1J5-OS-18b-6

# プライバシ保護クラウドソーシング

Privacy Preserving Crowdsourcing

梶野 洸[*1] 馬場 雪乃[*2*3] 鹿島 久嗣[*4]
Hiroshi Kajino Yukino Baba Hisashi Kashima

[*1]東京大学大学院情報理工学系研究科数理情報学専攻
Department of Mathematical Informatics, Graduate School of Information Science and Technology, The University of Tokyo

[*2]国立情報学研究所ビッグデータ数理国際研究センター
Global Research Center for Big Data Mathematics, National Institute of Informatics

[*3]JST, ERATO, 河原林巨大グラフプロジェクト
JST, ERATO, Kawarabayashi Large Graph Project

[*4]京都大学大学院情報学研究科知能情報学専攻
Department of Intelligence Science and Technology, Graduate School of Informatics, Kyoto University

We propose a qualitative evaluation method for privacy preserving crowdsourcing methods. It measures the privacy preservation capability and the task performance preservation capability of a method. We also propose a privacy preserving crowdsourcing method based on the idea of clipping instances. We experimentally investigated the properties of our method. The experimental results show that our method can execute tasks with privacy preserved if a task can be executed locally and a privacy definition depends on an instance globally.

## 1. 序論

クラウドソーシングとは、不特定多数のワーカーに仕事を依頼する仕組みである。依頼者は、クラウドソーシングサービスに仕事を登録し、ワーカーはその仕事を処理して得られる成果物を依頼者に返すことで報酬を得る。人の知能が必要な仕事を安価かつ大量に処理可能という利点がある一方、品質の問題 [Lease 11] を含む様々な問題が指摘されている。本研究では、仕事依頼時にワーカーに渡すデータからのプライバシ漏洩に関する問題を取り扱う。

データ処理タスクでは、仕事依頼時に依頼者はワーカーにデータ（以降、インスタンスと呼ぶ）を渡し、ワーカーはタスクの指示に従ってそのデータを処理し、処理結果を成果物とする。例として、写真のタグ付けや音声書き起こしなどが挙げられる。しかし、インスタンスに不特定多数に公開すべきでない情報が含まれる場合、プライバシの問題が生じるため、クラウドソーシングを用いるのは不適切となる。この問題に対する既存研究は主に二つ挙げられる。一つは、医療カルテ書き起こしに特化した研究である [Little 11]。カルテの雛形を利用して各項目を分割し、各項目を別々のワーカーに渡すことでプライバシ侵害を防ぐ。この手法は、インスタンスに雛形がなければ適用できないという欠点がある。もう一つは、プライバシ保護能力と成果物の品質のトレードオフを解析した研究である [Varshney 12]。一般に、プライバシ保護を強めるほど成果物の品質は下がると考えられ、Varshney はそのトレードオフを数理的に解析した。この研究は、実際のクラウドソーシングでの検証を行っていない点で不十分である。

本研究では、一般的なデータ処理タスクを対象に、プライバシ保護クラウドソーシング手法のプライバシ保護能力と成果物の品質のトレードオフの評価手法を提案し、さらにインスタンスの雛形を必要としない切抜型プライバシ保護手法を提案し、写真中の顔を隠すタスクを例に実験的評価を行った。

連絡先: 梶野 洸, hiroshi_kajino@mist.i.u-tokyo.ac.jp

## 2. 設定

クラウドソーシングでは、依頼者とワーカーという二人のやりとりを考える。依頼者は、仕事を依頼する人で、インスタンス $I \in \mathcal{I}$ に対してタスクを適用した結果 $R \in \mathcal{R}$ を成果物として受け取る。ワーカーは、仕事を実行する人で、インスタンス $I$ に対してタスクを適用して、結果 $R$ を生成し、報酬と引き換えに結果 $R$ を成果物として依頼者に返す。例えば、タスクは「画像が鳥ならば yes、鳥でないならば no を選択して下さい」という仕事の指示に相当し、インスタンスは各画像、成果物は {yes, no} のいずれかに相当する。本研究では、このタスク処理過程の他に、ワーカーによるプライバシ侵害過程を考える。つまり、ワーカーはインスタンス $I$ から秘匿情報 $S \in \mathcal{S}$ を抽出するとする。

