様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　５ 年　 ６ 月 １５ 日現在

機関番号：１０１０１
研究種目：若手研究
研究期間：2019～2022
課題番号：１９Ｋ２０２６９

研究課題名（和文）Achieving Differential Privacy under Spatiotemporal Correlations

研究課題名（英文）Achieving Differential Privacy under Spatiotemporal Correlations

研究代表者
曹　洋（Cao, Yang）

北海道大学・情報科学研究院・准教授

研究者番号：６０８３６３４４

交付決定額（研究期間全体）：（直接経費）　　3,100,000円

研究成果の概要（和文）：差分プライバシーは、プライバシー標準として広く研究・展開されています。本プロジェクトでは、時空間データに対して差分プライバシーを適用する際の潜在的なリスクと効用の不十分さを示します。時空間データのための新しい柔軟なプライバシー概念、例えば、Geo-graph-indistinguishability（DBSec 2019、IEICE 2023）、時空間イベントプライバシー（IEEE ICDE 2019、IEEE TKDE 2019）、およびポリシーベースの位置プライバシー（ESORICS 2020）を提案し、プライバシーと厳密性と柔軟性のトレードオフをより良く達成できます。

研究成果の学術的意義や社会的意義
時空間データの収集と分析は、多くの研究分野や新興産業の基盤となっており、例えば、スマートシティ、交通予測、人流統計、クラウドソーシング、自動運転などがあります。しかし、プライバシーは無視できない障壁となることが多いです。本研究の成果は、時空間データ駆動型の科学技術の発展を支援することができます。

研究成果の概要（英文）：Differential Privacy has been extensively studied and deployed as the de facto privacy standard for preserving data privacy during collection and analysis. In this project, we demonstrate the potential risks and utility insufficiency of Differential Privacy when applied to spatiotemporal data. We propose new, flexible privacy notions for spatiotemporal data, such as Geo-graph-indistinguishability (DBSec 2019, IEICE 2023), Spatiotemporal Event Privacy (IEEE ICDE 2019, IEEE TKDE 2019), and Policy-based Location Privacy (ESORICS 2020) to achieve a better privacy-utility tradeoff.

研究分野：データベース

キーワード：差分プライバシ　時空間データ

１．研究開始当初の背景

In recent decades, privacy breaches have been on the rise, leading to significant financial losses for companies and a decrease in trust towards big data technologies. This has prompted researchers to study how to formally define privacy and how to achieve it. Introduced by Dwork in 2006, Differential Privacy (DP) has emerged as a provable privacy definition and has been widely studied across multiple research communities, ranging from cryptography to database management. It has become the de facto standard for privacy-preserving data analysis. For instance, Google, Apple, and the U.S. Census Bureau have recently adopted DP techniques for collecting large-scale personal data.

Differential Privacy (DP) as a rigorous privacy definition has garnered increasing attention. Many studies employ traditional DP mechanisms, such as the Laplace mechanism and Exponential mechanism, as primitives to release information from sensitive data. These mechanisms implicitly assume that data are independent. However, real-life data tend to be spatiotemporally correlated, and such correlations may result in unexpected privacy leakage. Determining how to properly achieve DP under spatiotemporal correlations remains an open problem.

２．研究の目的

We are examining the scenario of using DP to collect and analyze spatiotemporal data. The central question is: how can we properly achieve ε-DP under spatiotemporal data? We have divided this question into three sub-questions as follows:

- What are the privacy risks associated with traditional DP mechanisms when considering spatiotemporal correlations?
- How can we enhance traditional DP mechanisms to mitigate these privacy risks?
- Do the enhanced DP mechanisms offer composability and maintain bounded privacy leakage?

３．研究の方法

Method 1. Assessing attack-based privacy risks of DP under spatiotemporal data.

We evaluate the vulnerability of traditional ε-DP mechanisms by simulating adversaries who possess knowledge of spatiotemporal correlations and various adversarial goals. For example, given a specific ε, we assess the probability of users' true data being accurately inferred (i.e., the adversarial goal is data reconstruction) from the perturbed data output by a traditional ε-DP mechanism.

Method 2. Designing enhanced DP mechanisms to mitigate the above privacy risks and improve data utility.

We aim to develop a generic method called "sensitivity analysis on correlated data" for calculating an appropriate (smaller) ε to strengthen traditional DP mechanisms and control the aforementioned privacy risks. Since a smaller ε implies larger perturbation, it is necessary to enhance the utility of the perturbed data.

