

LINE、データ関連分野の国際学会「SIGMOD2022」「VLDB2022」にて論文採択

2022.06.15 技術情報

プライバシー保護を前提とした機械学習およびデータベース領域の新技术を開発

LINE株式会社（本社：東京都新宿区、代表取締役社長：出澤 剛）は、データ関連分野における世界トップレベルの国際会議「SIGMOD2022」および「VLDB2022」にて、プライバシー保護の研究成果に関する論文が採択されましたので、お知らせいたします。

SIGMOD (The ACM Special Interest Group on Management of Data、主催：Association for Computing Machinery)、VLDB (48th International Conference on Very Large Databases、主催：Very Large Data Base Endowment Inc.)は、いずれもデータベース・データ工学分野における世界トップレベルの国際会議です。なおLINEとしてはSIGMOD、VLDBともに今回が初の論文採択となります。このたび採択された論文は、京都大学大学院情報学研究科 吉川正俊教授、曹洋特定准教授との共同研究成果であり、6月12日（日）～17日（金）開催の「SIGMOD2022」および9月開催予定の「VLDB2022」にて、それぞれ発表を行います。

■LINEが注力するプライバシー保護技術の研究開発

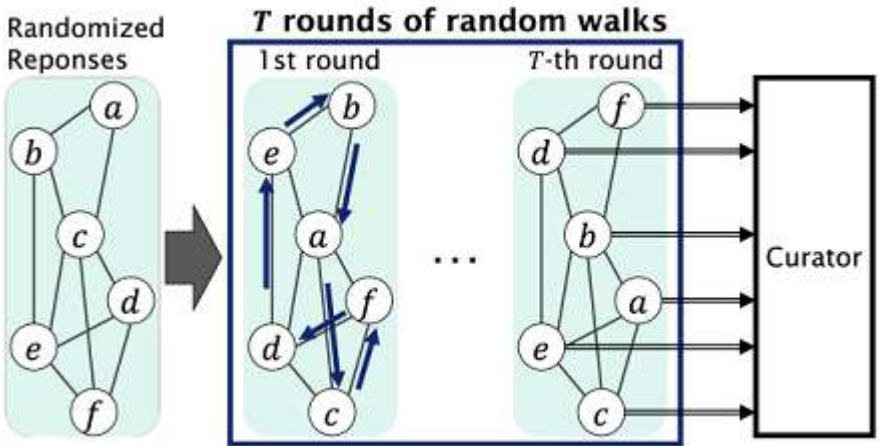
LINEでは、ユーザーデータを活用したパーソナライゼーションに力を入れており、同時にデータを扱う際のプライバシーへの配慮についても重要視しています。近年、国際的な規制の整備も進み、プライバシー保護の技術や考え方もめざましい発展を遂げています。時流に即した最適なプライバシーモデルの追求と導入は、プラットフォームとしての重要な責任です。LINEでは、連合学習や差分プライバシー、秘密計算等の先端的なプライバシー保護型機械学習技術の検証や実装を推進し、十分なプライバシーへの配慮と多様なユーザーへ深いパーソナライズの両立を目指しています。プライバシー保護に関する研究開発の成果として、これまでLINEとしては、「ICDE2021」や「ICLR2022」といった世界トップレベルの国際会議で、差分プライバシーに関する論文の採択実績があります。差分プライバシーはユーザーデータ収集・活用にあたって、所定のノイズやランダム性を追加することによって、あらゆる人と見分けがつかない出力結果とする際に用いられる、数学的に厳密なプライバシー基準です。現在、LINEでは差分プライバシーによるプライバシー保護型データ活用の実用化に向け、研究開発に取り組んでいます。

