

科学研究費助成事業 研究成果報告書

平成 30 年 6 月 15 日現在

機関番号：34310

研究種目：基盤研究(B) (一般)

研究期間：2014～2017

課題番号：26280115

研究課題名(和文)ビッグデータ活用アプリケーションのためのデータアクセスサービスに関する研究

研究課題名(英文)A Study on A Data Access Service for Applications utilizing Big Data Analytics

研究代表者

波多野 賢治 (Hatano, Kenji)

同志社大学・文化情報学部・教授

研究者番号：80314532

交付決定額(研究期間全体)：(直接経費) 12,100,000円

研究成果の概要(和文)：本研究では、ビッグデータを扱うアプリケーションの容易な開発をサポートするために、これまではアプリケーション側の創意工夫で実現していたデータアクセス部分を、一つのサービスとして捉え、そのサービスを介してデータアクセスするための基盤技術の開発を行った。特に、1) ビッグデータに対する効率的なアクセスを実現するためのデータストアサービス、2) 用途に応じたデータビュー導出サービス、3) ビッグデータの利活用を容易にするサービス指向データアクセス API の開発に力を入れ、各アプリケーションが API を介したデータアクセスが可能となったことで、データアクセス部の開発をせずに済むようになった。

研究成果の概要(英文)：In this study, we have developed a data access service for applications utilizing big data analytics. This is because the developers of the applications had implemented their own data access services using their inventive approach; however, this task was a factor in preventing the efficient software development. We implemented three types of base technologies for an efficient data access service from the applications utilizing big data analytics, and confirmed the effectiveness of our data access service.

研究分野：データ工学, メディア情報学

キーワード：サービス指向アーキテクチャ 社会サービス データアクセス基盤技術

1. 研究開始当初の背景

IoT デバイスの進化と社会への浸透を踏まえ、そうしたデバイスから得られるいわゆるビッグデータの利活用は、今や研究開発の分野に止まらず、実際の社会サービスやアプリケーションの一つとして認識されるようになった。

ビッグデータを利活用する際は、そうしたサービスやアプリケーションが、まずデータベースシステムにアクセスし、必要なデータだけを抽出する処理が発生するが、その際には、1) 効率的なデータアクセス機能を提供する旧来のデータベースシステム (RDBMS) が存在しない、2) 効率的データアクセス機能を有するデータベースシステム (NoSQL DBMS) は、データアクセス機能は貧弱で、社会サービスやアプリケーションの開発者が多大な労力を割く必要がある、といった問題が生じる。

2. 研究の目的

前述した問題を解決するためには、RDBMS の問合せ言語 SQL のような標準的なデータアクセス機能を NoSQL DBMS に持たせることになるが、NoSQL DBMS に用意されている問合せ言語はその DBMS に特化したものとなっており、それらの標準化は非常に難しい。

そこで本研究では、そのような問題を解決するために、本研究の三本柱として (1) ビッグデータに対する効率的なアクセスを実現するためのデータストアサービス、(2) 用途に応じたデータビュー導出サービス、(3) ビッグデータの利活用を容易にするサービス指向データアクセス API の開発、を設定し、それらを実現することで、ビッグデータ活用アプリケーションのためのデータアクセスサービスを構築する研究に着手することになった。

3. 研究の方法

研究の肝となる部分は、昨今、ビッグデータに対するデータアクセスに頻繁に用いられる NoSQL DBMS の使用を各種サービスやアプリケーションから容易にデータアクセスできるようにするための技術開発である。なぜなら、RDBMS の SQL のような標準的な問合せ言語があれば、サービス/アプリケーション開発者は、それほどの苦労することもなく、自身の開発に注力出来るためである。

乱立するの問合せ言語を SQL のように標準化することは非常に困難であると一般的に考えられ得るため、本研究では NoSQL DBMS 内に格納されているデータを開発者がハンドリングし易くする環境を整えるため、RDBMS におけるビューの概念を導入し、開発コストの増大を防ぐこととなった。

そこで本研究では、RDBMS と NoSQL DBMS をハイブリッド運用するための方法

論として、データのデータ指向格納のためのデータストア技術の開発、や要求・用途に応じたデータの分散ビュー導出技術の開発、を軸に要素技術の開発を行うこととし、それに並行して導出された分散ビューに対してさまざまなアプリケーションが共通のデータアクセス API を用いてデータアクセスする方法を確立する技術開発を行った。

