Novelty Assessment Report

Paper: Lipschitz Bandits with Stochastic Delayed Feedback
PDF URL: https://openreview.net/pdf?id=dfoN64vP4Q
Venue: ICLR 2026 Conference Submission
Year: 2026
Report Generated: 2026-01-04

Abstract

The Lipschitz bandit problem extends stochastic bandits to a continuous action set defined over a metric space, where the expected reward function satisfies a Lipschitz condition. In this work, we introduce a new problem of Lipschitz bandit in the presence of stochastic delayed feedback, where the rewards are not observed immediately but after a random delay. We consider both bounded and unbounded stochastic delays, and design algorithms that attain sublinear regret guarantees in each setting. For bounded delays, we propose a delay-aware zooming algorithm that retains the optimal performance of the delay-free setting up to an additional term that scales with the maximum delay $\tau_{\max}$. For unbounded delays, we propose a novel phased learning strategy that accumulates reliable feedback over carefully scheduled intervals, and establish a regret lower bound showing that our method is nearly optimal up to logarithmic factors. Finally, we present experimental results to demonstrate the efficiency of our algorithms under various delay scenarios.

Disclaimer

This report is **AI-GENERATED** using Large Language Models and WisPaper (a scholar search engine). It analyzes academic papers' tasks and contributions against retrieved prior work. While this system identifies **POTENTIAL** overlaps and novel directions, **ITS COVERAGE IS NOT EXHAUSTIVE AND JUDGMENTS ARE APPROXIMATE**. These results are intended to assist human reviewers and **SHOULD NOT** be relied upon as a definitive verdict on novelty.

Note that some papers exist in multiple, slightly different versions (e.g., with different titles or URLs). The system may retrieve several versions of the same underlying work. The current automated pipeline does not reliably align or distinguish these cases, so human reviewers will need to disambiguate them manually.

If you have any questions, please contact: mingzhang23@m.fudan.edu.cn

Core Task Landscape

This paper addresses: **Lipschitz Bandits with Stochastic Delayed Feedback**

A total of **6 papers** were analyzed and organized into a taxonomy with **7 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Continuous Action Space Bandits with Delays**
- **Multi-Agent and Game-Theoretic Learning with Delays**
- **General Sequential Decision Making with Delayed Feedback**

Complete Taxonomy Tree

- Lipschitz Bandits with Stochastic Delayed Feedback Survey Taxonomy
- Continuous Action Space Bandits with Delays
  - Lipschitz Bandits with Bounded Stochastic Delays ★ (1 papers)
  - [0] Lipschitz Bandits with Stochastic Delayed Feedback (Anon et al., 2026) View paper
  - Bandit Convex Optimization with Delayed Feedback (1 papers)
  - [2] Improved regret for bandit convex optimization with delayed feedback (Mingli Song, 2024) View paper
- Multi-Agent and Game-Theoretic Learning with Delays
  - Continuous Games with Delayed Rewards (1 papers)
  - [3] Gradient-free online learning in continuous games with delayed rewards (AmÃ©lie HÃ©liou, 2020) View paper
  - Cooperative Multi-Player Bandit Optimization (1 papers)
  - [5] Cooperative multi-player bandit optimization (Ilai Bistritz, 2020) View paper
- General Sequential Decision Making with Delayed Feedback
  - Reduction-Based Frameworks for Delayed Sequential Decisions (1 papers)
  - [1] A Reduction-based Framework for Sequential Decision Making with Delayed Feedback (Yang Yun-Chang, 2023) View paper
  - Distributed Stochastic Learning with Delays (1 papers)
  - [4] Distributed and Delayed Online Learning (Qiu, 2025) View paper
  - Network Analysis and Transport Methods in Bandit Problems (1 papers)
  - [6] Novel Methodologies in Network Analysis, Adversarial Learning, and Bandit Problems (Yi, 2025) View paper

Narrative

Core task: Lipschitz bandits with stochastic delayed feedback. This field addresses sequential decision-making problems where an agent selects actions from a continuous space under Lipschitz smoothness assumptions, but observes rewards only after random delays. The taxonomy organizes research into three main branches. The first, Continuous Action Space Bandits with Delays, focuses on algorithms that exploit smoothness or structure in infinite action sets while handling stochastic or adversarial delays; representative works include Lipschitz Bandits Delayed[0] and Bandit Convex Delayed[2], which develop regret bounds under bounded delay distributions. The second branch, Multi-Agent and Game-Theoretic Learning with Delays, examines settings where multiple learners interact, possibly in competitive or cooperative scenarios, and feedback arrives asynchronously; examples are Gradient-free Delayed Games[3], Cooperative Multi-player Bandit[5], and Distributed Delayed Learning[4]. The third branch, General Sequential Decision Making with Delayed Feedback, captures broader frameworks that apply to various bandit and reinforcement learning problems with delayed observations, such as Sequential Delayed Framework[1] and Network Adversarial Bandit[6].

