Novelty Assessment Report

Paper: Efficient algorithms for Incremental Metric Bipartite Matching

PDF URL: https://openreview.net/pdf?id=wnIanx0r0w

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2025-12-29

Abstract

The minimum-cost bipartite matching between two sets of points R and S in a metric space has a wide range of applications in machine learning, computer vision, and logistics. For instance, it can be used to estimate the 1-Wasserstein distance between continuous probability distributions and for efficiently matching requests to servers while minimizing cost. However, the computational cost of determining the minimum-cost matching for general metrics spaces, poses a significant challenge, particularly in dynamic settings where points arrive over time and each update requires re-executing the algorithm. In this paper, given a fixed set S, we describe a deterministic algorithm that maintains, after I additions to R, an $O(1/\delta^{0.631})$-approximate minimum-cost matching of cardinality I between sets R and S in any metric space, with an amortized insertion time of $\widetilde{O}(n^{1+\delta})$ for adding points in R. To the best of our knowledge, this is the first algorithm for incremental minimum-cost matching that applies to arbitrary metric spaces.

Interestingly, an important subroutine of our algorithm lends itself to efficient parallelization. We provide both a CPU implementation and a GPU implementation that leverages parallelism. Extensive experiments on both synthetic and real world datasets showcase that our algorithm either matches or outperforms all benchmarks in terms of speed while significantly improving upon the accuracy.

Core Task Landscape

This paper addresses: **Incremental Minimum-Cost Bipartite Matching in Metric Spaces**

A total of **26 papers** were analyzed and organized into a taxonomy with **11 categories**.

Taxonomy Overview

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

- **Online Matching with Immediate Irrevocable Decisions**
- **Online Matching with Delays**
- **Online Matching with Recourse**
- **Incremental and Dynamic Matching Algorithms**
- **Algorithmic Foundations and Related Problems**

Complete Taxonomy Tree

- Incremental Minimum-Cost Bipartite Matching in Metric Spaces Survey Taxonomy
- Online Matching with Immediate Irrevocable Decisions
  - General Metric Spaces
  - Deterministic Algorithms (2 papers)
    - [6] An optimal online algorithm for weighted bipartite matching and extensions to combinatorial auctions (Thomas Kesselheim, 2013) View paper
    - [17] A Robust and Optimal Online Algorithm for Minimum Metric Bipartite Matching (Raghvendra, 2016) View paper
  - Randomized Algorithms (4 papers)
    - [19] Randomized online algorithms for minimum metric bipartite matching (Adam Meyerson, 2006) View paper
    - [21] A Randomized O(log2k)-Competitive Algorithm for Metric Bipartite Matching (N. Bansal, 2014) View paper
    - [23] An O(log2k)-Competitive Algorithm for Metric Bipartite Matching (Bansal, 2007) View paper
    - [24] A randomized algorithm for the on-line weighted bipartite matching problem (BÃ©la Csaba, 2008) View paper
  - Line and Euclidean Metrics (2 papers)
  - [4] The power of greedy for online minimum cost matching on the line (Eric Balkanski, 2023) View paper
  - [7] Online minimum cost matching with recourse on the line (Megow, 2020) View paper
  - Stochastic and Beyond-Worst-Case Models (4 papers)
  - [1] Online Metric Matching: Beyond the Worst Case (Mingwei Yang, 2025) View paper
  - [10] Smoothed Analysis of Online Metric Matching with a Single Sample: Beyond Metric Distortion (Li Yingxi, 2025) View paper
  - [15] Stochastic online metric matching (Anupam Gupta, 2019) View paper
  - [18] Online minimum matching with uniform metric and random arrivals (Sharmila Duppala, 2021) View paper
  - Specialized Objective Functions (1 papers)
  - [3] Online Bottleneck Matching Problem with Two Heterogeneous Sensors in a Metric Space (Man Xiao, 2022) View paper
- Online Matching with Delays
  - General Metric Spaces with Delays (4 papers)
  - [8] Online matching: haste makes waste! (Yuval Emek, 2016) View paper
  - [13] Online Metric Matching with Delay (Deryckere, 2023) View paper
  - [20] A match in time saves nine: Deterministic online matching with delays (Bienkowski, 2017) View paper
  - [22] Polylogarithmic bounds on the competitiveness of min-cost (bipartite) perfect matching with delays (Azar, 2016) View paper