## 3. プライバシ問題の定式化

本章では、序論で議論したクラウドソーシングでのプライバシ問題の定式化を行う。前章で導入したタスク処理過程とプライバシ侵害過程のモデル化し、それを元にプライバシ保護手法の評価指標を提案する。

### 3.1 モデル化

プライバシ保護手法の定量評価を行うために、前章で導入した二つの過程のモデル化を行う。

#### 3.1.1 タスク処理モデル

タスク処理を確率分布を用いてモデル化する。つまり、インスタンス $I$ と成果物 $R$ をそれぞれ $\mathcal{I}, \mathcal{R}$ 上の確率変数とし、タスク処理モデルを、条件付き確率分布 $p_t(R \mid I)$ として定義する。また、インスタンス $I$ に対するタスクの実行を、$p_t(R \mid I)$ からのサンプリングとしてモデル化する。成果物の品質に関す

**プロトコル 1** タスク処理プロトコル

1: 依頼者は、インスタンス $I$ を用いてタスクを依頼。
2: ワーカーは、$p_t(R \mid I)$ から成果物 $R$ をサンプリング。
3: ワーカーは、成果物 $R$ を依頼者に送信。
4: ワーカーは、$p_p(S \mid I)$ から秘匿情報 $S$ をサンプリング。

る議論を反映させるために、タスク実行を確率分布からのサンプリングとしてモデル化した。

#### 3.1.2 プライバシ侵害モデル

プライバシ侵害過程も、秘匿情報を成果物とみなし、タスク処理過程と同様にモデル化する。秘匿情報 $S$ を $\mathcal{S}$ 上の確率変数とし、プライバシ侵害モデルを条件付き確率分布 $p_p(S \mid I)$ として定義する。また、インスタンス $I$ に対するプライバシ侵害過程を、$p_p(S \mid I)$ からのサンプリングとしてモデル化する。

### 3.2 プロトコル

2. 章での二つの過程は、前節で導入されたモデルを用いてタスク処理プロトコル (プロトコル 1) のように記述される。タスク処理プロトコルでは、インスタンスをそのままワーカーに渡すため、プライバシ侵害が起こることに注意する。

タスク処理プロトコルから派生したプライバシ保護クラウドソーシングプロトコルでは、確率変数 $R, S, I$ を共有しているが、依頼者は $p'_t(R \mid I)$ からのサンプルを取得し、ワーカーは $p'_p(S \mid I)$ からのサンプルを取得するとする。プライバシ保護機構があるため、プロトコルに登場する確率分布が異なる。

### 3.3 評価指標

プライバシ保護クラウドソーシングプロトコルの汎用的な評価指標を提案する。プロトコルには主に次の二つ要件がある。まず、秘匿情報がプライバシ侵害モデル $p'_p(S \mid I)$ から漏洩しないことが必要である。また、プライバシ保護プロトコルで得られる成果物は、出来る限りタスク処理プロトコルでの成果物と近いことが必要である。この二つの面を評価するために、二つの評価手法を導入する。

まず、成果物の品質を評価する指標として、タスク情報損失を定義 1 のように定義する。この指標は、タスク処理プロトコルの代わりにプライバシ保護プロトコルを用いて成果物を得た場合の情報損失である。

**定義 1** (タスク情報損失)**.** タスク処理プロトコルのタスク処理モデル $p_t(R \mid I)$ と、プライバシ保護プロトコルのタスク処理モデル $p'_t(R \mid I)$ が与えられた元で、タスク情報損失を

$$L_t(p'_t) := \mathbb{E}_{p(I)}[\mathrm{KL}(p_t(R \mid I) \parallel p'_t(R \mid I))],$$