また、開発したビッグデータ格納に対応したデータストアサービスとそのサービスから導出される分散ビュー、そしてその分散ビューに対してデータアクセス API を介してアクセスする一連の機構を、一つのサービス「データアクセスサービス」として展開し、それらが実運用に耐えうるか否かを、本研究グループで運用している簡易なスマートシティサービスを開発したデータアクセスサービス上で運用可能かを評価した。

4. 研究成果

研究の目的で明らかにしたように、本研究では主に三つの柱を中心に研究を行った。それぞれの研究テーマにおいて得られた研究成果は以下の通りである。

- (1) ビッグデータに対する効率的なアクセスを実現するためのデータストアサービス: このテーマでは、主にデータストアサービス内のデータベース自体で実行されるクエリ処理の効率化と、その分散環境における実現である。前者はクエリ処理時におけるメモリ内に読み込まれるデータをクエリ処理に関連するデータに限定することにより、メモリ使用量を 43.1%削減できたが、クエリ処理時間は 4.3%の削減に留まった。一方、分散環境上に実装したデータストアサービスでは、分散環境に特化したクエリ処理実行アルゴリズムを開発したことにより、メモリ使用量の削減は実現できなかったがクエリ処理時間を 25%削減できた。
- (2) 用途に応じたデータビュー導出サービス: 大規模なデータであるセンサデータを解析する場合は、(1) のアプローチのように分散環境における DBMS の効率化が必要である。このとき DBMS に必要な機能は多次元データに対するクエリ処理であるが、通常の DBMS ではその処理は非効率である。そこで、データの管理を RDB と NoSQL DBMS のハイブリッド運用することで、クエリの種類にも依るが処理時間を最大 97.2%削減することができた。一方、データ導出ビューにおいては完全な自動化を図ることは難しく、現時点ではデータアクセスパターンに応じて、ビューを半自動的に構築する手法の開発に留まった。
- (3) ビッグデータの利活用を容易にするサービス指向データアクセス API の開発: ビッグデータを活用したスマートサービスを容易に行えるように各種 API の実装

を行った。コンテキストウェアサービスや位置情報統合サービス、センシングデータ収集サービスなど各種固有の API の実装を始め、(1)、(2) で開発された各 DBMS にアクセスするための API も開発し、アプリケーションと DBMS 間を緩やかに繋ぐサービスが実現できた。これらは全て、サービス指向コンピューティングの概念に従った開発が行われ、API の共通化が図られており、これまで以上に低コストにアプリケーション開発が行われるための一助となった。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 10 件)

