

早稲田大学大学院 基幹理工学研究科

博士論文概要

論文題目

Fair-Trade Crowdsourcing:
Predicting the Working Times of Microtasks

公正取引クラウドソーシング：マイクロタスクの
作業時間推定

申請者

Susumu	SAITO
斎藤	奨

Department of Computer Science and Communications Engineering
Research on Perceptual Computing
情報理工・情報通信専攻 知覚情報システム研究

2019年11月

クラウドソーシングにおけるマイクロタスク作業時間推定システムを提案する。本論文は、*i)* サードパーティツールやオンラインコミュニティの利用を通じたクラウドワーカーの労働方法についての定量的な分析、*ii)* マイクロタスク作業時間推定システムのプロトタイピング、*iii)* 推定エラーに対するワーカーの感覚尺度に基づいたモデル最適化および評価、の3つから主に構成される。

クラウドソーシングのプラットフォームにおいて、多くのワーカーの賃金が低いことが問題となっている。この主な原因として、リクエスタ・ワーカー間の参入障壁の高さに大きな差が生じていることが挙げられる。リクエスタ側では、マイクロタスクのUI作成・説明文などのメタ情報の付与・報酬額の設定などを行う柔軟な機能が与えられており、独自の手法や裁量をもとにマイクロタスクの依頼を比較的容易に行うことができる。その反面、ワーカー側ではそうした（一定のノイズを含みうる）一連のマイクロタスク情報から肌感覚で総合的に判断し、「割の良い」マイクロタスクを探し当てることが求められる。しかしこれは多くのワーカーにとって非常に困難であるため、ワーカーの低賃金問題に繋がるとされている。こうした背景を踏まえ、クラウドソーシングの持続性を保持していくためには、ワーカーの賃金を上げるための支援方法を考えていくことが必要不可欠であるといえる。これに対し申請者が最も効果的と考える方法は、マイクロタスクの完了に必要な時間（作業時間）の見積もりを行うことである。これをタスク単価と比較することで、タスクの割の良さを事前に計算し、次に行うマイクロタスクをより適切に選択可能となることが期待される。

本論文では、プラットフォーム上に投稿されたあらゆるマイクロタスクに適用可能な、機械学習のアプローチに基づくマイクロタスク作業時間推定手法について検討する。まずツール・オンラインコミュニティの利用状況に関するアンケートをワーカーに対し実施することで本手法の重要性の裏付けをとった後、推定手法の構築における鍵となるデータ収集方法や特徴量設計を中心に検討する。さらに、回帰による秒単位での作業時間の推定結果が一定の誤差を伴うとき、推定値および実測値の秒数の絶対的な大きさとそれらの距離との関係から、誤差がワーカーにとってどのような意味合いを持つかという点に着目し、これを作業時間に対するワーカーの心理尺度として定量化する方法を検討する。この尺度を作業時間推定モデルの設計に適用することで、モデルをよりワーカーにとって推定結果が有益となるように最適化し、かつその基準においてシステムを評価することを試みる。

本論文は5章から構成される。以下に各章の概要を示す。

第1章では、まずワーカーの低賃金問題に触れ、マイクロタスクの作業時間を推定する重要性について述べる。続けて、それを実現するための問題点を挙げながら、本研究における研究課題を3つに分けて述べる。

第2章では、サードパーティツールおよびオンラインコミュニティの利用を通じたクラウドワーカーの労働方策の分析を行った。ここでの目的は、ワーカーの労働環境において何が可能・不可能であるかを理解し、マイクロタスク作業時間推定の必要性を浮き彫りにすることで

