

# クラウドソーシングにおけるインセンティブ設計への 多腕バンディットの適用

Applying Multi-armed Bandits for Incentive Design in Crowdsourcing

伊藤 聡比古<sup>\*1</sup>      松原 繁夫<sup>\*1</sup>  
Akihiko Itoh      Shigeo Matsubara

<sup>\*1</sup> 京都大学大学院 情報学研究科  
Graduate School of Informatics, Kyoto University

We have examined the possibility of modeling the incentive design in crowdsourcing as the Multi-armed bandits (MAB) model. So far, researchers studied the effect of a variety of incentives including financial and social ones. The effectiveness of these incentives, however, may differ from worker to worker or task to task, and thus finding an optimal incentive out of the whole incentive space is difficult. To overcome this difficulty, we consider a data-driven approach is promising instead of trying to find a universally effective incentive. More specifically, we consider applying MAB algorithms to the incentive design is effective. The MAB model is a theoretical model of exploration-exploitation tradeoffs in learning. To examine this possibility, we evaluated the effect of nine incentives on Mechanical Turk by using word puzzle tasks. Based on the obtained data, we further evaluated the performance of four MAB algorithms by simulation. The results show that applying the MAB algorithms are effective to obtain the appropriate incentive design.

## 1. はじめに

クラウドソーシングは不特定多数に問題解決を依頼するプロセスであり、近年、人工知能技術の適用対象として注目を集めている。請負者（ワーカーと呼ぶ）が計算機ではなく人であり、また、不特定多数への依頼であるため、成果物の品質に関して、作業態度に起因するばらつきが生じる恐れがある。よって、依頼者が高品質な成果物を得ようとすれば、ワーカーをどのように動機付けるかという点でのインセンティブ設計が重要となる。

クラウドソーシングにおけるインセンティブの与え方としては、タスクの社会的意義を強調したり、報酬を作業品質に連動させたりするなど様々なものが存在し、内因性動機／外因性動機、金銭的動機／社会的動機といった観点で多くの研究がなされている [Rogstadius 11, Shaw 11]。しかし、既存研究の多くは特定タスクに関する実験から結論を導いているため、インセンティブの効果が特定タスクに固有のものであるのか、一般的に成り立つのかを判断することは難しい。また、クラウドソーシングで依頼されるタスクは多様であるため、各種インセンティブの効果を網羅的に調べておくことも難しい。

この問題を解決するため、本研究では、タスク毎に様々なインセンティブを試しながら、適切なインセンティブ設計を得るというアプローチを考える。これは、インセンティブに対する一般的な知見を得ることをあきらめる立場とも言える。クラウドソーシングでは、最小のタスク単位を Human Intelligence Task (HIT) と呼ぶ。分類タスクやアノテーションタスクでは、複数の HIT が集まって、一つのタスクが構成されている。よって、上記のアプローチを取ることが可能である。このような状況では、依頼者は適切なインセンティブの探索と、収集したデータに基づく活用の両立が必要である。この課題の解決には多腕バンディットモデルの適用が有望である。

多腕バンディットモデルは、以下のような問題を記述するモデルである。複数のアームのそれぞれに対して利得の確率分布が決められており、プレイヤーがあるアームを選ぶと利得が確率的に得られる。ただし、プレイヤーは各アームの確率分布を事前

に知らされてはいない。この前提の元で、多腕バンディット問題は、アームの選択回数が与えられたときに、利得の総和を最大化するようにアームを選択する問題である。また、この問題を解くための多腕バンディットアルゴリズムに関して、様々な提案がなされている [Kuleshov 14]。

本稿では、クラウドソーシングにおけるインセンティブ設計への多腕バンディットモデルの適用可能性を検討する。具体的には、Amazon Mechanical Turk で様々なインセンティブ設計の元でワードパズルタスクを依頼し、その結果を元に多腕バンディットアルゴリズムの性能をシミュレーションを用いて評価する。得られた結果は、多腕バンディットモデルの適用が有効であることを示している。