■世界初、差分プライバシーの増幅を分散型プロトコルで実現

今回の「SIGMOD」では、クライアントからサーバーへのデータ収集時におけるプライバシー保護に関して、クライアント上で保証した差分プライバシーの強度を、クライアント間のセキュアなデータ交換によって増幅させる技術（プライバシー増幅*1）の新しい方式として「Network Shuffling」を発表しました。従来のプライバシー増幅は、送信元の匿名化処理であるシャッフリングを、中央集権的（Centralized）かつ信頼できる第三者によって実施する必要がありました。提案法「Network Shuffling」は、クライアント間がE2E (End-to-end)で暗号化された通信を行うソーシャルネットワークであることを前提として、様々なクライアント間でデータを秘密裏に交換し合うことにより、第三者による中央集権的なサーバーを必要とせず、分散型（Decentralized）のシャッフリングを実現します。このクライアント間のデータ交換をグラフ上のランダムウォーク*2として定式化し、既存のシャッフリングと同様のプライバシー増幅効果が得られることを示しました。

分散型のプライバシー増幅手法「Network Shuffling」の概念図

(Tラウンドのセキュアなデータ交換 (ランダムウォーク) によって、送信元を匿名化)



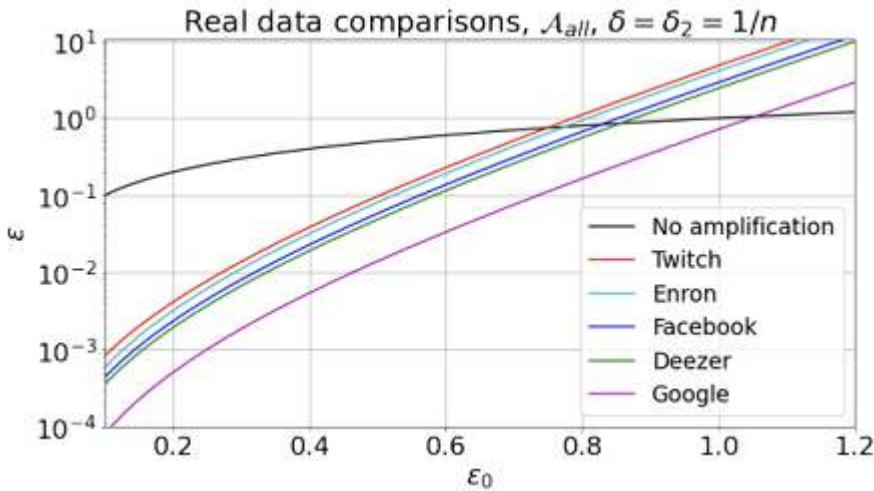
プライバシー増幅の効果

(「Network Shuffling」は、既存のUniform Shufflingと同様のプライバシー増幅効果を持つことを証明)

Mechanism	Privacy Amplification
No amplification [18]	ϵ_0
Uniform subsampling [1, 33]	$O(e^{\epsilon_0}/\sqrt{n})$
Uniform shuffling [22]	$O(e^{3\epsilon_0}/\sqrt{n})$
Uniform shuffling (w/ clones) [25]	$O(e^{0.5\epsilon_0}/\sqrt{n})$
Network shuffling (ours)	$O(e^{1.5\epsilon_0}/\sqrt{n})$

グラフデータ毎のプライバシー増幅効果

(グラフの規模が大きいくほど、プライバシー増幅効果が高い傾向)

