

科学研究費助成事業 研究成果報告書

令和 6 年 5 月 31 日現在

機関番号：33917

研究種目：基盤研究(C) (一般)

研究期間：2017～2023

課題番号：17K00432

研究課題名(和文) プライバシー保護データマイニングを用いたコンテキストベース位置情報サービス

研究課題名(英文) Context-base Geographic Information Service using Privacy Preservation Data Mining

研究代表者

河野 浩之 (Kawano, Hiroyuki)

南山大学・理工学部・教授

研究者番号：70224813

交付決定額(研究期間全体)：(直接経費) 3,300,000円

研究成果の概要(和文)：「位置情報・環境情報・操作行動履歴・プロフィール等」のデータに基づくコンテキストベースの情報サービスが提供されつつあるが、個人情報を含むデータ利活用に関する議論が重要である。

本研究では、プライバシー保護データマイニング(Privacy Preserving Data Mining)技術を用いて、位置情報サービスプロバイダにおけるデータセキュリティを担保した上で、より高精度なコンテキストベースサービスを提供するデータマイニングアルゴリズムの利用を提案した。例えば、プライバシー保護環境下の経路探索や経路推薦を、提案アルゴリズムを用いるソーシャルフィルタリングなどの応用が考えられる。

研究成果の学術的意義や社会的意義

本研究は、位置情報サービスプロバイダにおけるデータセキュリティを担保した上で、より高精度なコンテキストベースサービスを提供するプライバシー保護データマイニング(Privacy Preserving Data Mining)を提案した。秘密計算技術の応用を検討したことが、学術的意義である。

また、秘密計算技術の社会的実装も進みつつある。かつ、EU一般データ保護規則(General Data Protection Regulation: GDPR)や人工知能に関する包括的規制などが議論されている。これらの観点から、本研究のプライバシー保護データマイニング技術には、社会的意義があると考えている。

研究成果の概要(英文)：Context-based location information services based on "location information, environmental information, operation history, profiles, etc." are growing up, but it is important to discuss the privacy of data including personal information.

In this study, we propose the use of a data mining algorithm that provides more accurate context-based location services while ensuring data privacy and security for location-based service providers by using "Privacy Preserving Data Mining (PRDM)" technology.

We have proposed the use of a data mining algorithm that provides more accurate context-based services while ensuring data security for location-based service providers. For example, our proposed algorithm can be used for social filtering of route discovery and route recommendation in privacy preserving environments.

研究分野：情報システム

キーワード：プライバシー保護データマイニング 位置情報サービス コンテキストベースサービス データセキュリティ 交通データウェアハウス ビッグデータ

※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

様式 C-19、F-19-1 (共通)

1. 研究開始当初の背景

(1) 研究開始にあたって、「プライバシー保護データマイニングを用いた交通データウェアハウス解析」(基盤研究 C, 研究代表者, H24-27)において、基礎的なプライバシー保護機能をもつクラウドのプロトタイプシステム構築を行った。研究開始時点の構築システムは、計算性能に課題を残しているが、HElib(完全準同型暗号の実装プログラム)を用いたクラウドシステムのプロトタイプシステムの実現可能性を検証した。

(2) 加えて、コンテキストベース位置情報サービスに用いる行動履歴データ収集技術として、「プローブ技術を援用したデータフュージョン理論による総合的交行動調査の高度化」(基盤研究(A), 研究分担者, H21-24)におけるプローブカーデータの収集技術を活用し、スマートフォン等を利用したデータ収集を検討した。

2. 研究の目的

(1) 位置情報サービスが普及するにつれ、各種デバイスにセンサーを搭載することで、より正確かつ詳細な「位置情報・環境情報・操作行動履歴・プロフィール等」のデータに基づくコンテキストベースの情報サービスが可能になりつつある。本研究では、プライバシー保護データマイニング(Privacy Preserving Data Mining)技術を用いて、位置情報サービスプロバイダにおけるデータセキュリティを担保した上で、より高精度なコンテキストベースサービスを提供するデータマイニングアルゴリズムの研究を目的とする。

(2) 具体的には、携帯デバイスによる行動履歴データ、都市高速道路において蓄積されている交通データウェアハウス、プローブ機能を備えた車両から得られる複数データベースを用いて、アルゴリズムの提案・実装ならびに評価の実施を目指す。

3. 研究の方法

(1) ビッグデータやIoTが非常に注目され関心を集めると共に、履歴データなどを用いたサービスが急増している。例えば、データセンターやデータクラウドなどにおける大規模データベースに蓄積したデータから、データマイニング技術を用いてルールやパターンを抽出し、情報推薦(Recommendation)などを行う各種サービスが相当する。しかし、処理速度の高速化が進む一方、個人情報を含むデータベースに対するデータセキュリティに配慮したデータマイニング技術への関心は高いと言えない。