ある。先行研究ではワーカーが賃金向上を目的として様々な種類のツールやコミュニティ等の補助機能を利用していることが言及されてきた。しかし、それらの利用により何が可能となり、何がそうでないのか、こういった機能性が誰に評価されているのか、等についてはまだ整理されていない。こうした状況では、ワーカーの需要が正確に把握できず、効果的なワーカー支援が行えない可能性がある。そこで、2019年現在最も規模の大きいプラットフォームである Amazon Mechanical Turk においてワーカーへツール・コミュニティ利用についての聞き取り調査を行った。ワーカー360人の回答結果から、ワーカーとしての収入の上位10%、上位50%、下位50%、のグループに分けて分析を行ったところ、収入の多いワーカーほど頻繁に「マイクロタスク時給」を計算するツールやコミュニティを用いていることが分かった。これは暗にワーカーがマイクロタスクの作業時間を意識しているということの意味しており、作業時間推定の必要性が示される結果となった。

第3章では、マイクロタスク作業時間推定システム「TurkScanner」について述べる。既存のサードパーティツールやオンラインコミュニティにおいて提供される、マイクロタスクの作業時間または時給の目安（＝平均値）を提示する機能は、そのマイクロタスクを他のワーカーが既に行っており、その際の作業時間が履歴としてシステムに蓄積されていることが動作の必要条件となる。この場合、最低一人のワーカーが作業履歴をシステムに提供しなければ作業時間の提示をすることができないため、マイクロタスクの適用範囲が狭いという問題がある。これを踏まえ、提供情報の有無にかかわらずあらゆるマイクロタスクの作業時間を推定可能なシステムを構築した。入力をワーカーに公開されたマイクロタスク関連情報とし、出力を秒単位の作業時間として回帰推定する機械学習モデルの構築を目指し、具体的には次の3点について検討した。一点目に、作業時間の測定方法の定義を行った。ワーカーがマイクロタスクを行う際には、途中で休憩のため席を離れる、複数のマイクロタスクを同時に進める、外部サイトを閲覧する（マイクロタスクに関連する場合としない場合のどちらもあり得る）、などの様々な振る舞いが想定されるため、あらゆる場合にも正しく作業時間を測定可能とする画一的な方法は存在しないと考えた。そこで、想定されるワーカーの振る舞いを少なくともいずれか1つでカバー可能となるような2種類の自動計測方法と2種類の（ワーカー操作による）手動計測手法を策定した。二点目に、モデル学習および評価のためのデータ収集を行った。ワーカーが実際に作業したマイクロタスクの情報を実際の作業時間とともに収集するにはワーカーの協力が必要不可欠だが、公開データセットや利用可能なデータ収集スクリプトは存在しない。そこで、独自にデータ収集用の Chrome 拡張機能を開発し、AMT のワーカーにインストールさせる方法を策定した。ワーカーが作業したマイクロタスクそれぞれにおいて、前述の測定方法に基づき作業時間をラベリングすることでデータを収集した結果、83人のワーカーの貢献により7303件のマイクロタスク作業データを獲得した。三点目に、回帰推定モデルの設計および評価を行った。まずシステム入力として、①マイクロタスク情報（公開されたメタ情報やHTML内の要素など）、②ワーカー情報（今までの作業履歴やプロフィールなど）、③リクエスト情報（外部サイトから取得可能な評判情報）、を含む計101次元の特徴量を設計した。モデルには、デ

ータセットのサイズや説明可能性を考慮して **Gradient Boosting Decision Tree** を採用した。クロスバリデーションによる評価結果を混同行列のヒートマップとして可視化することで、およそのマイクロタスク作業時間のオーダーを推定する（短い作業時間を短く、長い作業時間を長く推定する）ことが可能であることが分かった。