Method 3. Formalizing theoretical properties (composability and upper bound) of the above mechanisms.

These properties have been well-established in traditional DP but remain open questions under spatiotemporal correlations. Our new composition theorem demonstrates the overall privacy guarantee of a complex algorithm that uses the enhanced DP mechanisms as primitives. The upper bound indicates the worst leakage of the enhanced DP mechanisms given varying strengths of spatiotemporal correlations.

４．研究成果

Building on the methods discussed earlier, we have proposed the following three flexible privacy notions for spatiotemporal data:

1. Geo-graph-indistinguishability for spatiotemporal data on road networks (DBSec 2019, IEICE 2023),
2. Spatiotemporal Event Privacy for customizable privacy preferences (IEEE ICDE 2019, IEEE TKDE 2019), and

3. Policy-based Location Privacy for a unified location privacy framework (ESORICS 2020).

More details are presented below.

4.1 Numerous methods have been proposed to protect location privacy over the past few decades. In particular, perturbation methods based on Geo-Indistinguishability (GeoI), which randomly perturb a true location to a pseudo-location, have garnered attention due to their strong privacy guarantee inherited from differential privacy. However, GeoI is based on the Euclidean plane, even though many LBSs are based on road networks (e.g., ride-sharing services). This results in unnecessary noise and an insufficient tradeoff between utility and privacy for LBSs on road networks. To address this issue, we propose a new privacy notion, Geo-Graph-Indistinguishability (GeoGI), for locations on a road network to achieve a better tradeoff. We introduce the Graph-Exponential Mechanism (GEM), which satisfies GeoGI. Furthermore, we formalize the optimization problem for finding the optimal GEM in terms of the tradeoff. However, the computational complexity of a naive method for finding the optimal solution is prohibitive, so we propose a greedy algorithm to find an approximate solution within an acceptable timeframe. Finally, our experiments demonstrate that our proposed mechanism outperforms GeoI mechanisms, including the optimal GeoI mechanism, with respect to the tradeoff.

4.2 Location privacy-preserving mechanisms (LPPMs) have been extensively studied for protecting users' location privacy by releasing a perturbed location to third parties, such as location-based service providers. However, when a user's perturbed locations are released continuously, existing LPPMs may not protect sensitive information about the user's real-world activities, such as visiting a hospital in the past week or regularly commuting between locations A and B every weekday (it is easy to infer that location A and location B may be home and office). We refer to this as spatiotemporal events. In this paper, we first formally define spatiotemporal events as Boolean expressions between location and time predicates, and then we define ε-spatiotemporal event privacy by extending the notion of differential privacy. Second, to understand how much spatiotemporal event privacy existing LPPMs can provide, we design computationally efficient algorithms to quantify the spatiotemporal event privacy leakage of state-of-the-art LPPMs. It turns out that existing LPPMs may not adequately protect spatiotemporal event privacy. Third, we propose a framework, PriSTE, to transform an existing LPPM into one that protects spatiotemporal event privacy by calibrating the LPPM's privacy budgets. Our experiments on real-life and synthetic data verify that the proposed method is effective and efficient.

4.3 Location privacy has been extensively studied in the literature. However, existing location privacy models are either not rigorous or not customizable, which limits the trade-off between privacy and utility in many real-world applications. To address this issue, we propose a new location privacy notion called PGLP, i.e., Policy Graph-based Location Privacy, providing a rich interface to release private locations with customizable and rigorous privacy guarantees. First, we design rigorous privacy for PGLP by extending differential privacy. Specifically, we formalize location privacy requirements using a location policy graph, which is expressive and customizable.

５．主な発表論文等

〔雑誌論文〕　計3件（うち査読付論文　3件／うち国際共著　3件／うちオープンアクセス　2件）

１．著者名 Cao Yang、Xiao Yonghui、Xiong Li、Bai Liquan、Yoshikawa Masatoshi	４．巻 33
２．論文標題 Protecting Spatiotemporal Event Privacy in Continuous Location-Based Services	５．発行年 2020年
３．雑誌名 IEEE Transactions on Knowledge and Data Engineering	６．最初と最後の頁 3141-3154
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/TKDE.2019.2963312	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 該当する