1. Kazuma Kusu, Izuru Kume, and Kenji Hatano, A Graph Partitioning Approach for Efficient Dependency Analysis using a Graph Database System, *International Journal on Advances in Networks and Services*, Vol. 12, No. 3&4, pp. 82-92, 2017. with review.
2. Hikaru Inomoto, Sachio Saiki, Masahide Nakamura, and Shinsuke Matsumoto, Design and Evaluation of Mission-Oriented Sensing Platform with Military Analogy, *International Journal of Pervasive Computing and Communications*, Vol. 13, No. 1, pp. 1-17, 2017. with review.
3. Long Niu, Sachio Saiki, Shinsuke Matsumoto, and Masahide Nakamura, WIF4InL: Web-Based Integration Framework for Indoor Location, *International Journal of Pervasive Computing and Communications*, Vol. 12, No. 1, pp. 49-65, 2016. with review.
4. Hiroki Takatsuka, Seiki Tokunaga, Sachio Saiki, Shinsuke Matsumoto, and Masahide Nakamura, KULOCS: Unified Locating Service for Efficient Development of Location-Based Applications, *International Journal of Pervasive Computing and Communication*, Vol. 12, No. 1, pp. 154-172, 2016. with review.
5. Seiki Tokunaga, Kazunari Tamamizu, Sachio Saiki, Masahide Nakamura, and Kiyoshi Yasuda, VirtualCareGiver: Personalized Smart Elderly Care, *International Journal of Software Innovation*, Vol. 5, No. 1, pp. 30-43, 2016. with review.
6. Seiki Tokunaga, Sachio Saiki, Shinsuke Matsumoto, and Masahide Nakamura, On Estimating Quality of Elderly Monitoring Service Based on Sensor Reliability, *Intelligent Automation & Soft Computing*, 2016. with review.
7. Ben Yan, Hua-Ping Yao, Masahide Nakamura, and Shinsuke Matsumoto, A Proposal for Deriving Timing Constraint Context on Multiple Sensor Web Services in Service Oriented Home Network, *International Journal of Smart Home*, Vol. 9, No. 8, pp. 151-162, 2015. with review.
8. Hiroki Takatsuka, Sachio Saiki, Shinsuke Matsumoto, and Masahide Nakamura, RuCAS: Rule-Based Framework for Managing Context-Aware Services with Distributed Web Services, *International Journal of Software Innovation*, Vol. 3, No. 3, pp. 57-68, 2015. with review.
9. Yuichi Watanabe, Shinsuke Matsumoto, Sachio Saiki, and Masahide Nakamura, Applying Log Data to Context-Awareness in Home Network System, *EAI Endorsed Transactions on Context-aware Systems and Applications*, Vol. 15, No. 3, pp. 1-10, 2015. with review.
10. Shintaro Yamamoto, Shinsuke Matsumoto, Sachio Saiki, and Masahide Nakamura, Design and Evaluation of Materialized View as a Service for Smart City Services with Large-Scale House Log, *IEICE Transactions on Information and Systems*, Vol. E97, No. 7, pp. 1709-1718, 2014. with review.

[学会発表] (計 58 件)

1. 若月 駿亮, 櫻 惇志, 宮崎 純, GPU 上で辞書を利用した N-gram クエリ尤度の効率的な事前計算, 第 10 回データ工学と情報マネジメントに関するフォーラム, 2018.
2. 張 涵, 渡 佑也, 櫻 惇志, 宮崎 純, 中村 匡秀, 分散環境における多次元データに対する集約演算結果の推定とその評価, 第 10 回データ工学と情報マネジメントに関するフォーラム, 2018.
3. 柳本 晟熙, 櫻 惇志, 宮崎 純, GPU 上の MapReduce による大規模データ処理の最適な分割粒度の動的推定, 第 10 回データ工学と情報マネジメントに関するフォーラム, 2018.
4. Takamitsu Shioi, Kenji Hatano, and Haruo Yokota, SCMAT: A Mechanism Presuming SCMs to Efficiently Enable both OLAP and OLTP, *IEEE BigData Congress 2017*, 2017.

