledgments

This paper is supported by the Italian MIUR PRIN 2022 Project "Targeted Learning Dynamics: Computing Efficient and Fair Equilibria through No-Regret Algorithms", by the FAIR (Future Artificial Intelligence Research) project, funded by the NextGenerationEU program within the PNRR-PE-AI scheme (M4C2, Investment 1.3, Line on Artificial Intelligence), and by the EU Horizon project ELIAS (European Lighthouse of AI for Sustainability, No. 101120237).

References

[1] D. F. V. M. M. M. E. S. L. S. I. S. E. T. A. U. M. V. Riccardo De Benedictis, Matteo Castiglioni, Preface to the Italian Workshop on Planning and Scheduling, RCRA Workshop on Experimental evaluation of algorithms for solving problems with combinatorial explosion, and SPIRIT Workshop on Strategies, Prediction, Interaction, and Reasoning in Italy (IPS-RCRA-SPIRIT 2023), in: Proceedings of the Italian Workshop on Planning and Scheduling, RCRA Workshop on Experimental evaluation of algorithms for solving problems with combinatorial explosion, and SPIRIT Workshop on Strategies, Prediction, Interaction, and Reasoning in Italy (IPS-RCRA-SPIRIT 2023) co-located with 22th International Conference of the Italian Association for Artificial Intelligence (AI* IA 2023), 2023.

[2] M. Bernasconi, M. Castiglioni, A. Celli, A. Marchesi, F. Trovò, N. Gatti, Optimal rates and efficient algorithms for online Bayesian persuasion, in: A. Krause, E. Brunskill, K. Cho, B. Engelhardt, S. Sabato, J. Scarlett (Eds.), Proceedings of the 40th International Conference on Machine Learning, volume 202 of *Proceedings of Machine Learning Research*, PMLR, 2023, pp. 2164–2183.

[3] E. Kamenica, M. Gentzkow, Bayesian persuasion, American Economic Review 101 (2011) 2590–2615.

[4] P. Bro Miltersen, O. Sheffet, Send mixed signals: earn more, work less, in: Proceedings of the 13th ACM Conference on Electronic Commerce, 2012, pp. 234–247.

[5] Y. Emek, M. Feldman, I. Gamzu, R. PaesLeme, M. Tennenholtz, Signaling schemes for revenue maximization, ACM Transactions on Economics and Computation 2 (2014) 1–19.

[6] A. Badanidiyuru, K. Bhawalkar, H. Xu, Targeting and signaling in ad auctions, in: Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, 2018, pp. 2545–2563.

[7] M. Castiglioni, G. Romano, A. Marchesi, N. Gatti, Signaling in posted price auctions, Proceedings of the AAAI Conference on Artificial Intelligence 36 (2022) 4941–4948.

[8] F. Bacchiocchi, M. Castiglioni, A. Marchesi, G. Romano, N. Gatti, Public signaling in bayesian ad auctions, in: Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI 2022, Vienna, Austria, 23-29 July 2022, ijcai.org, 2022, pp. 39–45.

[9] Y. Cheng, H. Y. Cheung, S. Dughmi, E. Emamjomeh-Zadeh, L. Han, S.-H. Teng, Mixture selection, mechanism design, and signaling, in: 56th Annual Symposium on Foundations of Computer Science, 2015, pp. 1426–1445.