第4章では、ワーカーの作業時間推定エラーに対する感覚の定量化手法「**CrowdSense**」について述べる。前章で設計したモデルでは、それぞれのマイクロタスクの作業時間の推定結果がワーカーにとってどのような意味合いを持つかを議論することができない。例えば、完了に30秒かかるマイクロタスクに対する作業時間推定ではどれだけの誤差がワーカーに許容され得るか、あるいは100秒や500秒かかるマイクロタスクではどうか、といった法則は自明ではなく、単純な秒単位での距離や割合では測ることができないものである。しかし実際には、ワーカーが割の良いマイクロタスクを選択する支援という目的を踏まえると、推定モデルはこうした基準で最適化され、また評価されるべきである。そこで、マイクロタスク作業時間推定の誤差に対するワーカーの感覚量を実証的に定量化し、これを用いて構築したマイクロタスク作業時間推定モデルの目的関数および評価関数の設計と評価を行った。まず、あるマイクロタスク作業時間[秒]に対する許容可能な最大誤差[秒]の関係を定式化するためにAMTのワーカーに対するアンケートを実施した。マイクロタスク作業時間推定が可能な架空のワーカー支援ツールの利用時を想定した上で、作業時間の予測値と実測値を被験者に提示し、その推定誤差が許容可能かどうかの聞き取り調査を行った。計918通りの異なる提示パターンにそれぞれについておよそ100回答ずつワーカーから収集し、回答の境界に対して曲線近似を行うことで評価関数を導出した。さらに、ここで得られた近似関数の逆関数を積分することで、任意の作業時間に対するワーカーの「感覚作業時間」の関係を定式化した。これを目的関数における損失の計算時に用い、ワーカーの感覚上の推定誤差を最小とするようなモデル最適化方法を策定した。以上の設計により **TurkScanner** のモデルを再学習し評価を行った結果、①テストデータの約73%においてワーカーの許容誤差範囲内の推定が可能であること、②ベースライン手法と比べてより広い範囲の作業時間スケールに渡って正確な推定が可能であること、の2点においてモデル改善を達成した。

第5章では、本研究の結論を述べる。提案したマイクロタスク作業時間推定システムがクラウドソーシング市場の透明性を向上させ得ること、人間による主観的な感覚に基づきワーカー支援を行う手法を実現したことなどを中心に、クラウドソーシング研究における貢献を考察する。続けて、提案手法の他の用途への適用可能性および今後の課題として、ワーカーの労働の支援だけでなく、リクエストのマイクロタスク作成の支援や、リアルタイムなクラウドソーシングにも適用でき得ることなどを議論する。

早稲田大学 博士（工学） 学位申請 研究業績書

氏名 齋藤 奨 印

(2020年 2月 現在)