A particularly active line of work contrasts single-agent continuous bandits with multi-agent or distributed settings. In the former, the main challenge is balancing exploration of a large action space with the uncertainty introduced by delayed rewards, often requiring careful discretization or tree-based search strategies. In the latter, coordination and communication constraints add complexity, as seen in Cooperative Multi-player Bandit[5] and Distributed Delayed Learning[4]. Lipschitz Bandits Delayed[0] sits squarely within the Continuous Action Space branch, emphasizing bounded stochastic delays and Lipschitz continuity to derive tight regret guarantees. Its focus on stochastic delay models distinguishes it from works like Bandit Convex Delayed[2], which may consider adversarial delays or convex loss functions, and from the game-theoretic emphasis of Gradient-free Delayed Games[3]. Overall, the field remains open regarding optimal delay-dependent regret rates and the interplay between smoothness assumptions and delay distributions.

Related Works in Same Category

No sibling papers were found in the same taxonomy leaf. A taxonomy-subtopic-level comparison will be produced instead.

Taxonomy-Level Summary

Both subtopics address bandit optimization problems with Lipschitz continuity assumptions under delayed feedback. The original leaf focuses specifically on stochastic delays that are bounded by a known maximum, while the sibling addresses convex loss minimization with delays. Both exclude settings without Lipschitz structure and defer unbounded delay scenarios to other methods.

Similarities: - Both assume Lipschitz continuity in the reward/loss structure - Both address the challenge of delayed feedback in bandit settings - Both exclude unbounded delay scenarios and non-Lipschitz structures - Both operate in sequential decision-making frameworks where observations arrive with temporal lag

Differences: - The original leaf focuses on general Lipschitz bandits with bounded stochastic delays, while the sibling specifically targets convex loss minimization - The original leaf explicitly requires a known maximum delay parameter, whereas the sibling's delay model is less specified - The sibling excludes game-theoretic multi-agent scenarios explicitly, while the original leaf does not mention this exclusion - The original leaf contrasts with convex optimization methods, suggesting it may handle non-convex objectives, while the sibling is restricted to convex losses

Suggested Search Directions: - Investigate whether bounded stochastic delay techniques from the original leaf can be adapted to convex bandit settings - Explore the boundary between general Lipschitz bandits and convex Lipschitz bandits under delays - Examine whether convex optimization methods can leverage known maximum delay bounds for improved regret

Sibling Subtopics

- **Bandit Convex Optimization with Delayed Feedback** (leaves: 1, papers: 1)
- Scope: Focuses on convex loss minimization in bandit settings where Lipschitz continuity applies and feedback is subject to delays.
- Exclude: Excludes non-convex objectives and game-theoretic multi-agent scenarios; see continuous games and general sequential decision making.

Contributions Analysis

Overall novelty summary. The paper addresses Lipschitz bandits with stochastic delayed feedback, proposing algorithms for both bounded and unbounded delay settings. According to the taxonomy, this work resides in the 'Lipschitz Bandits with Bounded Stochastic Delays' leaf, which currently contains only this paper as a sibling. This positioning indicates a relatively sparse research direction within the broader continuous action space bandits literature, suggesting the specific combination of Lipschitz continuity and stochastic delays has received limited prior attention in the examined literature.

The taxonomy reveals neighboring work in adjacent leaves: 'Bandit Convex Optimization with Delayed Feedback' addresses convex loss minimization under delays, while the 'Multi-Agent and Game-Theoretic Learning with Delays' branch explores strategic interactions with delayed rewards. The paper's focus on single-agent Lipschitz bandits distinguishes it from game-theoretic settings and from convex optimization approaches that may not exploit metric space structure. The taxonomy's scope notes explicitly exclude non-Lipschitz reward structures and adversarial delay models, clarifying that this work occupies a distinct methodological niche emphasizing stochastic delay distributions over continuous metric spaces.

Among the eight candidates examined through semantic search, none were found to refute the paper's three main contributions. The delay-aware zooming algorithm examined two candidates with zero refutations; the phased learning strategy for unbounded delays examined three candidates with zero refutations; and the instance-dependent lower bound examined three candidates with zero refutations. This limited search scope suggests that within the top-ranked semantically similar papers, no direct prior work addressing the same algorithmic framework for Lipschitz bandits with stochastic delays was identified, though the small candidate pool limits the strength of this signal.

Based on the constrained literature search of eight candidates and the sparse taxonomy leaf containing only this paper, the work appears to occupy a relatively unexplored intersection of Lipschitz continuity and stochastic delays. However, the limited search scope means potentially relevant work outside the top semantic matches may exist. The analysis covers immediate neighbors in the taxonomy but does not exhaustively survey all bandit or delay-related literature.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: Delay-aware zooming algorithm for bounded delays

Description: The authors extend the classic zooming algorithm to handle bounded stochastic delays in Lipschitz bandits. Their Delayed Zooming algorithm achieves a regret bound of Õ(T^(dz+1)/(dz+2) + τmax·T^(dz/(dz+2))), matching the optimal delay-free rate with an additive delay-dependent term.