5. Toshiaki Wakatsuki, Atsushi Keyaki, and Jun Miyazaki, A Case for Term Weighting using a Dictionary on GPUs, DEXA 2017, 2017.
6. Takuhiro Kagawa, Sachio Saiki, and Masahide Nakamura, Visualizing and Analyzing Street Crimes Using Personalized Security Information Service PRISM, iiWAS 2017, 2017.
7. 塩井 隆円, 横田 治夫, 波多野 賢治, SCM の利用を想定した OLAP 高速化のための列指向バッファ更新用インデックスの提案, 第 9 回データ工学と情報マネジメントに関するフォーラム, 2017.
8. 楠 和馬, 久米 出, 波多野 賢治, プログラムの動的解析効率化のための参照頻度を考慮したグラフ属性分割格納法, 第 9 回データ工学と情報マネジメントに関するフォーラム, 2017.
9. 柳本 晟熙, 櫻 惇志, 宮崎 純, GPU を用いた大規模な文書に対する高精度検索のための高速な重み付け計算, 第 9 回データ工学と情報マネジメントに関するフォーラム, 2017.
10. 張 涵, 渡 佑也, 櫻 惇志, 宮崎 純, ヒストグラムとカーネル密度推定を組み合わせた集約演算結果の推定, 第 9 回データ工学と情報マネジメントに関するフォーラム, 2017.
11. 植田 聖司, 櫻 惇志, 宮崎 純, プレイリスト生成における遷移確率を用いたスコアリング手法の提案, 第 9 回データ工学と情報マネジメントに関するフォーラム, 2017.
12. 渡 佑也, 櫻 惇志, 宮崎 純, 多次元データに対する集約演算の効率化手法におけるデータ挿入スループットの向上, 第 9 回データ工学と情報マネジメントに関するフォーラム, 2017.
13. 若月 駿亮, 櫻 惇志, 宮崎 純, GPU 向け簡潔データ構造による辞書を利用した効率的な語の重み計算, 第 9 回データ工学と情報マネジメントに関するフォーラム, 2017.
14. 佐々木 夢, 櫻 惇志, 宮崎 純, 検索対象データの事前インデックスを必要としない Top-k 検索アルゴリズムの提案と評価, 第 9 回データ工学と情報マネジメントに関するフォーラム, 2017.
15. 鈴木 崇弘, 櫻 惇志, 宮崎 純, Copula を用いたユーザプロファイリング手法の提案, 第 9 回データ工学と情報マネジメントに関するフォーラム, 2017.
16. 楠 和馬, 久米 出, 波多野 賢治, アクセスパターンを利用した同一節点へのグラフ走査回数削減法, 平成 29 年度情報処理学会関西支部支部大会, 2017.
17. 柳本 晟熙, 櫻 惇志, 宮崎 純, GPU 上の MapReduce による大規模データの処理におけるソートアルゴリズムの影響と評価, 情報処理学会データベースシステム, 2017.
18. 渡 佑也, 櫻 惇志, 宮崎 純, 中村 匡秀, RDB と KVS を相互に活用した大規模多次元データに対する集約演算の効率化, 情報処理学会データベースシステム, 2017.
19. Takamitsu Shioi and Kenji Hatano, Rule- and Cost-Based Optimization of OLAP Workloads on Distributed RDBMS with Column-Oriented Storage Function, FICloud 2016, 2016.
20. Yume Sasaki, Takuya Komatsuda, Atsushi Keyaki, and Jun Miyazaki, A New Readability Measure for Web Documents and its Evaluation on an Effective Web Search Engine, iiWAS 2016, 2016.
21. Seiki Tokunaga, Hiroyasu Horiuchi, Sachio Saiki, Shinsuke Matsumoto, Masahide Nakamura, and Kiyoshi Yasuda, Implementation and Evaluation of Interactive Memory-Aid Agent Service for People with Dementia, HCI 2016, 2016.
22. Kazunari Tamamizu, Seiki Tokunaga, Sachio Saiki, Shinsuke Matsumoto, Masahide Nakamura, and Kiyoshi Yasuda, Towards Person-Centered Anomaly Detection and Support System for Home Dementia Care, HCI 2016, 2016.
23. Seiki Tokunaga, Hiroyasu Horiuchi, Hiroki Takatsuka, Sachio Saiki, Shinsuke Matsumoto, Masahide Nakamura, and Kiyoshi Yasuda, Towards Personalized and Context-Aware Reminder Service for People with Dementia, IEEE WCCI 2016, 2016.
24. Seiji Sakakibara, Sachio Saiki, Masahide Nakamura, and Shinsuke Matsumoto, Indoor Environment Sensing Service in Smart City Using Autonomous Sensor Box, ICIS 2016, 2016.
25. Hikaru Inomoto, Sachio Saiki, Masahide Nakamura, and Shinsuke Matsumoto, Mission-Oriented Large-Scale Environment Sensing Based on Analogy of Military System, iiWAS 2016, 2016.
26. Kazuma Kusu, Izuru Kume, and Kenji Hatano, A Trace Partitioning Approach for Efficient Trace Analysis, ACIT 2016, 2016.
27. 玉水 一柔, 徳永 清輝, 堀内 大祥, 榎本 真佑, 佐伯 幸郎, 中村 匡秀, 安田 清, 環境センシングに基づく在宅認知症者

