

令和 5 年 6 月 9 日現在

機関番号：14401

研究種目：基盤研究(A)（一般）

研究期間：2018～2022

課題番号：18H04095

研究課題名（和文）ビッグデータ時代の多様な検索要求を満たす統一インデックス基盤の実現

研究課題名（英文）Implementation of Unified Indexing Mechanisms Satisfying Various Retrieval Requirements for Big Data Era

研究代表者

原 隆浩 (Hara, Takahiro)

大阪大学・大学院情報科学研究科・教授

研究者番号：20294043

交付決定額（研究期間全体）：（直接経費） 32,000,000円

研究成果の概要（和文）：IoTを始めとして多種多様なビッグデータが発生しており、データ集合の特徴抽出や指定条件に合致するデータの取得など、検索技術へのニーズも多様化している。既存研究では、多種多様な検索要求に対して、個別にデータ構造（インデックス）とオンライン検索アルゴリズムを構築しており、メモリ・ストレージのコストや構築・メンテナンスの手間が膨大となる。本研究では、この問題を解決するため、（課題1）多様な検索要求を満たす統一インデックス構造の構築、および（2）統一インデックス構造上のオンライン検索アルゴリズムの構築について研究開発を行った。本研究の成果は、多数の最高権威の論文誌や国際会議で公表された。

研究成果の学術的意義や社会的意義

従来研究では、個別のインデックス構造を構築し、そのうえでのみ効率的に動作する検索アルゴリズムを考案しているため、実システム上ではほとんど用いられていない。本研究は、多様な検索要求を満たすことができる統一インデックス構造とオンライン検索アルゴリズムを考案しており、実システム上の実現可能性が高い。さらに、従来技術よりも数十倍から数百倍高速に実行するアルゴリズムを考案しており、社会的な意義が極めて大きい。本研究の成果は、多数の最高権威の論文誌や国際会議で公表されており、学術的な意義も極めて大きい。

研究成果の概要（英文）：Recently, a variety of bigdata generated from different data sources such as IoT devices has caused diversified needs for advanced search including specific feature extraction and data retrieval based on specific conditions. Most existing studies individually design special index structures and online processing algorithms to meet the assumed search conditions, which results in significant increases in memory/storage costs and index maintenance costs. In this project, to solve this problem, we have conducted research on (i) constructing unified index structures to meet a variety of search needs, and (ii) designing online search algorithms on the constructed index structures. Our research achievements have been published in a number of top-ranked journals and conferences.

研究分野：データ工学

キーワード：ビッグデータ インデックス データ検索

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

IoT (モノのインターネット) や M2M (Machine-to-Machine) を始めとして多種多様なビッグデータが発生しており、データ分析の重要度は社会的・産業的に益々増大している。その中で、ビッグデータ解析の根幹を支える技術として、発生したデータ集合の特徴を抽出したり、ユーザが指定する条件に合致するデータを効率的に取得するなど、データ検索技術へのニーズの多様化が急速に進んでいる。実際、データベース分野などを中心にビッグデータに対する多様な検索技術が考案されているが、それぞれの技術に特化した独自のインデックス構造を用いているため、単独のシステム上で多様な検索技術を実現するのは、実装やメンテナンス、メモリやストレージのコスト的に現実的ではない。その結果、ほとんどの既存技術が実際のシステム上では使われていないのが現状である。研究代表者らは、このような状況に危機感を覚え、ビッグデータ検索技術をより現実的なものに発展させることを目的に、本研究を着想した。

2. 研究の目的

本研究では、上記の従来技術の問題を解決するために、多様な検索要求を満たすことができる統一的なインデックス基盤を確立することを目的とする。具体的には、複数の異なる検索要求に対して共通的に利用可能なインデックス構造を検討し、さらにその上で動作可能なオンライン検索アルゴリズムを個々の検索要求に対して検討する。構築したインデックスと検索アルゴリズムの有効性を、実データを用いた大規模な評価実験によって検証する。

3. 研究の方法

本研究では、上記の目的を達成するために、以下のように二つの研究課題を推進した。なお、これらの課題はそれぞれがある程度の独立性を持ちながらも、一方の開発内容が他方に影響を及ぼすため、相互に考慮しながら研究開発を進めた。研究全体の概要を図1に示す。

課題1: 多様な検索要求を満たす統一的なインデックス構造の構築

本研究では、多様な検索要求に対して、それぞれの検索の種類で効率化 (探索空間の枝刈り・絞込みなど) に用いることのできる数学的・理論的アイデアを整理もしくは新たに定義した。そして、複数の異なる検索要求において共通して利用できる効率化手段を検討した。ここで、異なる種類の検索は相互に全く異なる性質をもち、さらに単発的な検索やモニタリング (連続) 検索などが存在するため、一つのインデックス構造が全ての検索要求を満たすことは現実的に不可能である。そのため、実際には、異なる検索要求の中で、インデックスを統一利用可能な検索要求クラスにグループ化し、それぞれのクラスにおいて統一的なインデックスを検討する。