- Bipartite Matching with Delays (3 papers)
  - [9] Min-cost bipartite perfect matching with delays (Ashlagi, 2017) View paper
  - [11] Deterministic min-cost matching with delays (Y. Azar, 2020) View paper
  - [14] Minimum cost perfect matching with delays for two sources (Y. Emek, 2017) View paper
  - Non-Linear Delay Cost Functions (1 papers)
  - [16] Impatient online matching (Liu Xing-wu, 2018) View paper
- Online Matching with Recourse (2 papers)
  - [2] Online Metric Matching on the Line with Recourse (Nicole Megow, 2025) View paper
  - [12] Online minimum cost matching on the line with recourse (Nicole Megow, 2020) View paper
- Incremental and Dynamic Matching Algorithms ★ (1 papers)
  - [0] Efficient algorithms for Incremental Metric Bipartite Matching (Anon et al., 2026) View paper
- Algorithmic Foundations and Related Problems (3 papers)
  - [5] Online Duet between Metric Embeddings and Minimum-Weight Perfect Matchings (Bhore, 2023) View paper
  - [25] Input Sensitive Analysis of a Minimum Metric Bipartite Matching Algorithm (Krati Nayyar, 2017) View paper
  - [26] Empirical Analysis of Algorithms for the k-Server and Online Bipartite Matching Problems (Rutvij Sanjay Mahajan, 2018) View paper

Narrative

Core task: incremental minimum-cost bipartite matching in metric spaces. The field addresses how to match arriving requests to servers or resources in a metric space while minimizing total cost, under varying constraints on when and how matching decisions can be made or revised. The taxonomy organizes work into several main branches: Online Matching with Immediate Irrevocable Decisions focuses on classical models where each request must be matched instantly and permanently, often yielding competitive ratio bounds that depend on metric properties (e.g., Randomized Minimum Metric[19], Log2k Competitive Metric[21]). Online Matching with Delays permits deferring decisions for a bounded time, trading off waiting costs against better information (e.g., Bipartite Matching Delays[9], Polylogarithmic Delays[22]). Online Matching with Recourse allows limited changes to prior assignments, exploring how much improvement recourse can provide (e.g., Min Cost Recourse[7], Line with Recourse[2]). Incremental and Dynamic Matching Algorithms study settings where the input evolves more gradually or where offline-style reoptimization is feasible (e.g., Optimal Weighted Bipartite[6]). Finally, Algorithmic Foundations and Related Problems encompass broader theoretical tools, smoothed analysis, and connections to problems like k-server (e.g., Empirical k-Server[26], Smoothed Single Sample[10]).

A particularly active theme contrasts the power of delays versus recourse: works like Metric Matching Delay[13] and Deterministic Delays[11] show that even modest waiting can significantly improve competitive ratios, while studies such as Online Line Recourse[12] and Haste Makes Waste[8] explore how allowing a few reassignments can yield similar gains. Another line examines input-sensitive or beyond-worst-case guarantees (Beyond Worst Case[1], Input Sensitive Analysis[25]), moving away from purely adversarial models. The original paper, Incremental Metric Bipartite[0], sits within the Incremental and Dynamic Matching Algorithms branch, suggesting a focus on settings where the input arrives more predictably or where reoptimization is permitted. Compared to immediate-decision models like Randomized Minimum Metric[19] or delay-based approaches like Bipartite Matching Delays[9], Incremental Metric Bipartite[0] likely emphasizes maintaining or updating matchings as the instance grows, rather than committing irrevocably or waiting within strict time windows.

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

The original leaf focuses on incremental algorithms that maintain approximate matchings as points arrive, emphasizing efficient update operations without reconsidering past decisions. The sibling subtopics address complementary aspects: foundational techniques (embeddings, empirical methods) that support matching algorithm design, and online matching with recourse that explicitly allows modifying previous decisions at bounded cost. Together, these represent different layers of the matching problem space—foundational tools, incremental maintenance, and recourse-based flexibility.