- のための異常検知・対応サービスの検討, 電子情報通信学会知的環境とセンサネットワーク研究会, 2016.
28. 堀内 大祥, 徳永 清輝, 榎本 真佑, 佐伯 幸郎, 中村 匡秀, 安田 清, 認知症高齢者の記憶補助のための対話型エージェントシステムの評価, 電子情報通信学会知的環境とセンサネットワーク研究会, 2016.
 29. 榎原 誠司, 佐伯 幸郎, 榎本 真佑, 中村 匡秀, 自律センサボックスを活用したスマートシティ屋内環境センシングサービスの開発, 電子情報通信学会ライフインテリジェンスとオフィス情報システム研究会, 2016.
 30. 楠 和馬, 久米 出, 波多野 賢治, グラフ処理を用いたトレース解析システムの構築, 2016 年電子情報通信学会総合大会, 2016.
 31. 塩井 隆円, 波多野 賢治, 列指向型データ格納 RDBMS を利用した分散 OLAP の最適化, 第 8 回データ工学と情報マネジメントに関するフォーラム, 2016.
 32. 渡 佑也, 櫻 惇志, 宮崎 純, RDB と KVS を相互に利用した多次元データに対する集約演算の効率化, 第 8 回データ工学と情報マネジメントに関するフォーラム, 2016.
 33. 佐々木 夢, 小松田 卓也, 櫻 惇志, 宮崎 純, 文書の可読性を考慮した Web 検索に関する一考察, 第 8 回データ工学と情報マネジメントに関するフォーラム, 2016.
 34. 若月 駿亮, 櫻 惇志, 宮崎 純, 効率的なテキスト処理を目指した簡潔データ構造を用いたトライ木の GPU 上での実装, 第 8 回データ工学と情報マネジメントに関するフォーラム, 2016.
 35. 岸田 脩平, 櫻 惇志, 宮崎 純, スカイライン演算を用いたユーザ思考を考慮した情報推薦のランキング手法の精度改善, 第 8 回データ工学と情報マネジメントに関するフォーラム, 2016.
 36. 植田 聖司, 櫻 惇志, 宮崎 純, プレイリスト生成におけるグラフモデルを用いたスコアリング手法の提案, 第 8 回データ工学と情報マネジメントに関するフォーラム, 2016.
 37. 塩井 隆円, 波多野 賢治, 列指向型データストア機能を有する RDBMS の問合せ処理コスト計算法, 第 15 回情報科学技術フォーラム, 2016.
 38. 楠 和馬, 久米 出, 波多野 賢治, デバッグ支援を目的とした大域的動的依存性解析の効率化, 第 15 回情報科学技術フォーラム, 2016.
 39. Seiki Tokunaga, Hiroyasu Horiuchi, Sachio Saiki, Shinsuke Matsumoto, Masahide Nakamura, and Kiyoshi Yasuda, Design of Memory Aids Agent Service Based Location and Time for People with Dementia, 2015 Alzheimer's Disease Congress, 2015.
 40. Masahide Nakamura and Lydie du Bousquet, Constructing Execution and Life-Cycle Models for Smart City Services with Self-Aware IoT, ICAC 2015, 2015.
 41. Arisa Hayashi, Sachio Saiki, Seiki Tokunaga, Shinsuke Matsumoto, and Masahide Nakamura, Formulating Device-Independent Pass-By Rendezvous, APSITT 2015, 2015.
 42. Hikaru Inomoto, Sachio Saiki, Satoshi Takatori, Seiki Tokunaga, Shinsuke Matsumoto, and Masahide Nakamura, Implementation of Infrared as a Service in Cloud-Based HNS, APSITT 2015, 2015.
 43. Takamitsu Shioi and Kenji Hatano, Query Processing Optimization using Two Types of Storage Devices, 19th IDEAS 2015, 2015.
 44. Takamitsu Shioi and Kenji Hatano, Query Processing Optimization using Disk-based Row-store and Column-store, iiWAS 2015, 2015.
 45. Long Niu, Sachio Saiki, Shinsuke Matsumoto, and Masahide Nakamura, Implementation and Evaluation of Cloud-Based Integration Framework for Indoor Location, iiWAS 2015, 2015.
 46. Hiroki Takatsuka, Seiki Tokunaga, Sachio Saiki, Shinsuke Matsumoto, and Masahide Nakamura, Integrating Heterogeneous Locating Services for Efficient Development of Location-Based Services, iiWAS 2015, 2015.
 47. 塩井 隆円, 波多野 賢治, 列指向型データ格納 RDBMS の問合せ処理最適化の一検討, 2015 年電子情報通信学会総合大会, 2015.
 48. 今井 良一, 櫻 惇志, 宮崎 純, 中村 匡秀, マテリアライズドビューを利用した分散キーバリューストアへのアクセスの効率化, 第 7 回データ工学と情報マネジメントに関するフォーラム, 2015.
 49. 植田 聖司, 櫻 惇志, 宮崎 純, スカイライン演算を用いたユーザ嗜好を考慮した情報推薦のランキング手法の提案, 第 7 回データ工学と情報マネジメントに関するフォーラム, 2015.
 50. 小松田 卓也, 清水 伸幸, 田島 玲, 櫻 惇志, 宮崎 純, Copula を用いたスコア統合手法とその有効性の検証, 第 7 回データ工学と情報マネジメントに関するフォーラム, 2015.
 51. 井元 滉, 榎本 真佑, 佐伯 幸郎, 中村 匡秀, 軍隊ヒエラルキーに基づくスケー