[10] R. Alonso, O. Câmara, Persuading voters, American Economic Review 106 (2016) 3590–3605.
[11] M. Castiglioni, A. Celli, N. Gatti, Persuading voters: It's easy to whisper, it's hard to speak loud, in: The Thirty-Fourth AAAI Conference on Artificial Intelligence, 2020, pp. 1870–1877.
[12] M. Castiglioni, N. Gatti, Persuading voters in district-based elections, in: Proceedings of the AAAI Conference on Artificial Intelligence, volume 35, 2021, pp. 5244–5251.
[13] S. Vasserman, M. Feldman, A. Hassidim, Implementing the wisdom of waze, in: Twenty-Fourth International Joint Conference on Artificial Intelligence, 2015, pp. 660–666.
[14] U. Bhaskar, Y. Cheng, Y. K. Ko, C. Swamy, Hardness results for signaling in Bayesian zero-sum and network routing games, in: Proceedings of the 2016 ACM Conference on Economics and Computation, 2016, pp. 479–496.
[15] M. Castiglioni, A. Celli, A. Marchesi, N. Gatti, Signaling in bayesian network congestion games: the subtle power of symmetry, in: Proceedings of the AAAI Conference on Artificial Intelligence, volume 35, 2021, pp. 5252–5259.
[16] Z. Rabinovich, A. X. Jiang, M. Jain, H. Xu, Information disclosure as a means to security, in: Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, 2015, pp. 645–653.
[17] H. Xu, R. Freeman, V. Conitzer, S. Dughmi, M. Tambe, Signaling in Bayesian Stackelberg games, in: Proceedings of the 2016 International Conference on Autonomous Agents and Multiagent Systems, 2016, pp. 150–158.
[18] J. Wu, Z. Zhang, Z. Feng, Z. Wang, Z. Yang, M. I. Jordan, H. Xu, Sequential information design: Markov persuasion process and its efficient reinforcement learning, in: Proceedings of the 23rd ACM Conference on Economics and Computation, 2022, p. 471–472.
[19] J. Gan, R. Majumdar, G. Radanovic, A. Singla, Bayesian persuasion in sequential decision-making, in: Proceedings of the AAAI Conference on Artificial Intelligence, volume 36, 2022, pp. 5025–5033.
[20] M. Bernasconi, M. Castiglioni, A. Marchesi, N. Gatti, F. Trovò, Sequential information design: Learning to persuade in the dark, in: NeurIPS, 2022.
[21] I. Kremer, Y. Mansour, M. Perry, Implementing the "wisdom of the crowd", Journal of Political Economy 122 (2014) 988–1012.
[22] L. Cohen, Y. Mansour, Optimal algorithm for bayesian incentive-compatible exploration, in: Proceedings of the 2019 ACM Conference on Economics and Computation, 2019, pp. 135–151.
[23] Y. Mansour, A. Slivkins, V. Syrgkanis, Z. S. Wu, Bayesian exploration: Incentivizing exploration in Bayesian games, in: Proceedings of the 2016 ACM Conference on Economics and Computation, 2016, pp. 661–661.
[24] M. Sellke, A. Slivkins, The price of incentivizing exploration: A characterization via thompson sampling and sample complexity, in: Proceedings of the 22nd ACM Conference on Economics and Computation, 2021, pp. 795–796.
[25] Y. Mansour, A. Slivkins, V. Syrgkanis, Z. S. Wu, Bayesian exploration: Incentivizing exploration in Bayesian games, Operations Research 70 (2022) 1105–1127.
[26] S. Dughmi, H. Xu, Algorithmic Bayesian persuasion, in: Proceedings of the forty-eighth annual ACM symposium on Theory of Computing, 2016, pp. 412–425.

[27] S. Dughmi, H. Xu, Algorithmic persuasion with no externalities, in: Proceedings of the 2017 ACM Conference on Economics and Computation, 2017, pp. 351–368.
[28] H. Xu, On the tractability of public persuasion with no externalities, in: Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, 2020, pp. 2708–2727.
[29] Y. Zu, K. Iyer, H. Xu, Learning to persuade on the fly: Robustness against ignorance, Proceedings of the 22nd ACM Conference on Economics and Computation (2021) 927–928.
[30] M. K. Camara, J. D. Hartline, A. Johnsen, Mechanisms for a no-regret agent: Beyond the common prior, in: 2020 ieee 61st annual symposium on foundations of computer science (focs), IEEE, 2020, pp. 259–270.
[31] M. Castiglioni, A. Celli, A. Marchesi, N. Gatti, Online Bayesian persuasion, Advances in Neural Information Processing Systems 33 (2020) 16188–16198.
[32] M. Castiglioni, A. Marchesi, A. Celli, N. Gatti, Multi-receiver online bayesian persuasion, in: International Conference on Machine Learning, PMLR, 2021, pp. 1314–1323.
[33] Y. Babichenko, I. Talgam-Cohen, H. Xu, K. Zabarnyi, Regret-minimizing Bayesian persuasion, arXiv preprint arXiv:2105.13870 (2021).
[34] M. Castiglioni, A. Celli, A. Marchesi, N. Gatti, Regret minimization in online bayesian persuasion: Handling adversarial receiver's types under full and partial feedback models, Artificial Intelligence 314 (2023) 103821.
[35] M.-F. Balcan, A. Blum, N. Haghtalab, A. D. Procaccia, Commitment without regrets: Online learning in Stackelberg security games, in: Proceedings of the Sixteenth ACM Conference on Economics and Computation, 2015, p. 61–78.
[36] M. Castiglioni, A. Marchesi, N. Gatti, Bayesian persuasion meets mechanism design: Going beyond intractability with type reporting, 2022, pp. 226–234.
[37] M. Tambe, Security and game theory: algorithms, deployed systems, lessons learned, Cambridge university press, 2011.
[38] C. Daskalakis, V. Syrgkanis, Learning in auctions: Regret is hard, envy is easy, Games and Economic Behavior (2022).