Similarities: - All subtopics relate to bipartite matching in metric spaces with dynamic or online elements - All concern computational efficiency in handling sequential or evolving matching scenarios - All address approximate solutions rather than exact optimal matchings

Differences: - The original leaf focuses on incremental point arrivals with irrevocable matching decisions, while 'Online Matching with Recourse' explicitly allows bounded modifications to past decisions - The original leaf emphasizes maintenance algorithms and update operations, while 'Algorithmic Foundations' provides supporting techniques (embeddings, empirical analysis) rather than direct matching algorithms - 'Online Matching with Recourse' trades off solution quality against recourse cost, while the original leaf optimizes for update efficiency without reconsidering assignments - The original leaf excludes static models, while 'Algorithmic Foundations' may include static analysis techniques that inform dynamic algorithm design

Suggested Search Directions: - Investigate whether incremental algorithms can benefit from limited recourse (hybrid models) - Explore how metric embeddings from foundational work specifically enable efficient incremental updates - Examine the boundary between 'incremental maintenance' and 'recourse-based' approaches in terms of practical performance

Sibling Subtopics

- **Algorithmic Foundations and Related Problems** (leaves: 1, papers: 3)
- Scope: Foundational techniques including metric embeddings and empirical analysis supporting matching algorithm design.
- Exclude: Excludes direct matching algorithms; see online matching or incremental matching categories.
- **Online Matching with Recourse** (leaves: 1, papers: 2)
- Scope: Algorithms allowing modification of previous matching decisions at bounded recourse cost.
- Exclude: Excludes irrevocable or delay-based models; see immediate decisions or delays categories.

Contributions Analysis

Overall novelty summary. The paper presents a deterministic algorithm for maintaining an approximate minimum-cost bipartite matching in arbitrary metric spaces as points arrive incrementally. It occupies the 'Incremental and Dynamic Matching Algorithms' leaf of the taxonomy, which contains only this paper as a sibling. This isolation suggests the leaf represents a relatively sparse research direction within the broader field of online and dynamic matching, distinguishing itself from the more populated branches focused on immediate irrevocable decisions or delay-based models.

The taxonomy reveals that most prior work concentrates on online matching with immediate decisions (eight papers across deterministic and randomized approaches in general metrics) or matching with delays (eleven papers across general metrics, bipartite constraints, and non-linear costs). The original paper's position in a separate incremental branch indicates it diverges from these paradigms by allowing reoptimization as the input evolves, rather than committing irrevocably or trading off delay costs. Neighboring leaves like 'Online

Matching with Recourse' (two papers) share some conceptual overlap in permitting post-hoc adjustments, but the incremental setting appears to emphasize gradual input growth over bounded recourse budgets.

Among the three contributions analyzed, the literature search examined fourteen candidate papers total. The first contribution (constant-factor approximation with sublinear update time) examined two candidates with zero refutations. The second contribution (push-relabel framework for hierarchical metrics) also examined two candidates with zero refutations. The third contribution (parallelizable design with CPU/GPU implementations) examined ten candidates, again with zero refutations. These statistics suggest that within the limited search scope—focused on top semantic matches and citations—no prior work was identified that directly overlaps with the specific combination of incremental setting, general metric spaces, and the proposed algorithmic techniques.

Based on the top-fourteen semantic matches examined, the work appears to occupy a distinct niche: incremental maintenance of approximate matchings in arbitrary metrics with efficient updates. The absence of sibling papers in the same taxonomy leaf and zero refutations across all contributions indicate novelty relative to the surveyed literature. However, the limited search scope means the analysis does not cover the full landscape of dynamic graph algorithms or specialized metric structures that might contain related techniques outside the immediate matching literature.

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

Contribution 1: First constant-factor approximation algorithm for incremental metric bipartite matching with sublinear update time

Description: The authors introduce a deterministic algorithm that maintains an $O(1/\delta^{0.631})$-approximate minimum-cost matching in arbitrary metric spaces with amortized insertion time of $O(n^{1+\delta})$. This is the first algorithm for incremental minimum-cost matching that applies to arbitrary metric spaces while guaranteeing provably fast updates.

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. Maximum Matching on Trees in the Online Preemptive and the Incremental Graph Models

URL: View paper

Brief Assessment

Maximum Matching Trees[30] focuses on maximum matching problems on tree structures in online preemptive and incremental graph models, which is fundamentally different from the original paper's metric bipartite matching with cost minimization in arbitrary metric spaces.

---

2. Incremental -approximate dynamic matching in update time

URL: View paper

Brief Assessment

Incremental Approximate Dynamic[29] focuses on incremental $(1-\varepsilon)$-approximate bipartite matching in general graphs with edge insertions, not specifically on metric bipartite matching in arbitrary metric spaces. The candidate addresses a different problem setting (general bipartite graphs vs. metric spaces) and does not demonstrate prior work on constant-factor approximation for incremental metric bipartite matching.