種 類 別	題名, 発表・発行掲載誌名, 発表・発行年月, 連名者 (申請者含む)
論文誌 【国際】	○Predicting the Working Time of Microtasks Based On Workers' Perception To Prediction Errors, Human Computation Journal, vol. 6, pp.192-219, October 2019, Susumu Saito , Chun-Wei Chiang, Saiph Savage, Teppei Nakano, Tetsunori Kobayashi, and Jeffrey P. Bigham.
学会論文 【国際】	○MicroLapse: Measuring Workers' Leniency To Prediction Errors of Microtasks' Working Times, In Proc. of CSCW2019, pp. 1053-1057, November 2019, Susumu Saito , Teppei Nakano, Tetsunori Kobayashi, and Jeffrey P. Bigham. Two-stage calving prediction system: Exploiting state-based information relevant to calving signs in Japanese black beef cows, In Proc. of ECPLF 2019, pp.670-676, August 2019, Ryosuke Hyodo, Saki Yasuda, Yusuke Okimoto, Susumu Saito , Teppei Nakano, Makoto Akanabe, Tetsunori Kobayashi, and Tetsuji Ogawa. Calving prediction from video: Exploiting behavioral information relevant to calving signs in Japanese black beef cows, In Proc. of ECPLF 2019, pp.663-669, August 2019, Kazuma Sugawara, Susumu Saito , Teppei Nakano, Makoto Akabane, Tetsunori Kobayashi, and Tetsuji Ogawa. ○TurkScanner: Predicting the Hourly Wage of Microtasks, In Proc. of The Web Conference 2019, pp.3187-3193, May 2019, Susumu Saito , Chun-Wei Chiang, Saiph Savage, Teppei Nakano, Tetsunori Kobayashi, and Jeffrey P. Bigham. ○Striving to Earn More: A Survey of Work Strategies and Tool Use Among Crowd Worker, In Proc. of HCOMP2018, pp.70-78, July 2018, Susumu Saito* , Toni Kaplan*, Kotaro Hara, and Jeffrey P. Bigham. (*equal contribution) Waseda at TRECVID 2016: Ad-hoc Video Search, Notebook paper of the TRECVID 2016, pp.110-111, November 2016, Kazuya Ueki, Kotaro Kikuchi, Susumu Saito , and Tetsunori Kobayashi. ○ Evaluation of Collaborative Video Surveillance Platform: Prototype Development of Abandoned Object Detection, In Proc. of ICDCS2016, pp.172-177, September 2016, Susumu Saito , Teppei Nakano, Makoto Akabane, and Tetsunori Kobayashi. ○Towards a framework for collaborative video surveillance system using crowdsourcing, In Proc. of CSCW2016 (Poster), pp.393-396, February 2016, Susumu Saito , Teppei Nakano, and Tetsunori Kobayashi.
学会論文 【国内】	○TurkScanner: マイクロタスクの時給推定, 人工知能学会 第33回全国大会, 1K3-J-4-02, 2019年6月, 齋藤奨 , 中野鐵兵, 小林哲則. 肉牛の分娩検知システムにおけるクラウドソーシングを用いた誤通報の抑制, 人工知能学会 第33回全国大会, 1F4-OS-17b-05, 2019年6月, 沖本祐典, 齋藤奨 , 中野鐵兵, 赤羽誠, 小林哲則, 小川哲司. 画像から得られる牛の身体情報に基づく分娩予兆検知, 人工知能学会 第33回全国大会, 4H3-J-13-02, 2019年6月, 兵頭亮介, 齋藤奨 , 中野鐵兵, 赤羽誠, 小林哲則, 小川哲司.

早稲田大学 博士（工学） 学位申請 研究業績書

種 類 別	題名, 発表・発行掲載誌名, 発表・発行年月, 連名者(申請者含む)
学会論文 【国内】	<p>映像による肉牛分娩開始検知システムの早期運用のためのクラウドソーシングを用いた誤通報抑制, 電子情報通信学会 パターン認識・メディア理解研究会 (PRMU), Vol.2, pp.1-6, 2019年5月, 沖本祐典, 齋藤奨, 中野鐵兵, 赤羽誠, 小林哲則, 小川哲司.</p> <p>牛の分娩予兆として映像から観測可能な状態の検知, 人工知能学会 第32回全国大会, 2G3-OS-10c-03, 2018年6月, 沖本祐典, 菅原一真, 齋藤奨, 中野鐵兵, 赤羽誠, 小林哲則, 小川哲司.</p> <p>○クラウドソーシングを用いた協調型ビデオ監視システムフレームワークの提案, 第78回情報処理学会全国大会, Vol.1, pp.77-78, 2016年3月, 齋藤奨, 中野鐵兵, 小林哲則.</p> <p>○クラウドソーシングを用いた協調型ビデオ監視システムのプロトタイプ設計, 2016年電子情報通信学会全国大会, pp.252, 2016年3月, 齋藤奨, 中野鐵兵, 小林哲則.</p> <p>学習者の知識補完のためのオンライン講義コンテンツ連携フレームワーク, 情報処理学会 コンピュータと教育研究会 132回研究発表会, Vol.19, pp.1-6, 2015年12月, 齋藤奨, 小林哲則.</p>
特許 【国内】	状態監視システム, No. 2016-165131, 2016年8月.
その他	<p>リーディング大学院 実体情報学博士プログラム, 2015年4月～</p> <p>JSPS 学振特別研究員(DC2), 2019年4月～</p>