## 2. 関連研究

### 2.1 インセンティブの影響評価

報酬額を含めて各種インセンティブが作業品質にどのような影響を与えるかについて、これまで多くの報告がある。Mason らは、パズルタスクを用いて、金銭的報酬の多寡が処理されるタスク量と品質に与える影響を調べる実験を行い、金銭的報酬を上げることで処理量は増えるものの、品質は必ずしも向上しないことを示した [Mason 10]。Rogstadius らは、数え上げタスクを用いて、タスクの社会的意義を強調してワーカーの内因性動機を高めるようなインセンティブを与えると、報酬額が小さい（外因性動機が弱い）場合でも作業品質が向上することを示している [Rogstadius 11]。Shaw らは、ある web サイトの内容分析タスクを用いて、金銭的なもの社会的なものを含む全部で 14 個のインセンティブが作業品質に与える影響を調べ、その中で 2 つの金銭的インセンティブに効果があることを示している [Shaw 11]。

このように、既存研究の間でも必ずしも整合的な結果が得られているとは言えない。もちろんインセンティブと言ってもワーカーへの提示の仕方はタスクに依存する部分もあり、その影響と考えることも可能であるが、一方で、微小な差異に注意する必要があることになって、一つの知見を様々なタスクに適用することを困難にする。

## 2.2 クラウドソーシングにおける多腕バンディットの適用

Tran-Thanh らは、高度な専門知識を持つワーカーに対してタスクを割り当てる expert crowdsourcing において、ワーカーに割当て可能なタスク数や、依頼者の予算に制限があるという現実的な制約の元で、依頼者の利得を最適化するタスク割当てアルゴリズム bounded epsilon-first アルゴリズムを提案し、その有効性を示している [Tran-Thanh 12]。彼らの研究では、個々のワーカーがアームに対応するのに対し、本研究では各種インセンティブがアームに対応する点に違いがある。また、彼らの研究では、依頼者がワーカーに割り当てるタスクを選べることを仮定しているが、本研究ではワーカーにタスクの選択権があると仮定している点に違いがある。

## 3. 問題設定

### 3.1 タスク

本研究では、マイクロタスクと呼ばれる、一つの仕事に対して複数のワーカーが参加し、比較的小規模な作業を独立して行う形式のタスクを対象とする。例としては、画像のアノテーションや分類、文章の校正などが挙げられる。

### 3.2 依頼者

依頼者は個人から企業まで様々であるが、本研究では一度に数百から数千の HIT が含まれるタスクを依頼するような場合を想定する。また、依頼者がタスクを依頼し、ワーカーによる作業の品質を計測し、その結果に応じて次に依頼するタスクのインセンティブ設計を行うことを想定している。

### 3.3 ワーカー

本研究では、1 ワーカーが複数の HIT を処理することを禁じないが、ワーカーの作業履歴などを入手することはできないと仮定する。つまり、あるインセンティブが有効といった議論をする場合、それはワーカーの集団に対してであり、ワーカー個人に対するものではない。また、クラウドソーシングでは、真面目に作業せずに報酬を受け取ろうとする不誠実なワーカーの存在が知られている。本研究は、そのようなワーカーを誠実な作業に仕向けるような動機付けを対象とするのではない。様々な作業態度を含むワーカー集団に対して、より作業品質を高めるインセンティブの与え方を考察の対象としている。

### 3.4 インセンティブ

クラウドソーシングにおけるワーカーの行動動機に関しては、Hossain らが分類を与えている [Hossain 12]。これら以外にも、動機付けに関しては心理学や実験経済学などで多くの議論があり、すべてを列挙することは難しい。また、特に金銭的インセンティブに関しては、報酬額など多値変数を含み、それらを個別に列挙すれば、さらにインセンティブの数が増える。ただし、本研究の目的は個々のインセンティブそのものの理解ではなく、成果物の品質を向上させるインセンティブ設計の発見である。よって、考慮対象とするインセンティブをある程度絞って考えても、議論の妥当性を損なわないと考える。