課題2: 統一的なインデックス構造上のオンライン検索アルゴリズムの構築

本研究では、課題1で構築されたインデックス構造上で、対象となる個々の検索要求に対するオンライン検索アルゴリズムを構築した。また、ビッグデータの分野では、新しい検索要求自体が新しい研究課題として提案されることが一般的であり、急速な勢いで検索要求の多様化が進んでいる。本研究では、新しい検索要求の提案とその処理アルゴリズムの考案も研究課題の一部として捉え、それらの新しい検索手法を統一的なインデックス構造上で実現することで、拡張性と汎用性を実証した。

本研究を進めるなかで、共通的なインデックス構造の一つとしてグラフインデックスが有効であることが判明したため、本研究の後半では、主にグラフインデックスを想定して、新たな検索の概念を考案するとともに、その高速なオンライン検索アルゴリズムを設計・実装した。

さらに、本研究のとりまとめとして、これまでに考案した検索技術を整理し、インデックスの共通化の観点から体系化を進めた。本研究で開発したインデックスやオンライン検索アルゴリズムは、可能な範囲で積極的に公開し、他の研究者や技術者による利用を促進した。

4. 研究成果

本章では、本研究の代表的な成果をいくつか紹介する。

課題1の成果:

まず、課題1に対して、類似検索 (範囲検索やk最近傍検索) を要する処理に関して、一つのインデックスで種々の検索操作を高速に実行できることを示した。具体的には、近接グラフを用いることで、(1) 類似検索、(2) 距離に基づくアウトライア検出、(3) Top-N 重み付き k クリーク探索を高速に実行できることを本研究期間で明らかにした。あるデータ集合における近接グラフとは、一つのデータをグラフにおける頂点とし、距離が近いデータ (頂点) 同士に辺を加えたデータ構造である。特に、先述の3つの問題に対しては、k最近傍グラフを用いることで十分な高速化が可能であることを示した。k最近傍グラフとは、ある頂点は自身のk最近傍と辺を持つデータ構造 (インデックス) である。図2はその一例を示している。

後に述べるように、k最近傍グラフはいくつかのプリミティブオペレータの高速化に役立つものの、k最近傍グラフの構築自体に無視できない計算時間が必要である。特に、メトリック空間