と定義する。ここで $\mathrm{KL}(p \parallel q)$ を、$p$ から測った $q$ の KL 情報量とし、$p(I)$ を $\mathcal{I}$ 上の確率分布とした。

次に、プライバシ保護能力を評価する指標としてプライバシ情報利得を、定義 2 のようにインスタンスと秘匿情報間の相互情報量を用いて定義する。この指標は、プライバシ保護データ出版の分野での無情報原理 [Li 07] を元にしている。

**定義 2** (プライバシ情報利得)**.** プライバシ保護プロトコルのプライバシ侵害モデル $p'_p(S \mid I)$ が与えられた元で、プライバシ情報利得を

$$L_p(p'_p) := \mathbb{E}_{p(I)}[\mathrm{KL}(p'_p(S \mid I) \parallel p'_p(S))],$$

と定義する。

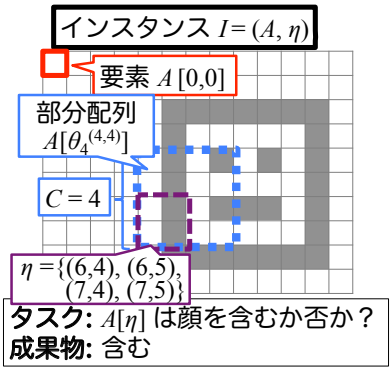


図 1: 提案プロトコルでの用語とその適用例。

タスク情報損失は、プライバシ保護プロトコルでの成果物の品質とタスク処理プロトコルでの品質の近さを反映している。また、インスタンスと秘匿情報が独立に近いほどプライバシ情報利得は小さくなる。インスタンスと秘匿情報が独立に近ければ、インスタンスから秘匿情報を推定することが困難になる。

提案指標の利点は、広い適用可能性にある。成果物や秘匿情報の定義に関わらず、3.1 節で導入したモデルが定義できれば提案指標が計算可能となる。また、正解データを利用せずに評価可能であることも利点の一つとして挙げられる。

#### 3.3.1 経験的推定

提案指標の実際の計算では、クラウドソーシングを用いてタスク処理モデルやプライバシ侵害モデルを経験的に推定し、それを用いて指標をプラグイン推定する。

推定は次のように行う。プロトコルを $M$ ($\geq 1$) 回繰り返し、$p_t(R \mid I)$ から $M$ 個のサンプル $\{R^{(m)}\}_{m \in \mathbb{Z}_M}$ が得られたとする。$p_t(R \mid I)$ の経験分布を

$$\hat{p}_t(R = r \mid I) \propto |\{m \in \mathbb{Z}_M \mid R^{(m)} = r\}| + \tau \ (\forall r \in \mathcal{R}),$$

と加法的平滑化を用いて推定する。ここで $\tau$ ($> 0$) を平滑化パラメタとし、$\mathbb{Z}_M := \{0, 1, \ldots, M-1\}$ と定義した。プライバシ侵害モデルも同様にして推定する。

## 4. プライバシ保護手法

本章では、切抜型プロトコルというプライバシ保護手法を提案する。提案手法では、インスタンスを切り抜くことでプライバシ保護を試みる。プロトコルの詳細を紹介した後に、その性質や適用可能性について議論する。

