

クラウドワーカを利用したメカニズム設計のための一考察

Mechanism Design by Utilizing Crowdworkers

櫻井 祐子 松田 昌史*2 小山 聡*3
Yuko Sakurai Masafumi Matsuda Satoshi Oyama

*1九州大学
Kyushu University

*2NTT コミュニケーション科学基礎研究所
NTT Communication Science Laboratories

*3北海道大学
Hokkaido University

We focus on the problem in applying mechanism design to crowdsourcing: crowdsourcing requesters are not specialized in mechanism design and it is difficult for them to choose an appropriate mechanism from a large variety of complicated mechanisms proposed by researchers. As a preliminary experiment for investigating acceptable crowdsourcing mechanisms, we perform a task that ask workers to propose a mechanism and to answer questions related to the value-intention scales and the just world scale. The result indicate that workers propose a mechanism according to their value-intention/just world scales.

1. はじめに

近年、ネットワーク環境の普及に伴い、インターネット上では様々なサービスが提供されている。特に、ここ数年で成長が著しいサービスの一つとして、Amazon Mechanical Turk に代表されるクラウドソーシングがある。クラウドソーシングは「群衆の知」の概念が根本にあり、多数の人々の知恵や力を利用することによって問題を解決するサービスである。クラウドソーシングのサービスを利用することで、世界中の人々に安価にタスクを実行させることが可能となっている。さらに、現在の計算機の能力では解決が難しいが、人間では容易に解決可能な問題などに対して、人間をプログラムの一部として機能させることで解決を目指すアイデアに基づく Human Computation など、今後の計算機科学における計算機と人間の融合に関する有望な研究領域になっている [Law 11]。

しかしながら、クラウドソーシングでは、安価に不特定多数のワーカを容易に働かせることが可能であるという利点がある一方で、得られる作業結果の品質が問題となっている。たとえば、画像のラベル付けなど、機械学習の訓練集合に用いるラベルをクラウドソーシングによって得る場合、大量のラベル付きデータが得られる一方で、ラベルの信頼性に関する問題が生じる。専門家等から得られるラベルとは異なり、不特定多数のワーカのラベルには、ワーカ自信の意図の有無に関わらず、ラベルに対して誤りが入る可能性がある。そこで、ワーカに対して適切なインセンティブを与えるメカニズム (ルール) を設計することによって、品質管理を実現する方法に関する研究が行われている [Liu 17]。

メカニズムデザインとは望ましい性質を満たすメカニズムを設計することである。メカニズムデザイナーがメカニズムを決める際、デザインの前提となる条件がいくつか存在する。例えば、クラウドソーシングでの一般的な金銭報酬設定では、大きく分けて定額か成果報酬かという2つの枠組みがある。また、成功報酬の与え方においても、非承認という形で罰するか追加報酬を与える等方法がある。このように様々な報酬の設定方法が考えられるが、クラウドソーシングのリクエスト及びワーカからはメカニズムデザインの専門家でないことを考慮すれば、複雑な報酬設定を含むメカニズムは受け入れられないと考える。

そこで、我々は、クラウドソーシングのリクエストが必ずしもメカニズムデザイナーの専門家ではないこと、メカニズムデザイナーの専門家が考案したメカニズムが必ずしも簡単ではないことに着目し、クラウドソーシングに適切なメカニズムデザインのフレームワークを提案することを目指す。この目標を達成するための初期検討として、本論文では、日本のクラウドソーシングサービスにおいて、リクエストとして予算制約がある下でどのようにメカニズムを考えるかというタスクを実施した。さらに、各ワーカが考案するメカニズムのフレームと、個人の価値志向性尺度の理論、経済、社会に関する尺度と正当世界尺度との関係性についても調べた。

2. クラウドソーシングを利用した実験

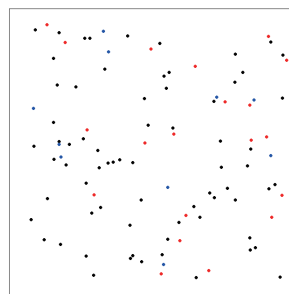


図 1: 各色の個数を数えるタスク

我々は日本のクラウドソーシングサービスのランサーズ (<http://www.lancers.jp/>) を利用してタスクを実施した。30名のワーカを集め、各ワーカには300円の報酬を払った。ワーカには図1の各色の個数を数えるタスクに対して、次のインストラクションのもとでメカニズムを考えて貰った。
ワーカへのインストラクション: あなたは、図に黒、赤、青の3つの点があくつ含まれるか、正確な個数を知りたい。1,000円の予算でクラウドソーシングサービスを利用してワーカから答えを得ようとする。このとき、何人のワーカを雇って、どのような報酬設定でどのように正解となる解を決めるかを記述してください。