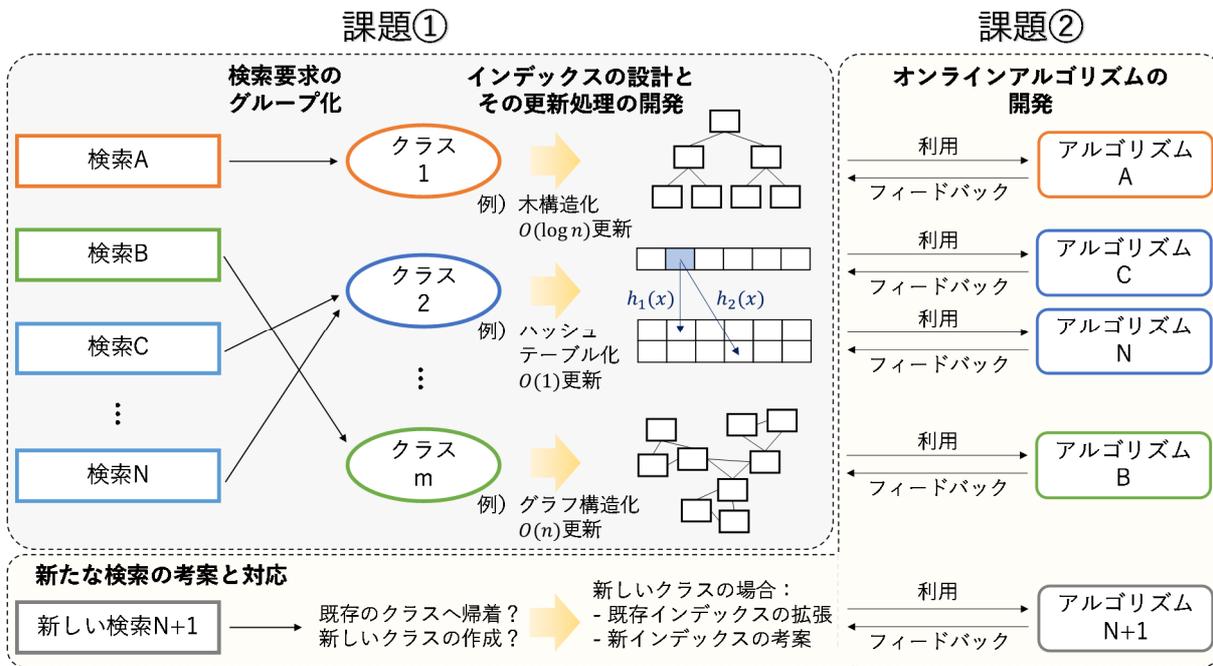


図 1：研究の全体像

において厳密な k 最近傍グラフを構築するには、データ数を n としたとき、 $O(n^2)$ 時間必要である。これはデータ数が大きい場合、例えば $n=1,000,000$ 、に何十時間以上の計算コストを要する。(低次元ユークリッド空間では、既存の木構造データ構造である R 木や kd 木を使えば高速に構築できるが、最悪時間は $O(n^2)$ のままである。) そのため、まずは k 最近傍グラフを高速に構築できるアルゴリズムを開発した。このアルゴリズムはメトリック空間を想定しており(メトリックを満たす任意の距離指標を利用可能) 低次元データだけでなく高次元データもサポートしている。ここで、任意のデータに対して厳密な k 最近傍グラフを構築するためには $O(n^2)$ 時間必要であることは変わらないため、高精度な「近似」 k 最近傍グラフを高速に構築することを目指した。

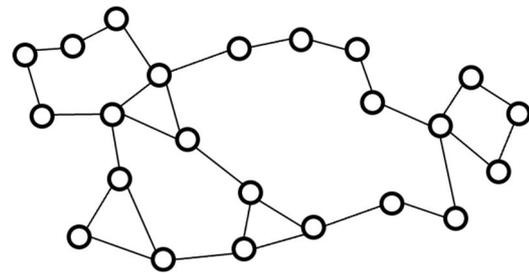


図 2： k 最近傍グラフ

本研究で開発したアルゴリズムは、既存の近似 k 最近傍グラフ構築アルゴリズムである NNDescent を拡張したものである。これはランダムなグラフを初期グラフとし、各頂点に対して自身の隣接点の隣接点との距離を計算し、より近い頂点に辺を貼り直すことを繰り返す(つまり、隣の隣を見て、距離が小さい頂点を見つける)。この「自身に似ているものの似ているものは、同様に自身に似ている」という考えによって反復を繰り返すことで、NNDescent は k 最近傍の精度を向上していく仕組みとなっている。ここで、このアルゴリズムを発表した論文ではこのアルゴリズムの計算量を解析していなかった。そこで我々はまずこのアルゴリズムの計算量を解析し、 $O(nk^2 \log k)$ であることを示した。つまり、このアルゴリズムはデータ数に比例する程度でしか計算時間は増加せず、厳密な k 最近傍グラフを計算するよりも (k が小さい場合は) 高速であることを保証している。また、実践的に高精度な近似 k 最近傍グラフを構築することも示されている。

しかし、このアルゴリズムは k 最近傍を更新しない無駄な頂点や辺へのアクセスが多いため、近似 k 最近傍の更新が起きない場合に無駄なアクセスを避けるアルゴリズム NNDescent+に拡張した。このアルゴリズムの計算量も $O(nk^2 \log k)$ のままであるものの、7つの実世界データセットを用いて実験し、NNDescent+は実践的に NNDescent に対して2倍以上の高速化を達成した。本研究は、SIGMOD2021 および The VLDB Journal で発表した内容の一部である。

課題 2 の成果：

次に、類似検索 (k 最近傍検索) 距離に基づくアウトライア検出、および Top-N 重み付き k クリーク探索に対する (近似) k 最近傍グラフを使ったアルゴリズムを開発した。

(1) 近似 k 最近傍検索の高速化：近年、埋め込みベクトルの普及によって、高速なベクトル探索の需要が急速に増しており、特に高次元空間における k 最近傍検索(クエリに最も距離に近い

k 個のデータを探索する問題)の高速化に高い注目が集まっている。高次元空間における k 最近傍検索はいくつかのアプローチが存在するが、近接グラフを使ったアプローチが最も実践的に高速であることが知られている。これは、図3のよう

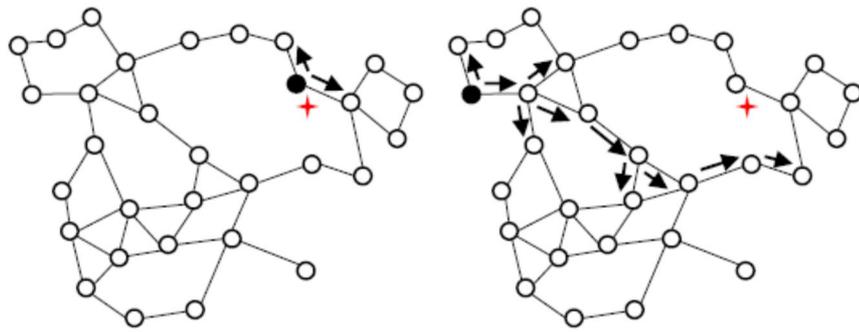


図3：近接グラフにおける k 最近傍検索

に、ある頂点から探索を開始し、クエリ(赤いマーカー)に最も近づく頂点を貪欲に辿るアプローチである。

近接グラフ上の近似 k 最近傍検索は様々なアルゴリズムが提案されているものの、その性能の理論的な解析はされてこなかった。そのため、どのような近接グラフを用いれば k 最近傍検索の時間や精度を理論的に保証できるかについて解析した。その結果、計算量の上限を保証できる近接グラフの存在は示したものの、そのための空間計算量が $O(n^2)$ となることも示し、理論的アルゴリズムをそのまま実装できないことを明らかにした。