今後、例えば、データ同定(Data Identification)を用いて、複数運用組織が構築するデータベースを安全に連携し活用するうえで、プライバシー保護データマイニングの技術研究が重要である。特に、2000年代以後、データマイニングとプライバシーの問題は、国内外で活発に議論されており、多数の提案がなされており、「完全準同型暗号(Fully Homomorphic Encryption)」が、プライバシー保護データマイニング研究に大きな影響を与えてきた。そこで、複数データベースを連携するコンテキストベース位置情報サービスに用いる情報推薦のデータマイニングアルゴリズムの研究調査を進める。

(2) プライバシー保護データマイニング技術として、垂直分割モデルと水平分割モデルや、匿名化と暗号化とランダム化などの手法が用いられる。その中でも、暗号化技術の可能性に着目し、交通データウェアハウスを対象に、完全準同型暗号を用いたデータクラウド上で車両データの平均速度の計算などを可能とするデータクラウドのプロトタイプシステム構築を進めた。

本研究では、複雑なデータマイニングアルゴリズムをプライバシー保護制約下で実装することを目標とする。具体的には、携帯デバイスを用いたプローブデータベースと、交通データウェアハウス等のデータベースのプライバシーを互いに保護しつつ、コンテキストベース位置情報サービスを実現するアルゴリズムの実装である。より一般的には、データ同定などを用いてデータ統合が可能な複数データベースに対し、コンテキストベース位置情報サービスを実現するアルゴリズムをプライバシー保護制約下で実行するクラウド環境の構築と位置づけた研究を進める。

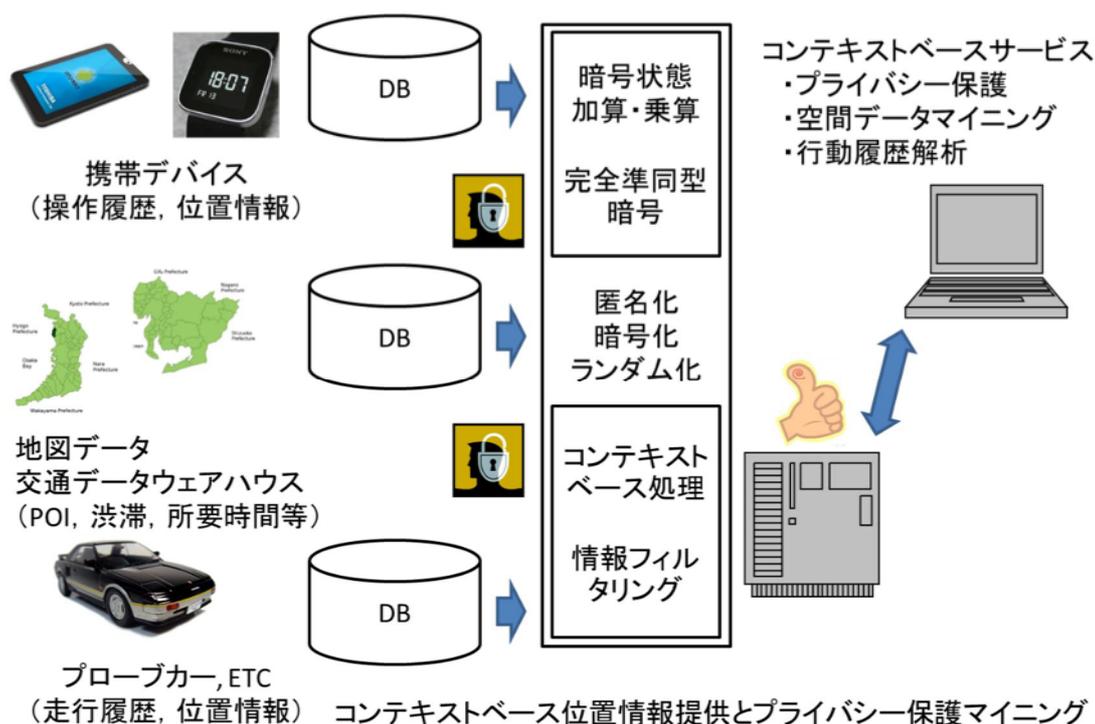
4. 研究成果

(1) 平成 29(2017)年度(初年度)は、プライバシー保護アルゴリズムの高速化に関する研究調査を進めた。特に、位置情報・加速度情報・渋滞情報などを含む行動履歴データに基づくコンテキストベース位置情報サービスの基礎となるデータマイニングアルゴリズムに関する研究調査を行なった。

加えて、より複雑な演算処理を含むアルゴリズムとなる情報プライバシー保護データマイニングの研究調査も進めた。さらに、プライバシー保護データマイニングアルゴリズムを実

装したプロトタイプシステムの実験環境整備を進めた。(下図)

なお、開始年度前年の年度末に、過年度までデータ処理実験環境を構築してきた機器にトラブルが生じたため、実験環境の再構築作業を円滑に進めることができず進捗が遅れが出た。



- (2) 平成 30(2018)年度 (第二年度) は、分散データベース環境におけるプライバシー保護アルゴリズムの高速化に関する研究調査を進めた。特に、複雑な演算処理を含むアルゴリズムとなる情報プライバシー保護データマイニングの研究調査を進めた。プライバシー保護データマイニングアルゴリズムの実験環境整備を進めつつある。