また、本研究ではインセンティブの順序効果を考慮しない。順序効果とは、例えば、最初の HIT では作業精度に応じて減額し、次の HIT で作業精度に応じて増額するのと、減額・増額の順序を逆にするので、この 2 HIT から得られる成果物の品質が異なることを指す。後述の評価実験では、1 ワーカーには 1 HIT しか割り当てないとしているため、順序効果を直接考慮する必要はないが、多腕バンディットアルゴリズムの適用を考える上でも、順序効果はないと仮定して議論を進める。

## 3.5 定式化

本研究では、各インセンティブをスロットマシンにおけるアーム、選択したインセンティブを与えたタスクに対するワーカーの作業品質をマシンからの利得、依頼者をプレイヤーと考えて、インセンティブ設計に多腕バンディットを適用する。

依頼者は  $N$  個の HIT 含むタスクを依頼する。依頼者が各 HIT に与えることができるインセンティブが  $K$  種類 ( $K < N$ ) あるとし、 $i_k (1 \leq k \leq K)$  で表す。  $n$  番目 ( $1 \leq n \leq N$ ) の HIT  $t_n$  に与えるインセンティブを  $I(t_n) \in \{i_1, i_2, \dots, i_K\}$  で表す。インセンティブ  $i_k$  を与えられた HIT に対する作業品質  $q_k$  は平均  $\mu_k$ 、分散  $\sigma_k^2$  の正規分布に従うものと仮定する。依頼者はタスク  $t_n, n = 1, 2, \dots, N$  においてインセンティブ  $I(t_n)$  を選択し、作業品質  $q_{I(t_n)}$  を得る。なお、ワーカーの能力は均一であり、作業品質の違いはインセンティブによってのみ生じると仮定する。

このとき、依頼者の戦略  $\pi : \langle \hat{\mu}_1, \hat{\mu}_2, \dots, \hat{\mu}_K \rangle \rightarrow i_k \in \{i_1, i_2, \dots, i_K\}$  は、過去の探索の結果から求められる各インセンティブの期待品質  $\hat{\mu}_k, k = 1, 2, \dots, K$  を入力として、次の HIT で与えるインセンティブ  $i_k$  を返す関数である。ここで、 $\hat{\mu}_k$  は、これまでの探索に基づく推定値であり、真の作業品質の期待値  $\mu_k$  とは異なるかもしれない。依頼者の目標は、 $\sum_{n=1}^N q_{I(t_n)}$  を最大化することである。

## 4. 多腕バンディットアルゴリズム

多腕バンディットアルゴリズムとして  $\epsilon$ -first や  $\epsilon$ -greedy, Softmax, Upper Confidence Bounds (UCB) などが代表的なものとして知られている [Kuleshov 14]。

$\epsilon$ -first は全試行回数のうち前半を探索に当て、後半では最も期待値の高いアームを引き続ける単純なアルゴリズムである。 $\epsilon$  は探索に当てる試行回数の、全試行回数に対する割合を表すパラメータである。

$\epsilon$ -greedy も単純なアルゴリズムで、それゆえに広く使われている。 $\epsilon$  は貪欲さを表すパラメータであり、期待値が最大のアームを確率  $1 - \epsilon + \epsilon/K$  で引き、それ以外のアームを  $\epsilon/K$  の等確率で引く。ここで、 $K$  はアームの総数である。

Softmax は過去の利得の平均値に応じて確率的にアームを引くことで探索と活用を行う。各アームの評価値が早く収束するが、一方で、各アームの試行回数を考慮せず、評価が低いことがわかっているアームにも一定の割合で探索機会を与える。

UCB は、各アームについてどれだけ知っているかを考慮し、評価の低いアームにも比較的多くの探索機会を与えるアルゴリズムである。

以下の評価実験ではこれら 4 つのアルゴリズムの性能を調べる。本稿は、多腕バンディットアルゴリズムの適用可能性を調べるのが目的であり、新規なアルゴリズムを提案するものではない点に注意されたい。

## 5. Mechanical Turk 実験

インセンティブ設計への多腕バンディットの適用可能性を評価するには、各インセンティブに対するワーカーの作業品質データが必要である。そこで、Amazon Mechanical Turk 上で実験を行った。

### 5.1 実験タスク

本研究では、評価タスクとしてワードパズルを用いる。ワーカーは  $12 \times 12$  のマスに書かれた文字の中から、10 個の単語を探そう求められる。ただし、回答方式として選択回答と記

Q3. Type all words you've found in the puzzle with LARGE CASE.