また、このアプローチは初期の頂点によって性能が大きく影響することに着目した。図3の左側に示すように、初期頂点とクエリが近ければ、アクセスする頂点が少なく精度も高いことが期待できる。一方、図3の右に示すように初期頂点がクエリに遠い場合、アクセスする頂点も増え、局所解に陥って精度が低下する可能性が高まってしまう。そのため、なるべくクエリに近い頂点から探索を始める必要があるが、近い頂点を低コストで見つけるという課題も生じる。

上の課題に対して、近似 k 最近傍グラフと局所鋭敏性ハッシュ関数を併用することにより、実践的に高速かつ精度の向上を確率的に保証できることを示した。また、実世界のベクトルデータを用いて実験を行い、提案アルゴリズムが既存の代表的な近接グラフである NSG および HNSW よりも高速かつ高精度であることを確認した。本研究の内容は DEXA2021 で発表し、ベストペーパー賞を受賞した。

(2) 距離に基づくアウトライア検出の高速化：この問題は、距離の閾値 r が与えられたとき、自身から r 以内に k 個のデータが存在しないデータをアウトライアと定義し、そのようなアウトライアを全て検出する。例えば図4において、 $k=5$ のとき、 p はアウトライアである。このアウトライア検出は、教師なしアウトライア検出において、非常に高い精度でアウトライアを検出することが過去の研究で明らかとなっている。

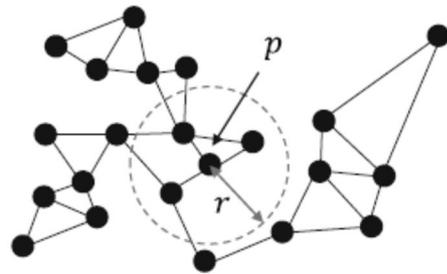


図4：近接グラフを用いたアウトライア判定

ここで、上記の定義から自身から距離 r 以内の辺を辿ることにより、自身がアウトライアかどうかを基本的に $O(k)$ 時間で判定できることが分かる。(既存アルゴリズムは $O(n)$ 時間必要であった。) このアイデアにより、アウトライアと判定できなかったデータに対してのみ線形スキャン等の $O(n)$ 時間アルゴリズムを用いてアウトライアかどうか計算すれば良い。また、近似 k 最近傍グラフに単調経路(距離が単調に減少する経路)を加えることによって、より多くのデータを $O(k)$ 時間でアウトライアではないと判定しやすくなることを示した。

また、k 最近傍までの距離が最も大きい N 個のデータをアウトライアとする定義も存在する。この定義においても、近似 k 最近傍グラフを用いることによって k 最近傍までの上界値を $O(k)$ 時間で計算することができ、正確な k 最近傍を計算する必要がないデータを枝刈りできることを示した。これにより、ほぼ $O(nN)$ 時間でアウトライアを判定できるアルゴリズムを設計した。本研究の内容は、データベース分野の最難関国際会議である SIGMOD02021 および最難関国際論文誌である The VLDB Journal にて発表した。

(3) Top-N 重み付き k クリーク探索の探索：近接グラフにおいて、辺の重みをその頂点間の距離とする。また、k クリークの重みをそれを構成する辺の重みの最大値とする。このとき、重みが最小の k クリークは、頂点間が互いに近いクリークであり、相互作用が生じやすいと考えられ

る。そのため、co-location pattern のマイニング等に応用可能である。

本問題を解く最も単純な方法は、全ての k クリークを列挙するものだが、これは NP 困難であり、非常に高い計算コストを要する。そこで、Top-N にならない k クリークおよび頂点を枝刈りし、無駄な k クリークの列挙を削減するアルゴリズムが求められる。これに対して、 k 最近傍グラフから、最も重みが小さい k クリークを推定して閾値を得ることで、大部分の k クリークを無視できるアルゴリズムを設計した。例えば $k=3$ のとき、図 5 において k 最近傍グラフから各頂点(データ)に対して k 最近傍によって構成されるクリークを簡単に ($O(k)$ 時間で) 列挙できる。これにより、図中の赤で表記しているような重みが小さい k クリークを小さい計算コストで求めることができる。このアイデアと三角不等式を用いて解に含まれない頂点を枝刈りし、厳密解を実践的に高速に計算する。

