



Title	ZDDを用いた大規模データ匿名化基盤の検討
Author(s)	白井, 康之
Citation	2010年度科学技術振興機構ERATO湊離散構造処理系プロジェクト講究録. p.377.
Issue Date	2011-06
Doc URL	<a href="http://hdl.handle.net/2115/48367">http://hdl.handle.net/2115/48367</a>
Type	conference presentation
Note	ERATO 湊離散構造処理系プロジェクトシンポジウム（第1回）：第9回情報科学技術フォーラム(FIT2010)イベント企画セッション. 2010年9月8日（水）. 九州大学伊都キャンパス.
File Information	05.FIT_shirai.pdf



[Instructions for use](#)

## ZDDを用いた大規模データ匿名化基盤の検討

2010年9月 JST-ERATO 白井康之

### データベースの匿名化

●事業者が持つデータ(個人情報)を組み合わせると意味のあるデータを作ることができる。  
 ●しかし、個別の(個人が特定・または推定できる)個人情報が流出するのは問題。

●個人情報とは、氏名、住所、生年月日等、個人を「直接的に」特定できる情報  
 ●しかし、実は直接的に特定できなくても「間接的」に識別できる情報が存在。

●そこで...  
 データを流用する場合には、情報の匿名化が必要である。  
 データを分析する場合には、分析結果のみを開示できる仕組みが必要である。  
 ●前者は個人情報匿名化、後者はセキュア計算と呼ばれる。

●情報の匿名化では、データの削除、一般化により、準識別情報からデータを与えられた条件(k)以下に特定できないようにデータを加工・編集する。

**医療情報**  
 投薬情報、病歴情報を共有し、医療機関、薬メーカ等で利用。薬剤メーカでニーズが高く、一部実現。  
 クレジット情報  
 過去のクレジット履歴を共有し、クレジットカード会社で利用(個人特定)。今後は存の枠組みを超えた情報共有も  
 購入履歴情報・サービス利用履歴  
 過去の他人の利用履歴からの推薦。  
 単一店舗でのリコメンデーションは実現されているが、危ういケースも。

(過去の主な関連研究 [k-匿名化関連])  
 P. Samarati, L. Sweeney, 1998, [L. Sweeney, 2002], [K. LeFevre, 2005]  
 [広島市立大学, 2007]  
 [JIPDEC, AIST情報セキュリティ研究センター, NTT情報流通プラットフォーム研究所, 三菱総研, 情報大航海 Prj. 2009]

### BDD/ZDDによる効率的な表現

BDD(Binary Decision Diagram), ZDD(Zero Suppressed BDD):  
 ▶論理関数を効率的に表現・計算するための手法  
 ▶共通部分集合の表現、スペースの構造の表現において圧縮効果が高い

二分決定木による  $F = (a \vee b) \wedge c$  の表現(左)とBDDによる表現(右)

S. Minato, "Binary Decision Diagram and Applications for VLSI CAD", 1996

### ZDDに基づくデータ表現と匿名化基盤への拡張

**ZDD処理系での実行例**

システム構成(例)

データの組み合わせの積と和表示

最小値、最大値(係数)

指定された値以上の係数を持つ項の表示

データの追加(更新)

データの抽象化レベルの更新

データ抽出・特徴抽出

配布データ(1) (2) (3) ...

Original Data

AbstLevel = (1,1,1)

Work0, Work1, Work2, Job categories

### BDD/ZDDによる効率・今後の課題

**既存研究の方法論**  
 データベースシステムのアプリケーションとしての構築  
 パッチ処理を前提

**今後の課題**

- 変数順序の最適化 (効率性・利用局面の考慮)
- 順序集合や組み合わせ集合への対応検討 (ex. 商品の購入の組み合わせや履歴(系列)など)
- ライブラリ機能(ZDD処理系)との連携・実装 (データ変換・抽出機能や問い合わせ機能の検討)
- 他の匿名化アルゴリズムとの連携検討 (効果的な匿名化基準やあいまい化手法との連携など)

**ZDDを用いた方法論の特徴**

疎なデータセットに対するコンパクトな表現とデータ抽出メカニズム  
 逐次的なデータ更新への対応