### 4.1 タスクに関する仮定

インスタンスと成果物に関する仮定を置く。まず、インスタンスは $D$ 次元配列 $A$ と配列のインデックス集合 $\eta$ から成るとし、それを $I = (A, \eta)$ と書く。$A$ を配列、$\eta$ を目標窓と呼ぶ。また、インスタンス $I$ の成果物を、配列 $A$ を $\eta$ で切り取った部分配列に対するラベルとする。配列 $A$ の $(h_1, \ldots, h_D)$ のインデックスの要素を $A[h_1, \ldots, h_D]$ と書き、インデックスの集合 $\eta$ で切り取った配列に対しても同じ表記 $A[\eta]$ を用いる。この表記を用いると、成果物 $R$ は、部分配列 $A[\eta]$ に対するラベルとなる。人間の頭部検出タスクの場合、配列は画像に相当し、目標窓は画像の特定の領域に相当する。図 1 では、紫破線の四角が目標窓 $\eta$ であり、成果物は、$A[\eta]$ に頭部が含まれるか否かのラベルとなる。目標窓を画像全体に動かすことで、画像全体のアノテーションが得られることに注意する。

**プロトコル 2** 切抜型プロトコル

1: 依頼者は、インスタンス $\{\phi_C(I;\theta)\}_{\theta\in\Theta_C}$ を用いて $|\Theta_C|$ 個のタスクを依頼。
2: **for** $\theta \in \Theta_C$ **do**
3: ワーカーは、$p_t(R \mid \phi_C(I;\theta))$ から成果物 $R^{(\theta)}$ をサンプリング。
4: ワーカーは、成果物 $R^{(\theta)}$ を依頼者に送信。
5: ワーカーは、秘匿情報を $p_p(S \mid \phi_C(I;\theta))$ から秘匿情報 $S$ をサンプリング。
6: **end for**
7: 依頼者は、サンプル集合 $\{R^{(\theta)}\}_{\theta\in\Theta_C}$ を $R$ のサンプルと見なす。

### 4.2 プロトコル

提案プロトコルでは、切抜関数（定義 3）を用いる。切り抜かれたインスタンスを用いてタスクを依頼することで、ワーカーが秘匿情報を抽出することを防止する。

**定義 3** (切抜関数)**.** インスタンス $I = (A,\eta)$ が与えられた元で、$C$ を偶数、$\theta_C^{(h_1,\ldots,h_D)} := \{(h_1+k_1,\ldots,h_D+k_D) \mid k_1,\ldots,k_D \in \mathbb{Z}_C\}$ を、一辺 $C$ の超立方体インデックス集合とし、$\Theta_C = \{\theta_C^{(h_1,\ldots,h_D)} \mid h_1,\ldots,h_D = 0, C/2, C, 3C/2,\ldots, \theta_C^{(h_1,\ldots,h_D)} \supseteq \eta\}$ と定義する。各切抜窓 $\theta\in\Theta_C$ に対して、切抜関数 $\phi_C(I;\theta)$ は $\phi_C(I;\theta) := (A[\theta],\eta).$ と定義される。$I[\theta] := (A[\theta],\eta)$ を部分インスタンスと呼ぶ。

ここで、$C$ は成果物の品質とプライバシ保護能力とのトレードオフを調整するパラメタである。$C$ を大きくすることで、成果物の品質は高くなるがプライバシ保護能力は低くなる。図 1 で、青点線の四角で書かれたものは、切抜窓 $\theta_4^{(4,4)}$ が 2 次元配列に適用されたものを示す。

提案手法でのプロトコルは、プロトコル 2 のように書ける。異なる $\theta\in\Theta_C$ に対して、各ワーカーは $A[\eta]$ のラベルを $p_t(R \mid I[\theta])$ からサンプリングする。切抜窓 $\theta\in\Theta_C$ は、目標窓 $\eta$ を含んでいるため、このサンプリングは実行可能であることに注意する。$p_t(R \mid I[\theta])$ は、$\theta$ に依存して互いに異なる分布となるが、これらの分布から得られるサンプルをすべて $R$ のサンプルとみなすというヒューリスティックを用いた。