実世界データを用いた実験により、提案アルゴリズムは既存の列挙アルゴリズムよりも 100 倍以上高速に解を出力できることを確認した。本研究の内容は、DASFAA2022、IEEE Access、および IEEE BigData workshop にて発表した。

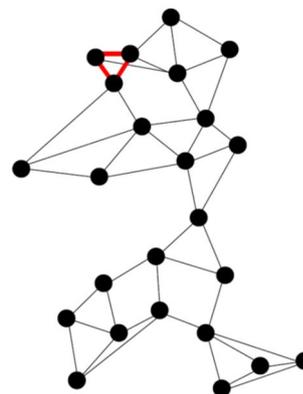


図 5: k 最近傍グラフから重みが小さい k クリークを推定

< 引用文献 >

- Dong, W., Moses, C., Li, K.: Efficient k -nearest neighbor graph construction for generic similarity measures. In: WWW, pp. 577-586 (2011).
- Amagata, D., Onizuka, M., Hara, T.: Fast and exact outlier detection in metric spaces: a proximity graph-based approach. In: SIGMOD, pp. 36-48 (2021).
- Amagata, D., Onizuka, M., Hara, T.: Fast, exact, and parallel-friendly outlier detection algorithms with proximity graph in metric spaces. The VLDB Journal, 31(4), pp. 36-48 (2021).
- Fu, C., Xiang, C., Wang, C., Cai, D.: Fast approximate nearest neighbor search with the navigating spreading-out graph. PVLDB 12(5), pp. 461-474 (2019).
- Malkov, Y.A., Yashunin, D.: Efficient and robust approximate nearest neighbor search using hierarchical navigable small world graphs. IEEE Trans. Pattern Anal. Mach. Intell. 42(4), pp. 824-836 (2020).
- Arai, Y., Amagata, D., Fujita, S., Hara, T.: LGTM: A Fast and Accurate k NN Search Algorithm in High-Dimensional Spaces, In: DEXA, pp. 220-231 (2021).
- Taniguchi, R., Amagata, D., Hara, T.: Efficient retrieval of top- k weighted spatial triangles, In: DASFAA, pp. 224-231 (2022).
- Taniguchi, R., Amagata, D., Hara, T.: Efficient retrieval of top- k weighted spatial triangles, IEEE Access 10, pp. 55298-55307 (2022).
- Taniguchi, R., Amagata, D., Hara, T.: Retrieving Top-N Weighted Spatial k -cliques, In: Big Data, pp. 4952-4962 (2022).

5. 主な発表論文等

〔雑誌論文〕 計10件（うち査読付論文 10件 / うち国際共著 0件 / うちオープンアクセス 3件）

1. 著者名 Amagata Daichi, Onizuka Makoto, Hara Takahiro	4. 巻 31(4)
2. 論文標題 Fast, Exact, and Parallel-friendly Outlier Detection Algorithms with Proximity Graph in Metric Spaces	5. 発行年 2022年
3. 雑誌名 The VLDB Journal	6. 最初と最後の頁 797-821
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/s00778-022-00729-1	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 加藤 慎也, 天方 大地, 原 隆浩	4. 巻 61(9)
2. 論文標題 多次元ストリーミング時系列データの効率的なモチーフモニタリングアルゴリズム	5. 発行年 2020年
3. 雑誌名 情報処理学会論文誌	6. 最初と最後の頁 1567-1576
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 加藤 慎也, 天方 大地, 西尾 俊哉, 原 隆浩	4. 巻 61(2)
2. 論文標題 ストリーミング時系列データに対するディスコードモニタリング	5. 発行年 2020年
3. 雑誌名 情報処理学会論文誌	6. 最初と最後の頁 510-519
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 鶴岡 翔平, 西尾 俊哉, 天方 大地, 原 隆浩	4. 巻 12(4)
2. 論文標題 Pub/Sub環境におけるkNNデータモニタリングの分散処理のためのクエリ割り当てアルゴリズム	5. 発行年 2019年
3. 雑誌名 情報処理学会論文誌データベース (IP SJ-TOD)	6. 最初と最後の頁 53-65
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 加藤 慎也, 天方 大地, 西尾 俊哉, 原 隆浩	4. 巻 60(7)
2. 論文標題 ストリーミング時系列データに対するモチーフモニタリング	5. 発行年 2019年
3. 雑誌名 情報処理学会論文誌	6. 最初と最後の頁 1260-1269
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Daichi Amagata, Takahiro Hara	4. 巻 -
2. 論文標題 Reverse Maximum Inner Product Search: Formulation, Algorithms, and Analysis	5. 発行年 2023年
3. 雑誌名 ACM Transactions on the Web	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1145/3587215	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Kohei Hirata, Daichi Amagata, Sumio Fujita, Takahiro Hara	4. 巻 11
2. 論文標題 Categorical Diversity-Aware Inner Product Search	5. 発行年 2023年
3. 雑誌名 IEEE Access	6. 最初と最後の頁 2586-2596
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ACCESS.2023.3234072	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Kohei Hirata, Daichi Amagata, Takahiro Hara	4. 巻 3
2. 論文標題 Cardinality Estimation in Inner Product Space	5. 発行年 2022年
3. 雑誌名 IEEE Open Journal of the Computer Society	6. 最初と最後の頁 208-216
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/OJCS.2022.3215206	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Ryosuke Taniguchi, Daichi Amagata, Takahiro Hara	4. 巻 10
2. 論文標題 Efficient Retrieval of Top-k Weighted Triangles on Static and Dynamic Spatial Data	5. 発行年 2022年
3. 雑誌名 IEEE Access	6. 最初と最後の頁 55298-55307
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ACCESS.2022.3177620	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Shunya Nishio, Daichi Amagata, Takahiro Hara	4. 巻 34
2. 論文標題 Lamps: Location-Aware Moving Top-k Pub/Sub	5. 発行年 2022年
3. 雑誌名 IEEE Transactions on Knowledge and Data Engineering	6. 最初と最後の頁 352 ~ 364
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/TKDE.2020.2979176	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