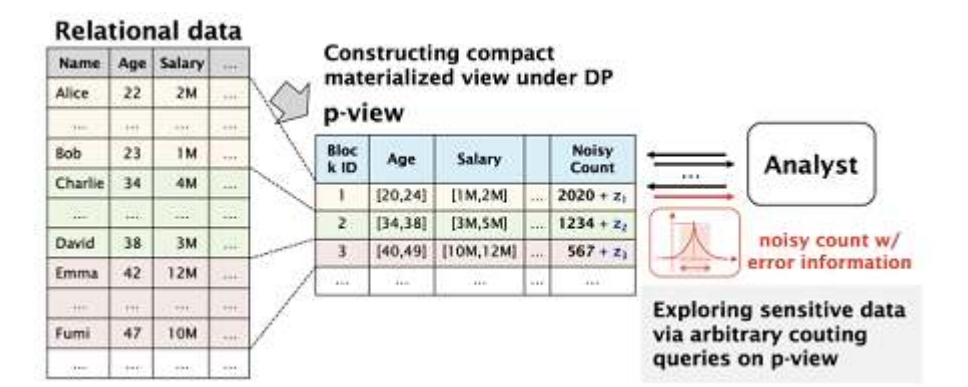


■有用性とプライバシーを両立した、空間効率の高いデータベース問合せを実現

「VLDB」では、データベースへの問合せ処理を差分プライバシーによって保護する際、任意のクエリに対し、少ないノイズで応答可能な中間データ（ビュー）を構築する手法「HDPView」を提案しました。差分プライバシーによる問合せ処理を実現する方法には、クエリ応答の度にノイズを加算する方法と、事前にノイズを加算した中間データを構築する方法があります。前者には問合せ回数等の制限があること、また後者の既存技術は事前に決められたクエリにのみ最適化されていることから、いずれも探索的なデータ解析には適していないという課題がありました。提案法「HDPView」は、これらの課題の解決を目指し、クエリに非依存、誤差の見積もりが可能、高次元データに適用可能、高い空間効率といった特徴を有しています。評価実験では、主に解析の誤差やビューのサイズの点で、既存の手法を大きく上回る性能を示しました。

「HDPView」が解決した課題

任意のクエリに対し、少ないノイズで応答するためのビュー (p-view) を構築



クエリの誤差とビューのサイズの比較

(「HDPView」は既存手法を上回る性能を示す)

	Identity ¹ [15]	Privtree [45]	HDMM [30]	PrivBayes [44]	HDPView (ours)
Average relative RMSE	1.94×10^7	7.05	35.34	3.79	1.00
Average relative size of p-view	4.59×10^{17}	5578.27	N/A	N/A	1.00

*1 クライアントで保証した差分プライバシーの強度を高めるために用いる匿名化処理。送信元を匿名化するシャッフリングによって所定の増幅効果を得る。

*2 次のステップの位置がランダムに決定される運動。本研究ではグラフの頂点上の移動を扱い、次の位置をグラフの接続関係の中からランダムに決定する。

採択された論文

●Network Shuffling: Privacy Amplification via Random Walks

Seng Pei Liew, Tsubasa Takahashi, Shun Takagi, Fumiyuki Kato, Yang Cao, Masatoshi Yoshikawa
(SIGMOD2022)

●HDPView: Differentially Private Materialized View for Exploring High Dimensional Relational Data

Fumiyuki Kato, Tsubasa Takahashi, Shun Takagi, Yang Cao, Seng Pei Liew, and Masatoshi Yoshikawa
(VLDB2022)

【LINEが注力する基礎研究について】

LINEでは、ユーザーのプライバシー保護をはじめとした適切なデータ運用、AI関連サービスや機能の創出を支える、さまざまな技術の基礎研究を組織的に推進しています。

基礎研究においては、機械学習を軸に、音声処理、言語処理、画像処理などの領域に注力しています。直近では、音声・音響信号処理分野の国際学会「ICASSP 2022」にて3本の論文*3、音声処理分野の国際学会「INTERSPEECH 2021」にて6本の論文*4、コンピュータービジョン分野の国際学会「ICCV 2021」にて2本の論文*5が採択されるなど、着実な成果を収めております。

*3 2022年2月17日発表プレスリリース： <https://linecorp.com/ja/pr/news/ja/2022/4120>

*4 2021年8月30日発表プレスリリース： <https://linecorp.com/ja/pr/news/ja/2021/3891>

*5 2021年7月28日発表プレスリリース： <https://linecorp.com/ja/pr/news/ja/2021/3843>

LINEでは、今後も事業発展やサービスの品質向上のための取り組みを積極的に行い、コミュニケーションプラットフォームとしてさらなる成長・拡大の可能性を広げてまいります。