This contribution was assessed against **2 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Efficient Algorithms for Lipschitz Bandits

URL: View paper

Brief Assessment

Efficient Lipschitz Bandits[7] focuses on memory-efficient algorithms for Lipschitz bandits without addressing delayed feedback, which is the core novelty of the original paper's contribution.

---

2. Distributed and Delayed Online Learning

URL: View paper

Brief Assessment

Distributed Delayed Learning[4] focuses on distributed online convex optimization and multi-armed bandits over communication networks, not on Lipschitz bandits with continuous action spaces or zooming algorithms for metric spaces.

Contribution 2: Phased learning algorithm for unbounded delays

Description: The authors introduce Delayed Lipschitz Phased Pruning (DLPP), a phased elimination-based algorithm that handles unbounded and potentially infinite delays. The algorithm achieves near-optimal regret with an additive dependence on delay distribution quantiles, supported by matching lower bounds.

This contribution was assessed against **3 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Bandit Learning Problems in Recommendation Systems: Self-Reinforcing User Preferences, Delayed Feedback, and Online Learning to Rank

URL: View paper

Brief Assessment

Recommendation Delayed Ranking[9] focuses on recommendation systems with delayed feedback in multi-armed bandits, not on Lipschitz bandits with continuous action spaces and phased elimination strategies as in the original paper.

---

2. Optimal Algorithm for Max-Min Fair Bandit

URL: View paper

Brief Assessment

Max-Min Fair Bandit[10] addresses multi-player multi-armed bandits with max-min fairness objectives, not Lipschitz bandits with delayed feedback. The phased elimination approach in Max-Min Fair Bandit[10] focuses on eliminating sub-optimal player-arm pairs to achieve fairness, while DLPP handles unbounded stochastic delays in continuous action spaces.

---

3. Delayed Adversarial Attacks on Stochastic Multi-Armed Bandits

URL: View paper

Brief Assessment

Delayed Adversarial Bandits[8] addresses adversarial attacks on multi-armed bandits with delayed feedback, not phased elimination algorithms for Lipschitz bandits with stochastic delays. The technical focus and problem settings are fundamentally different.

Contribution 3: Instance-dependent lower bound for delayed Lipschitz bandits

Description: The authors prove an instance-dependent regret lower bound for Lipschitz bandits with stochastic delays, demonstrating that their upper bounds are nearly tight. The lower bound shows that the regret must scale with both the zooming dimension and delay quantiles.

This contribution was assessed against **3 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Non-stationary Delayed Online Convex Optimization: From Full-information to Bandit Setting

URL: View paper

Brief Assessment

Non-stationary Delayed Convex[12] focuses on online convex optimization with delayed feedback in non-stationary environments, not on Lipschitz bandits with stochastic delays. The candidate paper does not address instance-dependent lower bounds for Lipschitz bandits.

---

2. de l'UniversitÃ© des Sciences et des Technologies de Lille

URL: View paper

Brief Assessment

Lille University[13] is a doctoral dissertation covering bandit theory, statistical learning, and reinforcement learning. While it addresses multi-armed bandits and includes lower bound analyses, it does not specifically address Lipschitz bandits with stochastic delays or provide instance-dependent lower bounds for this particular setting.

---

3. Recent advances in multiarmed bandits for sequential decision making

URL: View paper

Brief Assessment

Multiarmed Bandits Advances[11] is a survey paper covering recent advances in multi-armed bandits. The provided context fragments are too sparse to determine whether it discusses instance-dependent lower bounds for Lipschitz bandits with delays, and the fragments do not contain sufficient technical detail to refute the novelty of the original paper's contribution.

Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

References

- [0] Lipschitz Bandits with Stochastic Delayed Feedback View paper
- [1] A Reduction-based Framework for Sequential Decision Making with Delayed Feedback View paper
- [2] Improved regret for bandit convex optimization with delayed feedback View paper
- [3] Gradient-free online learning in continuous games with delayed rewards View paper
- [4] Distributed and Delayed Online Learning View paper
- [5] Cooperative multi-player bandit optimization View paper
- [6] Novel Methodologies in Network Analysis, Adversarial Learning, and Bandit Problems View paper
- [7] Efficient Algorithms for Lipschitz Bandits View paper
- [8] Delayed Adversarial Attacks on Stochastic Multi-Armed Bandits View paper
- [9] Bandit Learning Problems in Recommendation Systems: Self-Reinforcing User Preferences, Delayed Feedback, and Online Learning to Rank View paper
- [10] Optimal Algorithm for Max-Min Fair Bandit View paper
- [11] Recent advances in multiarmed bandits for sequential decision making View paper
- [12] Non-stationary Delayed Online Convex Optimization: From Full-information to Bandit Setting View paper
- [13] de l'UniversitÃ© des Sciences et des Technologies de Lille View paper