[学会発表] 計42件 (うち招待講演 0件 / うち国際学会 19件)

1. 発表者名 Yuchen Ji, Daichi Amagata, Yuya Sasaki, and Takahiro Hara
2. 発表標題 A Performance Study of One-dimensional Learned Cardinality Estimation
3. 学会等名 International Workshop on Design, Optimization, Languages and Analytical Processing of Big Data (DOLAP) (国際学会)
4. 発表年 2022年

1. 発表者名 Hayato Nakama, Daichi Amagata, and Takahiro Hara
2. 発表標題 Approximate Top-k Inner Product Join with a Proximity Graph
3. 学会等名 IEEE International Conference on Big Data (IEEE BigData) (国際学会)
4. 発表年 2021年

1 . 発表者名 Daichi Amagata, Shohei Tsuruoka, Yusuke Arai, and Takahiro Hara
2 . 発表標題 Feat-SKSJ: Fast and Exact Algorithm for Top-k Spatial-Keyword Similarity Join
3 . 学会等名 ACM SIGSPATIAL International Conference on Advances in Geographic Information System (SIGSPATIAL) (国際学会)
4 . 発表年 2021年

1 . 発表者名 Daichi Amagata and Takahiro Hara
2 . 発表標題 Reverse Maximum Inner Product Search: How to Efficiently Find Users Who Would Like to Buy My Item?
3 . 学会等名 ACM Recommender Systems Conference (RecSys) (国際学会)
4 . 発表年 2021年

1 . 発表者名 Yusuke Arai, Daichi Amagata, Sumio Fujita, and Takahiro Hara
2 . 発表標題 LGTM: A Fast and Accurate kNN Search Algorithm in High-dimensional Spaces
3 . 学会等名 International Conference on Database and Expert Systems Applications (DEXA) (国際学会)
4 . 発表年 2021年

1 . 発表者名 Daichi Amagata and Takahiro Hara
2 . 発表標題 Fast Density-Peaks Clustering: Multicore-based Parallelization Approach
3 . 学会等名 ACM SIGMOD International Conference on Management of Data (SIGMOD) (国際学会)
4 . 発表年 2021年

1. 発表者名 Daichi Amagata, Makoto Onizuka, and Takahiro Hara
2. 発表標題 Fast and Exact Outlier Detection in Metric Spaces: A Proximity Graph-based Approach
3. 学会等名 ACM SIGMOD International Conference on Management of Data (SIGMOD) (国際学会)
4. 発表年 2021年

1. 発表者名 谷口 凌亮, 天方 大地, 原 隆浩
2. 発表標題 空間データにおける重み付き三角形の効率的なTop-kモニタリング
3. 学会等名 第14回データ工学と情報マネジメントに関するフォーラム (DEIMフォーラム2022)
4. 発表年 2022年

1. 発表者名 平田 皓平, 天方 大地, 原 隆浩
2. 発表標題 高速かつ高精度な内積空間におけるカーディナリティ推定
3. 学会等名 第14回データ工学と情報マネジメントに関するフォーラム (DEIMフォーラム2022)
4. 発表年 2022年

1. 発表者名 谷口 凌亮, 天方 大地, 原 隆浩
2. 発表標題 空間データにおける重み付き三角形の効率的なTop-k検索アルゴリズム
3. 学会等名 FIT2021 第20回情報科学技術フォーラム
4. 発表年 2021年

1. 発表者名 平田 皓平, 天方 大地, 原 隆浩
2. 発表標題 内積空間におけるカーディナリティ推定
3. 学会等名 FIT2021 第20回情報科学技術フォーラム
4. 発表年 2021年