Contribution 2: Push-relabel framework for maintaining 1-feasible matchings across hierarchical metric levels

Description: The authors develop a novel push-relabel framework that maintains partial 1-feasible matchings simultaneously across multiple hierarchical metric levels, replacing the augmenting path approach used in prior static algorithms. This enables efficient incremental updates without scanning the entire graph for each arriving request.

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. Rebalancing Modular Transit Systems: A Hierarchical Graph-Based Optimization Framework for Fleet Sizing and Routing

URL: View paper

Brief Assessment

Rebalancing Modular Transit[28] applies push-relabel to bipartite matching for fleet sizing in transit systems, not to maintaining 1-feasible matchings across hierarchical metric levels for incremental minimum-cost matching.

---

2. Multi-target tracking by non-linear motion patterns based on hierarchical network flows

URL: View paper

Brief Assessment

Multi-target Hierarchical[27] focuses on multi-target tracking using network flows in video sequences, not on incremental metric bipartite matching with hierarchical metrics. The mention of 'push relabel algorithm' appears only as a reference to existing flow algorithms, not as a novel framework for maintaining 1-feasible matchings across metric hierarchies.

Contribution 3: Parallelizable algorithm design with CPU and GPU implementations

Description: The algorithm supports parallel request processing where new requests can begin execution while earlier ones are still running. The authors provide efficient implementations on both CPU and GPU, with core subroutines that parallelize naturally, making the algorithm particularly effective for batched-insertion scenarios.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. THE GREAT DIVIDE: WORK PARTITIONING OF DATABASE JOINS

URL: View paper

Brief Assessment

Great Divide Joins[40] focuses on database join operations with CPU-GPU work partitioning for datasets exceeding GPU memory. The original paper addresses incremental metric bipartite matching with parallel request processing capabilities. These are fundamentally different problem domains and algorithmic approaches.

---

2. CGCGraph: Efficient CPU-GPU Co-execution for Concurrent Dynamic Graph Processing

URL: View paper

Brief Assessment

CGCGraph[37] focuses on concurrent dynamic graph processing with CPU-GPU co-execution for graph snapshots, not on incremental metric bipartite matching with batched insertions as in the original paper.

3. A Parallel CPU-GPU Framework for Batching Heuristic Operations in Depth-First Heuristic Search

URL: View paper

Brief Assessment

Parallel Depth-First[35] focuses on depth-first heuristic search algorithms (IDA*, BTS) with batched neural network heuristic evaluations, not on incremental bipartite matching with push-relabel frameworks as in the original paper.

4. GPU-Accelerated Batch-Dynamic Subgraph Matching

URL: View paper

Brief Assessment

Batch-Dynamic Subgraph[39] focuses on subgraph matching in dynamic graphs with GPU-based DFS search and warp-level parallelization, not on bipartite matching or metric spaces. The algorithmic techniques (DFS-based warp-centric search, work stealing) and application domain (graph pattern matching) differ fundamentally from the original paper's push-relabel framework for incremental metric bipartite matching.

5. SIGMo: High-Throughput Batched Subgraph Isomorphism on GPUs for Molecular Matching

URL: View paper

Brief Assessment

SIGMo Subgraph[34] focuses on subgraph isomorphism for molecular matching on GPUs, not incremental bipartite matching algorithms. The technical domains and problem formulations are fundamentally different.

6. Accelerated parallel hybrid GPU/CPU hash table queries with string keys

URL: View paper

Brief Assessment

Parallel Hash Table[31] focuses on hash table query operations with string keys, not bipartite matching algorithms. The parallelization techniques and application domains are fundamentally different.

7. Efficient algorithms to solve atom reconfiguration problems. III. The bird and batching algorithms and other parallel implementations on GPUs

URL: View paper

Brief Assessment

Atom Reconfiguration GPU[32] addresses atom reconfiguration in quantum computing trap arrays, not general reinforcement learning frameworks. The parallelization techniques are domain-specific to spatial atom displacement problems on grids, not batched request-server matching in metric spaces.

8. Stadium hashing: Scalable and flexible hashing on gpus

URL: View paper

Brief Assessment

Stadium Hashing[38] focuses on GPU-based hash table operations for key-value storage, not on incremental bipartite matching algorithms. The parallelization strategies serve fundamentally different computational problems.