初年度はデータ処理環境の再構築を行なったが、二年度の途中で、クライアント側の機器故障が発生し、修復が不可能となったため、新たな機器導入と環境修復を行うための作業時間が発生した。一昨年度に引き続き、実験環境の再構築が生じたため、データ処理環境の導入に遅れが出た。

- (3) 令和元年度(2019) (第三年度) は、分散データベース環境におけるプライバシー保護アルゴリズムの特性に関する研究調査を進めた。研究では、Data Privacy, Data Security, Data Mining Privacy の観点から、各種クラウド構成に対するセキュリティ、データ分割手法による複合的な演算処理を含むアルゴリズムとなる情報プライバシー保護データマイニングの研究調査を進めた。

また、システム構成に関して、Centric Data Mining, Distributed Data Mining, Partitional Data Mining を考え、特に、クラウド構成ではなく、分散 DB 構成の場合のシステム特性に関して分析を行った。なお、COVID-19 のパンデミックの影響で、国内国際会議とともに多くの会議中止となった。

- (4) 令和 2 年度(2020) (最終年度) は、プライバシー保護データマイニング(Privacy Preserving Data Mining)技術を用いて、位置情報サービスプロバイダにおけるデータセキュリティを担保した上で、より高精度なコンテキストベースサービスを提供するデータマイニングアルゴリズムの研究を行った。

- (5) 令和 3(2021)年度、令和 4(2022)年度と、パンデミックが半導体市場に影響を与え、予定予算で実験環境構築にかかる機器発注が困難となった。加えて、渡航制限が行われ、会議自体の中止も続いた。オンライン開催の一部会議で発表を進めた。令和 5(2023)年度も、国際会議への円滑な準備ができなかった。結果的に、当初の本科研費の最終年度となる 2020 年度以後、補助事業期間延長された予算の一部を返還し研究期間を終了した。

5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 2件／うち国際共著 0件／うちオープンアクセス 0件）

1. 著者名 Megumi Yamamoto, Nobuo Umemura, Hiroyuki Kawano	4. 巻 19
2. 論文標題 Proposal of Japanese Vocabulary Difficulty Level Dictionaries for Automated Essay Scoring Support System Using Rubric	5. 発行年 2019年
3. 雑誌名 Journal of the Operations Research Society of China	6. 最初と最後の頁 pp. 1-17
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/s40305-019-00270-z	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 山本 恵, 梅村 信夫, 河野 浩之	4. 巻 2018-CE-145, 12
2. 論文標題 レポート自動採点支援用日本語語彙レベル辞書の提案—Wikipediaコーパスの利用—	5. 発行年 2018年
3. 雑誌名 情報処理学会研究報告（コンピュータと教育(CE)）	6. 最初と最後の頁 1-7
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 無
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 竹田拓哉, 河野浩之, 石原靖哲	4. 巻 L-012
2. 論文標題 スマートホームシステムにおけるDempster-Shafer Theoryに基づく異常検知手法の提案	5. 発行年 2018年
3. 雑誌名 第17回情報科学技術フォーラム	6. 最初と最後の頁 153-156
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 無
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 Megumi Yamamoto, Nobuo Umemura and Hiroyuki Kawano	4. 巻 727
2. 論文標題 Automated Essay Scoring System Based on Rubric	5. 発行年 2018年
3. 雑誌名 Studies in Computational Intelligence, Applied Computing & Information Technology	6. 最初と最後の頁 177-190
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 -

[学会発表] 計5件 (うち招待講演 0件/うち国際学会 2件)

1. 発表者名 林 亮佑, 大極康平, 河野浩之
2. 発表標題 CNNを用いた画像鮮明化システムの実装
3. 学会等名 第21回情報科学技術フォーラム (FIT2022)
4. 発表年 2022年

1. 発表者名 手塚悠, 河野浩之
2. 発表標題 3次元姿勢推定による警察官の手信号認識の学習データ量の削減
3. 学会等名 情報処理学会第85回全国大会
4. 発表年 2023年

1. 発表者名 野首優斗, 河野浩之
2. 発表標題 安価な筋電センサによる食感の異なる複数食品のカテゴリ分類
3. 学会等名 情報処理学会第85回全国大会
4. 発表年 2023年

1. 発表者名 Hiroyuki Kawano
2. 発表標題 Privacy-Preserving Data Mining Algorithms for Distributed Database
3. 学会等名 30th European Conference on Operational Research (EURO 30th) (国際学会)
4. 発表年 2019年

1. 発表者名 Megumi Yamamoto, Nobuo Umemura, Hiroyuki Kawano
2. 発表標題 Proposal of Japanese Vocabulary Words List for Automated Essay Scoring Support System Using Rubric
3. 学会等名 The Thirteenth International Symposium on Operations Research and Its Applications (ISORA 2018) (国際学会)
4. 発表年 2018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

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
---------	---------