

早稲田大学大学院 基幹理工学研究科

博士論文審査報告書

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

2020年2月

インターネット上の不特定多数の潜在的ワーカーに、公募形式で業務（タスク）を外注し労働力を得る「クラウドソーシング」が、タスク遂行の一般的形態のひとつとして浸透してきた。タスクを発注するリクエスタにとっては、労働力を迅速に得ることができ、また、タスクを請負うワーカーにとっては、いつでも報酬を得る機会を得られることから、双方にとって有益な労働力の供受給プラットフォームとなることが期待されている。

しかし、現状、クラウドソーシングにおけるリクエスタとワーカーの関係は対等とは言い難い。リクエスタは、タスク発注の際、独自裁量で値付けができる一方で、ワーカーにはこの値付けの正当性を確認する術がない。実行してみるまでは労働に見合う報酬が得られるかどうかは不明であって、結果としてワーカーは低賃金で働かされる傾向にある。このような不健全な実態を改善するために、リクエスタとワーカーの間の取引を公正なものとするのが求められている。

本研究では、タスクの完了に必要な時間を見積もる技術を開発し、見積もった時間をタスク選定時のワーカーに提示することで、公正な取引を実現することを提案している。ワーカーは、大凡の作業時間を知ることによって、当該タスクの「割の良さ」を知ることができ、納得感を持った上でタスクを請負う判断ができるとしている。

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章は、本論文の序章であり、まずワーカーの低賃金問題に触れ、タスクの作業時間を推定することの重要性について述べている。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3つの研究課題、すなわち、i)クラウドワーカーの作業傾向に関する分析、ii)タスク完了時間推定システムのプロトタイプング、iii)作業推定誤差におけるワーカーの心理尺度の導入、を定義している。

第2章では、クラウドワーカーを収入でクラス分けした上で、各クラスのワーカー毎に、サードパーティーツールおよびオンラインコミュニティ等の利用に関する作業傾向の分析を行っている。従来、ワーカーが様々な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知られているが、どういう層のワーカーがどういう使い方をしているかは必ずしも明らかでない。ここでは、現在最も規模の大きいプラットフォームである Amazon Mechanical Turk (AMT) のユーザを対象として、聞き取り調査を実施している。ワーカー 360 人の回答結果から、ワーカーとしての収入の多いワーカーは頻繁に「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を示し、この層のワーカーがタスクの作業時間を意識しながら作業をしていることを明らかにしている。またこのことをもって、作業時間を意識させることで、ワーカーの収入を増やす可能性があることを示唆したとしている。

第3章では、タスク完了時間推定システム「TurkScanner」について述べている。従来から、サードパーティツールやオンラインコミュニティを用いてタスク完了時間、または時給の目安（＝平均値）を知ることは可能であるが、これはそのタスクを他のワーカが既に実行しており、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条件となる。そこで、これらの情報提供の有無にかかわらず、あらゆるタスクの作業時間を推定可能なシステムを構築している。①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次元の特徴量を用いて、秒単位でタスク完了時間を回帰推定するモデルを、機械学習を導入して構築している。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 を採用している。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、おおよそのタスク作業時間のオーダーを推定することが可能であることを示している。

第4章では、ワーカのタスク完了時間に関する心理尺度の構成法「CrowdSense」について述べている。前章で述べた方法では、タスク完了時間は物理量である「秒」を単位して最適化が行われた。しかし、ワーカにとっての心理的な許容度をもとに推定誤差を考えると、これを物理量で最適化することは問題がある。例えば、同じ5秒の推定誤りも10秒を5秒誤るのと、100秒を5秒誤るのでは誤差の大きさの感じ方は異なるし、20秒を5秒多く25秒と誤るのと、5秒少なく誤るのでもまた感じ方が異なる。要するに、時間軸上の各点で、誤差に対する感度が等価になるよう補正を加えた心理時間尺度上で、最適化は行われるべきである。ここでは、聴覚心理尺度の構成法に習って、タスク完了時間の心理尺度を実験的に構成することを試みている。すなわち、ワーカに対しアンケートを行って、タスク完了時間推定の誤差に対する弁別閾（ワーカはどの程度まで誤差を許容するかの閾値）を実験的に求め、弁別閾の逆関数の積分値としてタスク完了時間の心理尺度を構成することを試みている。これをモデリングにおける損失計算に用いることで、①テストデータの約73%においてワーカ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点を確認している。

第5章は、結論であり、本研究での提案を概観するとともに、その価値を総括している。また、今後の展望として、本研究成果のリクエストの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への適用可能性について議論している。

以上、これを要するに、本論文は、健全なクラウドソーシングの運用を目指す上では、リクエストとワーカ間の取引が公正であることの重要性を指摘

し、これを実現するための重要要素技術として、ワーカの心理特性に従うタスク完了時間を推定する手法を実現したものである。提案は、クラウドソーシングの持続性の担保に資するとともに、リアルタイムクラウドへの応用の可能性を持つなど、その工学的意義は高い。よって、本論文は、博士（工学）（早稲田大学）の学位論文として相応しいものと認める。

2020年2月

審査員

主査	早稲田大学 教授	工学博士（早稲田大学）	小林 哲則
副査	早稲田大学 教授	工学博士（慶応義塾大学）	中島 達夫
	早稲田大学 教授	博士（工学）（早稲田大学）	小川 哲司
	筑波大学 准教授	博士（情報理工学）（東京大学）	馬場 雪乃