ラブルな IoT センシング基盤の検討, 電子情報通信学会ライフインテリジェンスとオフィス情報システム研究会, 2015.

52. 榎本 真佑, 佐伯 幸郎, 中村 匡秀, HTML/CSS/JavaScript に対するサイズ最適化リテラシの醸成に向けて, ソフトウェアエンジニアリングシンポジウム 2015, 2015.
53. 塩井 隆円, 波多野 賢治, 列指向型データ格納可能な RDBMS の問合せ内述語に基づく問合せ処理最適化, 第 14 回情報科学技術フォーラム, 2015.
54. 林 亜梨沙, 榎本 真佑, 佐伯 幸郎, 中村 匡秀, すれちがいフレームワークにおける標準データ生成・蓄積機構の検討, 電子情報通信学会情報ネットワーク研究会, 2015.
55. 鳥羽 祐輔, 榎本 真佑, 佐伯 幸郎, 中村 匡秀, 内野 智仁, 横山 知弘, 武林 靖浩, 聴覚障害者支援のための発話可視化モダルの開発支援, 電子情報通信学会福祉情報工学研究会, 2015.
56. Hiroki Takatsuka, Sachio Saiki, Shinsuke Matsumoto, and Masahide Nakamura, Design and Implementation of Rule-Based Framework for Context-Aware Services with Web Services, iiWAS 2014, 2014.
57. Yuichi Watanabe, Tetsuya Masuda, Shinsuke Matsumoto, Sachio Saiki, and Masahide Nakamura, Exploiting House Log of Home Network System to Derive Contexts with Past Situations, ICCASA 2014, 2014.
58. Tetsuya Masuda, Yuichi Watanabe, Shinsuke Matsumoto, Sachio Saiki, and Masahide Nakamura, Exploiting Contexts with House Log for Value-Added Smart TV Services, IEEE GCCE 2014, 2014.

〔図書〕(計 0 件)

〔産業財産権〕

出願状況(計 0 件)

名称 :
発明者 :
権利者 :
種類 :
番号 :
出願年月日 :
国内外の別 :

取得状況(計 0 件)

名称 :

発明者 :
権利者 :
種類 :
番号 :
取得年月日 :
国内外の別 :

6 . 研究組織

(1) 研究代表者

波多野 賢治 (HATANO, Kenji)
同志社大学・文化情報学部・教授
研究者番号: 80314532

(2) 研究分担者

宮崎 純 (MIYAZAKI, Jun)
東京工業大学・情報理工学院・教授
研究者番号 : 40293394

中村 匡秀 (NAKAMURA, Masahide)
神戸大学・大学院システム情報学研究科・准教授
研究者番号: 30324859

(3) 連携研究者 なし

(4) 研究協力者

SRIVASTAVA, Divesh
Head, Database Research Department,
AT&T Labs Research

CHAKRAVARTHY, Sharma
Professor, Computer Science &
Engineering Department,
The University of Texas at Arlington

REIFF-MARGANIEC, Stephan
Senior Lecturer, Department of
Informatics, University of Leicester

櫻 惇志 (KEYAKI, Atsushi)
東京工業大学・情報理工学院・助教
研究者番号: 00733958

佐伯 幸郎 (SAIKI, Sachio)
神戸大学・先端融合研究環・特命助教
研究者番号: 40549408

榎本 真佑 (MATSUMOTO, Shinsuke)
大阪大学・大学院情報科学研究科・助教
研究者番号: 90583948