### 4.3 定性的性質

切抜型プロトコルの性質を、成果物の品質とプライバシ保護能力の面から定性的に考察する。切抜型プロトコルでの成果物の品質は、タスクの局所性（ここでは成果物が確率的に依存するインスタンスの要素数）に依存する。一つの極端な例として、成果物 $R$、つまり $A[\eta]$ に対するラベルが $A[\eta]$ にのみ依存する場合を考える。この場合、切抜型プロトコルを適用しても、成果物の品質は下がらない。一方、成果物 $R$ が $A$ 全体に依存する場合、切抜型プロトコルを適用すると成果物の品質は大きく低下すると考えられる。インスタンスを切り抜くことで、ラベルを付与するのに重要な情報が欠けてしまうからである。以上の観察により、切抜型プロトコルは局所的なタスクに有用であると考えられる。同様の議論により、プライバシ保護能力の面からの性質も考察される。プライバシ侵害を防ぐためには、プライバシ侵害は局所的でないことが必要となる。以上より、切抜型プロトコルは、局所的なタスクと大局的なプライバシ定義の組に適する手法であると考えられる。

## 5. 実験

提案指標を用いて、異なる切抜窓の大きさ $C$ での切抜型プロトコルの性能を実験的に評価した。

### 5.1 実験対象

データセットは、Stanford 40 Action Dataset [Yao 11] を用いた。このデータセットでは、各画像中の人間の行動ラベルが与えられている。40 クラスある行動ラベルのうち、料理、魚釣り、ランニング、フリスビー、テレビ視聴、馬への餌やり、ギター演奏、携帯操作、PC 使用、ノート記述の 10 クラスを選択し、各行動につき 50 枚の画像を選択し、合計で 500 枚の画像を使用した。この画像集合を $\mathcal{A}$ とおく。各画像は $500\times500$ 画素に正規化した。

タスクは、人間の頭部特定タスクを用いた。画像を $S\times S$ 画素[*1]のブロックに分割し、各ブロックが頭部を含むか否かのラベルを得ることを目標とする。切抜型プロトコルを適用するために、タスクを次のように変換する。画像を 2 次元配列 $A$ としてみなし、$S\times S$ 画素の各ブロックのインデックス集合、つまり $H := \{\theta_S^{(h_1 S, h_2 S)} \mid h_1,h_2\in\mathbb{Z}_+\}$ の各要素[*2]を目標窓 $\eta$ として定義する。部分配列 $A[\eta]$ $(\eta\in H)$ に対するラベルを成果物としてみなす。ゆえに、$\{(A,\eta)\}_{\eta\in H}$ を用いてタスクを依頼することで、各画像全体に対するアノテーションを得るという目標が達成される。図 1 を例に説明すると、$A[\theta_4^{(4,4)}]$（青点線の四角）が与えられて、$A[\eta]$（紫破線の四角）に頭部が含まれるか否かを判断するタスクを切抜型プロトコルで実行する。

プライバシ定義に関しては、画像中の人間とその行動が結びつくことをプライバシ侵害として定義した。人間は頭部で識別可能と仮定すると、頭部を含む部分インスタンスからその人間の行動を抽出できないならばプライバシが保護されると定義した。例えば、女性がランニングをしている画像から切り抜いた頭部画像をワーカーに見せて、行動ラベルを推定できなければプライバシが保護されている。

### 5.2 タスク性能評価

まず切抜型プロトコルでのタスク性能に関する評価を行う。

#### 5.2.1 実験設定

本実験では、切抜型プロトコルにおけるタスク性能と切抜窓の大きさ $C$ との関係を調べた。切抜窓の大きさを $C=50$ から $C=300$ まで 50 ごとに増やして繰り返し実験を行った。各切抜窓の大きさ $C$ に関して、全インスタンス $\{I=(A,\eta)\}_{\eta\in H, A\in\mathcal{A}}$ に対するラベルを取得し、そのラベルを用いてタスク情報損失を計算した。タスク実行におけるコストを削減するために、二つのアイデアを用いた。まず、異なる画像から得られた複数の部分配列を図 2 のように組み合わせて、その画像を用いてタスクを依頼した（組み合わせて得られた画像を結合画像と呼ぶ）。結合画像の大きさがおおまかに $500\times500$ 画素となるように部分配列を組み合わせた。次に、結合画像の $S\times S$ 画素のブロックすべてのラベル付けを 1 人のワーカーに依頼した。結合画像 1 枚につき 1 人のワーカーを割り当てたため、$S\times S$ 画素のブロックのほとんどは 4 人のワーカー[*3]によってラベルが付与された。結合画像 1 枚の処理につき 0.5 円支払った。