1. 発表者名 仲摩 隼人, 天方 大地, 原 隆浩
2. 発表標題 近似最大内積探索問題を利用した効率的なアウトライア検出アルゴリズム
3. 学会等名 FIT2021 第20回情報科学技術フォーラム
4. 発表年 2021年

1. 発表者名 Shohei Tsuruoka, Daichi Amagata, Shunya Nishio, Takahiro Hara
2. 発表標題 Distributed Spatial-Keyword kNN Monitoring for Location-aware Pub/Sub
3. 学会等名 International Conference on Advances in Geographic Information System (SIGSPATIAL) (国際学会)
4. 発表年 2020年

1. 発表者名 新井 悠介, 天方 大地, 原 隆浩, 藤田 澄
2. 発表標題 高次元空間における効率的な近似k Nearest Neighbor検索アルゴリズム
3. 学会等名 第13回データ工学と情報マネジメントに関するフォーラム (DEIMフォーラム2021)
4. 発表年 2021年

1. 発表者名 鶴岡 翔平, 天方 大地, 新井 悠介, 原 隆浩
2. 発表標題 位置・キーワードに基づくTop-k Similarity Joinの高速アルゴリズム
3. 学会等名 第13回データ工学と情報マネジメントに関するフォーラム (DEIMフォーラム2021)
4. 発表年 2021年

1. 発表者名 Daichi Amagata, Takahiro Hara
2. 発表標題 Correlation Set Discovery on Time-series Data
3. 学会等名 International Conference on Database and Expert Systems Applications (DEXA) (国際学会)
4. 発表年 2019年

1. 発表者名 Shinya Kato, Daichi Amagata, Shunya Nishio, Takahiro Hara
2. 発表標題 Discord Monitoring for Streaming Time-series
3. 学会等名 International Conference on Database and Expert Systems Applications (DEXA) (国際学会)
4. 発表年 2019年

1. 発表者名 天方 大地, 原 隆浩
2. 発表標題 メトリック空間における近接グラフを用いた高速アウトライアー検出
3. 学会等名 第12回データ工学と情報マネジメントに関するフォーラム (DEIMフォーラム2020)
4. 発表年 2020年

1. 発表者名 新井 悠介, 天方 大地, 原 隆浩, 藤田 澄男
2. 発表標題 近接グラフを用いたメトリック空間における効率的な近似範囲検索アルゴリズム
3. 学会等名 第12回データ工学と情報マネジメントに関するフォーラム (DEIMフォーラム2020)
4. 発表年 2020年

1. 発表者名 鶴岡 翔平, 西尾 俊哉, 天方 大地, 原 隆浩
2. 発表標題 位置・キーワードに基づくkNNデータモニタリングのための分散アルゴリズム
3. 学会等名 第12回データ工学と情報マネジメントに関するフォーラム (DEIMフォーラム2020)
4. 発表年 2020年

1. 発表者名 中谷 諒平, 天方 大地, 原 隆浩
2. 発表標題 メトリック空間におけるDensity Peaks Clusteringに対する木構造を用いた効率的なアルゴリズム
3. 学会等名 第12回データ工学と情報マネジメントに関するフォーラム (DEIMフォーラム2020)
4. 発表年 2020年

1. 発表者名 仲摩 隼人, 天方 大地, 原 隆浩
2. 発表標題 位置およびキーワードに基づく近似逆Top-k検索のためのバッチ処理アルゴリズム
3. 学会等名 第12回データ工学と情報マネジメントに関するフォーラム (DEIMフォーラム2020)
4. 発表年 2020年

1. 発表者名 西尾 俊哉, 天方 大地, 原 隆浩
2. 発表標題 位置およびキーワードに基づく近似逆Top-k検索のためのバッチ処理アルゴリズム
3. 学会等名 第12回データ工学と情報マネジメントに関するフォーラム (DEIMフォーラム2020)
4. 発表年 2020年

1. 発表者名 西尾 俊哉, 天方 大地, 原 隆浩
2. 発表標題 位置およびキーワードに基づく近似逆k最近傍検索
3. 学会等名 情報処理学会 マルチメディア通信と分散処理ワークショップ (DPSWS2019)
4. 発表年 2019年

1. 発表者名 加藤 慎也, 天方 大地, 原 隆浩
2. 発表標題 多次元ストリーミング時系列データの効率的なモチーフモニタリングアルゴリズム
3. 学会等名 情報処理学会 マルチメディア通信と分散処理ワークショップ (DPSWS2019)
4. 発表年 2019年

1. 発表者名 Daichi Amagata, Takahiro Hara
2. 発表標題 Identifying the Most Interactive Object in Spatial Databases
3. 学会等名 IEEE International Conference on Data Engineering (ICDE) (国際学会)
4. 発表年 2019年

1. 発表者名 Daichi Amagata, Takahiro Hara, Chuan Xiao
2. 発表標題 Dynamic Set kNN Self-Join
3. 学会等名 IEEE International Conference on Data Engineering (ICDE) (国際学会)
4. 発表年 2019年