9. Efficient Batched CPU/GPU Implementation of Orthogonal Matching Pursuit for Python

URL: View paper

Brief Assessment

Batched Orthogonal Matching[36] focuses on sparse signal recovery in compressed sensing using Orthogonal Matching Pursuit (OMP), not general reinforcement learning frameworks or bipartite matching problems. The parallelization techniques are domain-specific to OMP optimization.

10. ABCDPlace: Accelerated batch-based concurrent detailed placement on multithreaded CPUs and GPUs

URL: View paper

Brief Assessment

ABCDPlace[33] focuses on VLSI detailed placement optimization, not general bipartite matching or request-server assignment problems. The parallelization techniques target placement-specific operations (independent set matching, global swap, local reordering) rather than the incremental metric bipartite matching framework described in the original paper.

Appendix: Text Similarity Detection

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

References

- [0] Efficient algorithms for Incremental Metric Bipartite Matching View paper
- [1] Online Metric Matching: Beyond the Worst Case View paper
- [2] Online Metric Matching on the Line with Recourse View paper
- [3] Online Bottleneck Matching Problem with Two Heterogeneous Sensors in a Metric Space View paper
- [4] The power of greedy for online minimum cost matching on the line View paper
- [5] Online Duet between Metric Embeddings and Minimum-Weight Perfect Matchings View paper
- [6] An optimal online algorithm for weighted bipartite matching and extensions to combinatorial auctions View paper
- [7] Online minimum cost matching with recourse on the line View paper
- [8] Online matching: haste makes waste! View paper
- [9] Min-cost bipartite perfect matching with delays View paper
- [10] Smoothed Analysis of Online Metric Matching with a Single Sample: Beyond Metric Distortion View paper
- [11] Deterministic min-cost matching with delays View paper
- [12] Online minimum cost matching on the line with recourse View paper
- [13] Online Metric Matching with Delay View paper

- [14] Minimum cost perfect matching with delays for two sources View paper
- [15] Stochastic online metric matching View paper
- [16] Impatient online matching View paper
- [17] A Robust and Optimal Online Algorithm for Minimum Metric Bipartite Matching View paper
- [18] Online minimum matching with uniform metric and random arrivals View paper
- [19] Randomized online algorithms for minimum metric bipartite matching View paper
- [20] A match in time saves nine: Deterministic online matching with delays View paper
- [21] A Randomized O(log2k)-Competitive Algorithm for Metric Bipartite Matching View paper
- [22] Polylogarithmic bounds on the competitiveness of min-cost (bipartite) perfect matching with delays View paper
- [23] An O(log2k)-Competitive Algorithm for Metric Bipartite Matching View paper
- [24] A randomized algorithm for the on-line weighted bipartite matching problem View paper
- [25] Input Sensitive Analysis of a Minimum Metric Bipartite Matching Algorithm View paper
- [26] Empirical Analysis of Algorithms for the k-Server and Online Bipartite Matching Problems View paper
- [27] Multi-target tracking by non-linear motion patterns based on hierarchical network flows View paper
- [28] Rebalancing Modular Transit Systems: A Hierarchical Graph-Based Optimization Framework for Fleet Sizing and Routing View paper
- [29] Incremental -approximate dynamic matching in update time View paper
- [30] Maximum Matching on Trees in the Online Preemptive and the Incremental Graph Models View paper
- [31] Accelerated parallel hybrid GPU/CPU hash table queries with string keys View paper
- [32] Efficient algorithms to solve atom reconfiguration problems. III. The bird and batching algorithms and other parallel implementations on GPUs View paper
- [33] ABCDPlace: Accelerated batch-based concurrent detailed placement on multithreaded CPUs and GPUs View paper
- [34] SIGMo: High-Throughput Batched Subgraph Isomorphism on GPUs for Molecular Matching View paper
- [35] A Parallel CPU-GPU Framework for Batching Heuristic Operations in Depth-First Heuristic Search View paper
- [36] Efficient Batched CPU/GPU Implementation of Orthogonal Matching Pursuit for Python View paper
- [37] CGCGraph: Efficient CPU-GPU Co-execution for Concurrent Dynamic Graph Processing View paper
- [38] Stadium hashing: Scalable and flexible hashing on gpus View paper
- [39] GPU-Accelerated Batch-Dynamic Subgraph Matching View paper
- [40] THE GREAT DIVIDE: WORK PARTITIONING OF DATABASE JOINS View paper