*1 本論文では $S=25$ 画素と固定した。
*2 $A[\eta]$ を定義できる $\eta\in H$ のみを考える。
*3 インスタンス $(A,\theta_{C/2}^{(C/2,C/2)})$ を用いて切抜型プロトコルを実行すると、目標窓を含む切抜窓として $\Theta_C = \{\theta_C^{(0,0)},\theta_C^{(C/2,0)},\theta_C^{(0,C/2)},\theta_C^{(C/2,C/2)}\}$ が使われ、$A[\theta_{C/2}^{(C/2,C/2)}]$ に対して 4 人のワーカーがラベルを与えることになる。

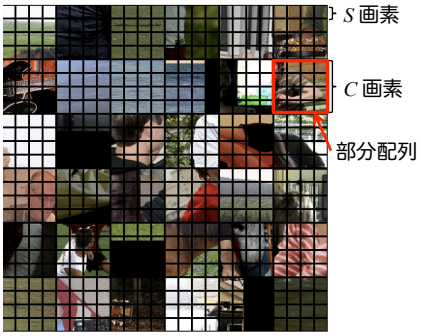


図 2: 部分配列を組み合わせることで、切抜型プロトコルの実行コスト削減を試みた。

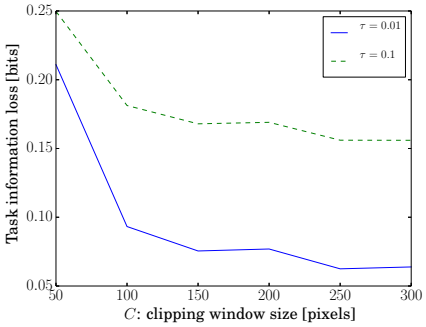


図 3: 異なる切抜窓の大きさでのタスク情報損失。

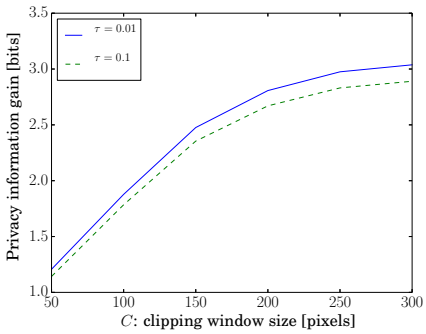


図 4: 異なる切抜窓の大きさでのプライバシ情報利得。

また、タスク処理プロトコルについても切抜窓の大きさ $C$ を変化させて、$\{I = (A, \eta)\}_{\eta \in H, A \in \mathcal{A}}$ に対する成果物のサンプルを得た。タスク実行におけるコストを削減するために、切抜型プロトコルで用いたアイデアの二つ目を用いた。つまり、画像 $A$ の $S \times S$ 画素のブロックすべてのラベル付けを 1 人のワーカーに依頼した。画像 1 枚につき 1 人のワーカーを割り当て、画像 1 枚の処理につき 1 円支払った。

各切抜窓の大きさ $C$ について、全成果物を用いてタスク情報損失を経験的に推定することで、切抜型プロトコルにおけるタスク性能を測った。平滑化パラメタは $\tau = 0.1$ と 0.01 の二つを用いた。