1. 発表者名 加藤 慎也, 天方 大地, 西尾 俊哉, 原 隆浩
2. 発表標題 ストリーミング時系列データの効率的なディスコードモニタリングアルゴリズム
3. 学会等名 第11回データ工学と情報マネジメントに関するフォーラム (DEIMフォーラム2019)
4. 発表年 2019年

1. 発表者名 天方 大地, 原 隆浩
2. 発表標題 空間データベースにおけるインタラクティブオブジェクトの高速検索
3. 学会等名 第11回データ工学と情報マネジメントに関するフォーラム (DEIMフォーラム2019)
4. 発表年 2019年

1. 発表者名 西尾 俊哉, 天方 大地, 原 隆浩
2. 発表標題 ユーザの移動を考慮した位置・キーワードに基づくTop-kデータモニタリング
3. 学会等名 第11回データ工学と情報マネジメントに関するフォーラム (DEIMフォーラム2019)
4. 発表年 2019年

1. 発表者名 鶴岡 翔平, 西尾 俊哉, 天方 大地, 原 隆浩
2. 発表標題 位置・キーワードに基づくkNNデータモニタリングの分散処理のためのクエリ割り当てアルゴリズム
3. 学会等名 第11回データ工学と情報マネジメントに関するフォーラム (DEIMフォーラム2019)
4. 発表年 2019年

1. 発表者名 Shunya Nishio, Daichi Amagata, Takahiro Hara
2. 発表標題 Lamps: Location-Aware Moving Top-k Pub/Sub (Extended abstract)
3. 学会等名 IEEE International Conference on Data Engineering (ICDE) (国際学会)
4. 発表年 2023年

1. 発表者名 Daichi Amagata
2. 発表標題 Diversity Maximization in the Presence of Outliers
3. 学会等名 AAAI Conference on Artificial Intelligence (国際学会)
4. 発表年 2023年

1. 発表者名 Daichi Amagata
2. 発表標題 Scalable and Accurate Density-Peaks Clustering on Fully Dynamic Data
3. 学会等名 IEEE International Conference on Big Data (国際学会)
4. 発表年 2022年

1 . 発表者名 Ryosuke Taniguchi, Daichi Amagata, Takahiro Hara
2 . 発表標題 Retrieving Top-N Weighted Spatial k-cliques
3 . 学会等名 IEEE International Conference on Big Data (国際学会)
4 . 発表年 2022年

1 . 発表者名 Daichi Amagata, Yusuke Arai, Sumio Fujita, Takahiro Hara
2 . 発表標題 Learned k-NN Distance Estimation
3 . 学会等名 ACM SIGSPATIAL International Conference on Advances in Geographic Information System (SIGSPATIAL) (国際学会)
4 . 発表年 2022年

1 . 発表者名 Kohei Hirata, Daichi Amagata, Sumio Fujita, Takahiro Hara
2 . 発表標題 Solving Diversity-Aware Maximum Inner Product Search Efficiently and Effectively
3 . 学会等名 ACM Recommender Systems Conference (RecSys) (国際学会)
4 . 発表年 2022年

1 . 発表者名 Ryosuke Taniguchi, Daichi Amagata, Takahiro Hara
2 . 発表標題 Efficient Retrieval of Top-k Weighted Spatial Triangles
3 . 学会等名 International Conference on Database Systems for Advanced Applications (DASFAA) (国際学会)
4 . 発表年 2022年

1. 発表者名 ラモス アレハンドロ, 天方 大地, 原 隆浩
2. 発表標題 埋め込み順序従属性の検証アルゴリズム
3. 学会等名 第14回データ工学と情報マネジメントに関するフォーラム (DEIMフォーラム2023)
4. 発表年 2023年

1. 発表者名 青山 和禎, 天方 大地, 藤田 澄男, 原 隆浩
2. 発表標題 公平性を考慮した高速な内積探索
3. 学会等名 第14回データ工学と情報マネジメントに関するフォーラム (DEIMフォーラム2023)
4. 発表年 2023年

1. 発表者名 平田 皓平, 天方 大地, 原 隆浩, 藤田 澄男
2. 発表標題 多様性を考慮した最大内積探索
3. 学会等名 FIT2022 第21回情報科学技術フォーラム
4. 発表年 2022年

1. 発表者名 新井 悠介, 天方 大地, 藤田 澄男, 原 隆浩
2. 発表標題 高速・高精度なk最近傍距離推定
3. 学会等名 FIT2022 第21回情報科学技術フォーラム
4. 発表年 2022年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分担者	天方 大地 (Amagata Daichi) (40770649)	大阪大学・情報科学研究科・助教 (14401)	
研究 分担者	前川 卓也 (Maekawa Takuya) (50447025)	大阪大学・情報科学研究科・准教授 (14401)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------