１．著者名 Ma Shuaicheng, Cao Yang, Xiong Li	４．巻 13
２．論文標題 Efficient logging and querying for Blockchain-based cross-site genomic dataset access audit	５．発行年 2020年
３．雑誌名 BMC Medical Genomics	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） 10.1186/s12920-020-0725-y	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 該当する

１．著者名 Caixia Yang ; Liang Tan ; Na Shi ; Bolei Xu ; Yang Cao ; Keping Yu	４．巻 8
２．論文標題 AuthPrivacyChain: A Blockchain-Based Access Control Framework With Privacy Protection in Cloud	５．発行年 2020年
３．雑誌名 IEEE Access	６．最初と最後の頁 70604-70615
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/ACCESS.2020.2985762	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 該当する

〔学会発表〕　計31件（うち招待講演　2件／うち国際学会　22件）

１．発表者名 Fumiyuki Kato, Yang Cao, Masatoshi Yoshikawa
２．発表標題 Preventing Manipulation Attack in Local Differential Privacy Using Verifiable Randomization Mechanism.
３．学会等名 DBSec（国際学会）
４．発表年 2021年

１．発表者名
Shuaicheng Ma, Yang Cao, Li Xiong

２．発表標題
Transparent Contribution Evaluation for Secure Federated Learning on Blockchain

３．学会等名
ICDE workshop（国際学会）

４．発表年
2021年

１．発表者名
Ruixuan Liu

２．発表標題
FedSel: Federated SGD under Local Differential Privacy with Top-k Dimension Selection

３．学会等名
DASFAA 2020（国際学会）

４．発表年
2020年

１．発表者名
Xiaolan Gu

２．発表標題
PCKV: Locally Differentially Private Correlated Key-Value Data Collection with Optimized Utility.

３．学会等名
USENIX Security 2020（国際学会）

４．発表年
2020年

１．発表者名
Yang Cao

２．発表標題
PANDA: Policy-aware Location Privacy for Epidemic Surveillance.

３．学会等名
VLDB 2020（国際学会）

４．発表年
2020年

１．発表者名
Yaowei Han

２．発表標題
Voice-Indistinguishability: Protecting Voiceprint in Privacy Preserving Speech Data Release.

３．学会等名
IEEE ICME 2020（国際学会）

４．発表年
2020年

１．発表者名
Yang Cao

２．発表標題
PGLP: Customizable and Rigorous Location Privacy through Policy Graph.

３．学会等名
ESORICS 2020（国際学会）

４．発表年
2020年

１．発表者名
Fumiyuki Kato

２．発表標題
Secure and Efficient Trajectory-Based Contact Tracing using Trusted Hardware.

３．学会等名
7th International Workshop on Privacy and Security of Big Data @IEEE BigData 2020（国際学会）

４．発表年
2020年

１．発表者名
Shun Takagi

２．発表標題
P3GM: Private High-Dimensional Data Release via Privacy Preserving Phased Generative Model.

３．学会等名
IEEE ICDE 2021（国際学会）

４．発表年
2021年

１．発表者名
Ruixuan Liu

２．発表標題
FLAME: Differentially Private Federated Learning in the Shuffle Model.

３．学会等名
AAAI 2021（国際学会）

４．発表年
2021年

１．発表者名
Wei Song

２．発表標題
Privacy-Preserving Polynomial Evaluation over Spatio-Temporal Data on An Untrusted Cloud Server.

３．学会等名
DASFAA 2021（国際学会）

４．発表年
2021年

１．発表者名
Shun Takagi, Yang Cao, Yasuhito Asano and Masatoshi Yoshikawa

２．発表標題
Geo-Graph-Indistinguishability: Protecting Location Privacy for LBS over Road Networks

３．学会等名
DBSec（国際学会）

４．発表年
2019年

１．発表者名
Yang Cao, Yonghui Xiao, Li Xiong, Liquan Bai

２．発表標題
PriSTE: From Location Privacy to Spatiotemporal Event Privacy

３．学会等名
IEEE ICDE short paper（国際学会）

４．発表年
2019年

１．発表者名
Chunmiao Li, Yang Cao, Zhenjiang Hu, Masatoshi Yoshikawa

２．発表標題
Blockchain-based Bidirectional Updates on Fine-grained Medical Data.