#### 5.2.2 実験結果

実験結果を図 3 に示す。平滑化パラメタを変えるとタスク情報損失の値は変わるが、値の変化の傾向は同じである。切抜窓の大きさ $C$ が 100 画素以上の場合には、タスク情報損失がほとんど変化しないことが読み取れる。すなわち、本実験設定では $C$ を 100 画素以上に設定することで、タスク処理プロトコルと同等のタスク性能を達成できる。

### 5.3 プライバシ保護能力評価

次に切抜型プロトコルでのプライバシ保護性能に関する評価を行う。

#### 5.3.1 実験設定

本実験では、切抜型プロトコルにおけるプライバシ保護性能と切抜窓の大きさ $C$ との関係を調べた。切抜窓の大きさを $C = 50$ から 300 まで 50 ごとに増加させ繰り返し実験を行った。

プライバシ侵害過程を模倣するために、十択の質問を用いた。ワーカーに、人間の頭部を含むような大きさ $C \times C$ 画素の画像 1 枚と、10 クラスの行動ラベルを与え、画像中の人間の行動を選択するように依頼した。前実験で頭部を含むと判断された $C \times C$ 画素の画像から、各行動ラベルにつき 25 枚選択し、計 250 枚の画像を実験に用いた。各画像につき 50 人のワーカーを割り当て、1 枚の画像の処理につき 0.2 円支払った。

各切抜窓の大きさ $C$ について、全成果物を用いてプライバシ情報利得を経験的に推定することで、プライバシ保護性能を測った。平滑化パラメタは $\tau = 0.1$ と 0.01 の二つを用いた。

#### 5.3.2 実験結果

図 4 に実験結果を示す。平滑化パラメタによらず、プライバシ情報利得の値は同じ傾向を示した。プライバシ情報利得は、$C$ を増加させるに従って単調に増加した。この結果は、大きい範囲を写した画像ほどプライバシ侵害が容易であるという直感と合致する。タスク情報損失と異なる点は、その変化速度である。タスク情報損失は、$C = 100$ 付近でほとんど変化しなくなるのに対し、プライバシ情報利得は常に増加している。タスクとしての性質が異なるためにこのような違いが生じると考えられる。つまり、頭部検出タスクは局所的なタスクである一方で、プライバシ侵害を行うためには広い範囲の画像を見る必要があったため、このような差異が生じる。このような差異により、本実験設定ではタスクを高品質に実行できると同時に、プライバシを保護することができると結論づけられる。

## 6. 結論

本論文では、クラウドソーシングにおけるプライバシ保護手法の評価方法の提案及び、プライバシ保護手法の提案を行った。提案手法では、インスタンスを切り抜いてワーカーへの情報流出を制限してプライバシ保護を行う。クラウドソーシングを用いた提案手法の評価により、局所的に実行可能なタスクとインスタンスに大局的に依存するプライバシ定義の場合に、プライバシを保護しつつタスクを実行可能であると示された。

## 参考文献

[Lease 11] Lease, M.: On quality control and machine learning in crowdsourcing, in *Proceedings of the Third Human Computation Workshop*, pp. 97–102 (2011)

[Li 07] Li, N., Li, T., and Venkatasubramanian, S.: t-closeness: Privacy beyond k-anonymity and l-diversity, in *Proceedings of 2007 IEEE 23rd International Conference on Data Engineering*, pp. 106–115 (2007)

[Little 11] Little, G. and Sun, Y.-A.: Human OCR: Insights from a complex human computation process, in *Proceedings of CHI 2011 Workshop on Crowdsourcing and Human Computation*, pp. 8–11 (2011)

[Varshney 12] Varshney, L. R.: Privacy and reliability in crowdsourcing service delivery, in *Proceedings of the 2012 Annual SRII Global Conference*, pp. 55–60 (2012)

[Yao 11] Yao, B., Jiang, X., Khosla, A., Lin, A. L., Guibas, L., and Fei-Fei, L.: Human action recognition by learning bases of action attributes and parts, in *Proceedings of 2011 IEEE International Conference on Computer Vision*, pp. 1331–1338 (2011)