Mechanical & Industrial Engineering
UNIVERSITY OF TORONTO

Master Project

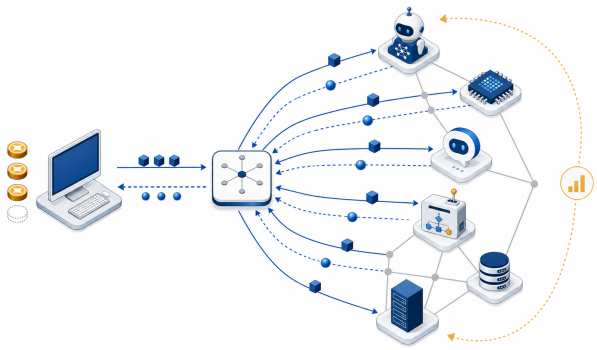
Learner-Robust Resource Allocation for Adaptive and Strategic AI Providers

Donya Jafari, Department of Mechanical & Industrial Engineering, University of Toronto
donya.jafari@mail.utoronto.ca

Peyman Mohajerin Esfahani, Department of Mechanical & Industrial Engineering, University of Toronto
p.mohajerinesfahani@utoronto.ca

Context

A client repeatedly assigns tasks or limited resources to a set of service providers. For each assignment, the client observes the cost and the quality of the result. The objective is to select providers that offer reliable service at a reasonable cost.



In a standard multi-armed bandit, every provider has a fixed but unknown reward distribution. The client learns this distribution by assigning tasks and observing outcomes. This assumption may not hold when providers are adaptive. A provider may learn from previous allocations, change its effort, misreport its cost or expected quality, or adjust its behaviour to receive more tasks.

The provider population may also be heterogeneous. Some providers may use reinforcement learning, no-regret learning, or language-model agents. Others may follow simple rules or remain fixed. The client usually does not know which learning rule each provider uses.

This project studies how to design an allocation and incentive mechanism that performs well across these different behaviours. The mechanism should achieve good quality and low cost while limiting the benefit that a provider can obtain through manipulation. The main goal is robustness to provider-learning strategies that were not known when the mechanism was designed.

The proposed approach is to evaluate and train mechanisms against a *population* of provider models rather than one fixed agent type. The population can include truthful providers, fixed heuristics, no-regret learners, reinforcement-learning policies, and language-model agents.

Each experiment runs a repeated interaction. The mechanism assigns a task or resource, the selected provider responds, and the outcome is recorded. Providers then update their policies using their own learning rules. This creates a feedback loop in which both the mechanism and the providers adapt over time.

The first objective is to build a benchmark for this setting. Existing allocation methods will be evaluated on familiar provider types and on held-out learning algorithms that were not used during mechanism design. The benchmark will measure output quality, cost, social welfare, regret, and manipulation gain.

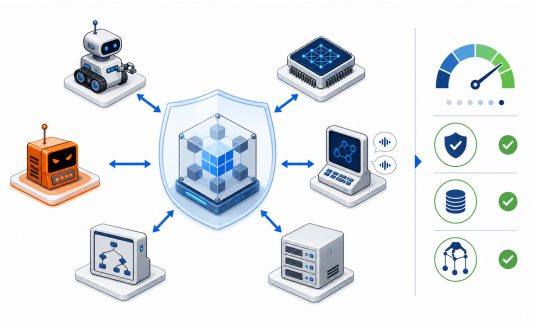


Figure 1: The mechanism is tested against different provider behaviours, including an adversarial response, and audited for quality, cost, and manipulation.

If time permits, the project will develop a population-robust mechanism or an auditing method. The mechanism may be trained against a mixture of provider models, with greater weight given to behaviours that currently exploit it. An independent auditor can search for profitable deviations and identify failure cases.

Project tasks

The project can roughly be divided into the following parts:

1. Review recent work on strategic bandits, online learning, multi-agent reinforcement learning, and dynamic mechanism design.

2. Build a repeated resource-allocation benchmark with several fixed, learning, and strategic provider models.

3. Implement and compare existing allocation or incentive mechanisms under familiar and held-out provider behaviours.

4. Develop and evaluate a population-robust mechanism or auditing method, if time permits.

The ideal candidate has some experience with Python and an interest in machine learning, optimization, or game theory. Familiarity with reinforcement learning, bandits, multi-agent systems, or mathematical optimization is helpful but not required.

Application: Please submit your CV, unofficial transcript, and a short paragraph describing your interest in the project by email to p.mohajerinesfahani@utoronto.ca and donya.jafari@mail.utoronto.ca.

Start Date: Fall 2026 / Winter 2027

Research Areas: Operations Research, Applied Machine Learning, Information Engineering