３．学会等名
BlockDM workshop at IEEE ICDE 2019（国際学会）

４．発表年
2019年

１．発表者名
高木 駿, 曹 洋, 浅野 泰仁, 吉川 正俊

２．発表標題
道路ネットワークにおける位置情報プライバシー

３．学会等名
DEIM 2019

４．発表年
2019年

１．発表者名
Xiaolan Gu, Ming Li, Yang Cao, Li Xiong

２．発表標題
Supporting both Range Queries and Frequency Estimation with Local Differential Privacy

３．学会等名
IEEE Conference on Communications and Network Security (CNS) 2019（国際学会）

４．発表年
2019年

１．発表者名
Maho Asada, Masatoshi Yoshikawa, Yang Cao

２．発表標題
When and where do you want to hide? Recommendation of location privacy preferences with local differential privacy.

３．学会等名
DBSec 2019（国際学会）

４．発表年
2019年

１．発表者名
Yang Cao, Yonghui Xiao, Li Xiong, Liquan Bai, Masatoshi Yoshikawa

２．発表標題
PriSTE: Protecting Spatiotemporal Event Privacy in Continuous Location-Based Services.

３．学会等名
VLDB 2019, demo track（国際学会）

４．発表年
2019年

１．発表者名
Xiaolan Gu, Ming Li, Yueqiang Cheng, Li Xiong and Yang Cao

２．発表標題
PCKV: Locally Differentially Private Correlated Key-Value Data Collection with Optimized Utility.

３．学会等名
USENIX Security 2020（国際学会）

４．発表年
2020年

１．発表者名
Xiaolan Gu, Ming Li, Li Xiong and Yang Cao

２．発表標題
ID-LDP: Providing Input-Discriminative Protection for Local Differential Privacy.

３．学会等名
IEEE ICDE 2020（国際学会）

４．発表年
2020年

１．発表者名
Ruixuan Liu, Yang Cao, Masatoshi Yoshikawa, Hong Chen

２．発表標題
Federated SGD under Local Differential Privacy with Top-k Dimension Selection.

３．学会等名
25th International Conference on Database Systems for Advanced Applications (DASFAA) 2020（国際学会）

４．発表年
2020年

１．発表者名
Yaowei Han, Sheng Li, Yang Cao, Qiang Ma, Masatoshi Yoshikawa

２．発表標題
Voice-Indistinguishability: Protecting Voiceprint in Privacy Preserving Speech Data Release.

３．学会等名
IEEE ICME 2020 (Oral)（国際学会）

４．発表年
2020年

１．発表者名
Shuyuan Zheng, Yang Cao, Masatoshi Yoshikawa

２．発表標題
Money Cannot Buy Everything: Trading Mobile Data with Controllable Privacy Loss.

３．学会等名
IEEE MDM 2020（国際学会）

４．発表年
2020年

１．発表者名
高木駿，曹洋，吉川正俊

２．発表標題
局所差分プライバシにおけるパラメータの秘匿について

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(DEIM2020)

４．発表年
2020年

１．発表者名
高木駿，高橋翼，曹洋，吉川正俊

２．発表標題
段階的学習を用いたプライバシ保護型深層生成モデル．

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(DEIM2020)．

４．発表年
2020年

１．発表者名
加藤郁之，曹洋，吉川正俊

２．発表標題
TEEに基づく差分プライバシの検証

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(DEIM2020).

４．発表年
2020年

１．発表者名
峯田初音，韓耀緯，曹洋，吉川正俊

２．発表標題
局所差分プライバシを用いた行列分解によるネッ　ト広告システムの提案

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(DEIM2020).

４．発表年
2020年

１．発表者名
加納英樹,加藤郁之,ティブシメディ,阿部正幸,曹洋

２．発表標題
プライバシ保護深層学習のための SGX分散処理の提案.

３．学会等名
2020年暗号と情報セキュリティシンポジウム(SCIS2020)

４．発表年
2020年

１．発表者名
成瀬真，高木駿，曹洋，吉川正俊

２．発表標題
道路ネットワークにおける位置情報プライバシを考 慮した軌跡データの評価に関する研究

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(DEIM2020).

４．発表年
2020年

１．発表者名 Yang Cao
２．発表標題 Towards Decentralized and Privacy-Preserving Personal Data Market
３．学会等名 IEEE Tokyo Blockchain Workshop 2019（招待講演）
４．発表年 2019年

１．発表者名 Yang Cao
２．発表標題 From Location Privacy to Spatiotemporal Event Privacy
３．学会等名 Wuhan University（招待講演）
４．発表年 2019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