連絡先: 櫻井 祐子, 九州大学, 福岡市西区元岡 744,
ysakurai@inf.kyushu-u.ac.jp

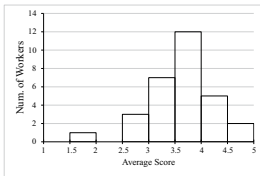


図 2: 理論

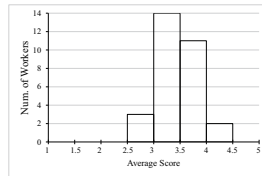


図 3: 経済

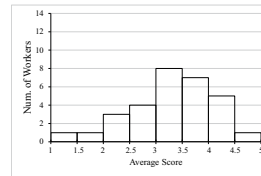


図 4: 社会

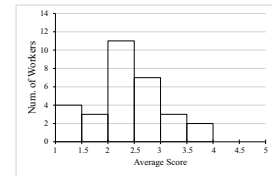


図 5: 正当世界

報酬	ワーカー数	理論	経済	社会	正当世界
定額	11	3.36	3.36	3.36	2.23
成果	17	3.64	3.50	3.26	2.30
従量	2	3.46	3.21	2.50	1.50

表 1: 各尺度の平均得点

ワーカーには、メカニズムを記述して貰った後、価値志向性尺度と世界正当性尺度を測るための質問にも回答させた。価値志向性尺度はシュプリンガー [Spranger 28] の提唱する 6 つの普遍的価値を個人がどの程度志向し、体験しているかを測定する尺度であり、酒井ら [Sakai 97] によって開発された。本実験では、6 つの価値志向性尺度のうち、理論、経済、社会を利用した。これら 3 つの尺度に関しては、12 項目ずつの質問を与えた。理論の価値尺度は理論が通じることや真理に対する傾向を測定する尺度、経済の価値尺度は時間や労力の節約に対する傾向を測定する尺度、社会の価値尺度は人の役に立つ行動をとる傾向を測定する尺度である。正当世界尺度は、正の投入には正の結果、負の投入には負の結果が伴うものとして世界を認知する傾向を測定する尺度である。本実験では、今野ら [Konno 98] によって開発された 4 項目の質問を利用した。それぞれの尺度は得点化（各価値をどの程度重要視するかの指標化）するために、5 段階の点数で答えさせた。

3. 実験結果

図 2 から図 5 に各尺度のワーカーの平均得点の分布を示す。各尺度の平均得点は、理論が 3.58、経済が 3.42、社会が 3.29、正当世界が 2.25 であった。さらに、表 1 に、30 名のワーカーにおいて報酬設定の違いによって、各尺度の平均得点にどのような傾向があるかを示す。各尺度での平均を算出し相関関係を調べた結果、経済と正当世界信念の相関はなく ($r = -0.0056$)、互いに独立であることより、この 2 つの指標に基づいてワーカーを分類した。2 つの尺度が共に平均以上のワーカーは 8 名存在した。そのうち、定額報酬が 1 名、成果報酬が 7 名であった。成果報酬を選んだ 7 名のワーカーが記述したメカニズムのうち、特徴的なものを載せる。

- 100 名のワーカーを雇用する。平均値を正解とする。各色の平均値との差が 10 個よりも大きい場合は非承認とする。
- 50 人のワーカーを雇用し、作業状況を動画で送ってもらう。多数決で解を決定する。作業状況が不適切なワーカーには報酬を与えない。
- 3 名のワーカーを雇用し、多数決で解を決定する。不正解者は非承認とする。3 人とも回答が違う場合は、全員非承認とし、再募集する。報酬は、正解者で均等割りで決定する。

また、興味深い結果として、成果報酬として、正解者には継続依頼を提示すると記載したワーカーが 2 名いた。この 2 名のワーカーは理論と経済の 2 つの尺度で平均以上であった。1 名のワーカーは経済尺度が最も高いワーカーであり、そのワーカーは社会尺度以外は全て平均以上であった。

実験の結果、1 回限りのゲームとして考えるワーカーがほとんどであり、非協力的なワーカーには非承認、すなわち罰することで協力を導こうとする考えに基づくメカニズムが多いことが分かった。

4. おわりに

今回、我々は、クラウドソーシングのリクエストが必ずしもメカニズムデザイナーの専門家ではないこと、メカニズムデザイナーの専門家が考案したメカニズムが必ずしも簡単ではないことに着目した。そこで、我々はクラウドソーシングワーカーに受け入れやすいメカニズムを提案するための初期実験として、クラウドソーシングワーカーにメカニズムを考えて貰うとともに価値志向尺度に関するアンケートに答えてもらうタスクを実施した。

実験結果として、ワーカーは価値志向尺度に応じたメカニズムを考案することができることが分かった。今後は、より複雑なタスクを対象に実験を行う予定である。

謝辞

JSPS 基盤研究 (B) (15H02751, 15H02782) の助成を受けました。深く感謝します。

参考文献

- [Konno 98] Konno, H. and Hori, H.: Effects of justice beliefs on injustice judgement, *Institute of Psychology, University of Tsukuba, Departmental Bulletin Paper*, Vol. 20, pp. 157–162 (1998), in Japanese
- [Law 11] Law, E. and Ahn, L. V.: *Human Computation*, Morgan & Claypool Publishers (2011)
- [Liu 17] Liu, Y. and Chen, Y.: Sequential Peer Prediction: Learning to Elicit Effort using Posted Prices, in *Proceedings of 31st AAAI Conference on Artificial Intelligence (AAAI-2017)* (2017), to appear
- [Sakai 97] Sakai, K. and Hisano, M.: Construction of value-intending mental act scale, *Educational Psychology Research*, Vol. 45, No. 4, pp. 388–395 (1997), in Japanese
- [Spranger 28] Spranger, E. and Pigors, P. J. W.: *Types of Men: The Psychology and Ethics of Personality* (1928)