Clues:

- the outer or farthest point of something
- something that is an exact copy of something else
- a door, gate, etc., by which you can enter a building or place
- having nothing inside
- a state of mind or attitude
- the instructions that control what a computer can do
- someone who collects and writes news stories and articles for newspapers, magazines, radio,
- the language that people speak in France, Belgium, parts of Canada and other countries
- having lived or existed for many years
- a piece of writing involving original study of a subject, esp. for a college or university degree

```

T E F D V O D Z T E T
S C T R P L X D T Y Y
I N E Y E D X A J G N
L A N R R R C W X D Z
A R T J A I C S N T G
N T G T L W P H H N N
R N Y P M I T E G D E
U E U M R L S F W V L
O D L I M I Y M O Y Q
J L T L S Y T R M S P

```

図 1: ワードパズルの問題例 (記述回答方式)

述回答の 2 種類を用意した。選択回答方式では、候補として 15 個の単語が与えられ、その内の 10 単語だけが実際にマスに含まれることを示して、10 個の単語の回答を求める。記述回答方式では、マスに含まれる 10 個の単語各々を説明する記述 (clue) \*1 を与え、その記述から単語を推測して、マスの中から探し出すことが求められる (図 1)。一般的には、選択回答方式が低難度で、記述回答方式が高難度と言える。

一つの HIT は、選択回答方式 2 題 (Q1, Q2 と呼ぶ)、記述回答方式 2 題 (Q3, Q4 と呼ぶ) の計 4 題から構成される。報酬額は \$0.08 の固定額で、次節に示す 9 種類のインセンティブに対し各 100HIT、合計 900HIT を依頼した。なお、少数のワーカが多数の HIT を実行することを防ぐため、1 人のワーカが実行できる HIT を 1 つに制限した。最初にワードパズルに関する説明を与え、その下に注釈 (Note) として “After this HIT has been completed, your answers to these questions will be reviewed for accuracy.” といった各インセンティブに対応する文を掲示した。

## 5.2 インセンティブ体系

Shaw らは 3 種類 14 個のインセンティブを調べている [Shaw 11]。本実験では、将来のタスク依頼の約束など、本実験の枠内で実現が難しい 5 個のインセンティブ (Normative priming questions, Reward Agreement, Punishment Agreement, Promise of future work, Betting on result) を除いた 9 個のインセンティブを用いた。なお、Shaw らが用いたインセンティブ記述にはタスクに依存する文言が含まれていたため、本実験タスクに合うよう適宜修正した。

### 5.2.1 社会的インセンティブ

社会的インセンティブは、他のワーカの存在やタスクを行うことの重要性・意義を意識させるものである。

**順位付け (Tournament Scoring)** 回答精度に関して他のワーカと比較されることを伝える。ただし、他のワーカとの優劣によって報酬額が増減することはない。これは、他の社会的インセンティブにおいても同様である。

**監視の意識付け (Cheap Talk – Surveillance)** 回答精度について審査を受けることを伝える。

**規範の意識付け (Cheap Talk – Normative)** 正確に回答することの重要性を強調する。

**感謝 (Humanization)** ワーカに対する感謝を示す。

\*1 各単語を説明する記述は、Cambridge Dictionaries Online (<http://dictionary.cambridge.org>) による

**信頼 (Trust)** 報酬がタスクの継続の有無によって変化しないことを伝えると同時に、残りの問題にも最善の努力で取り組むと信じていると伝える。

### 5.2.2 金銭的インセンティブ

金銭的インセンティブは、通常の報酬に加えてボーナスを支払う、あるいは、減額をするものである。

**精度に応じた報酬 (Reward Accuracy)** 少なくとも一つの問題に対して答え合わせを行い、その問題に正解していればボーナスを支払うと伝える。

**精度に応じた罰則 (Punishment Accuracy)** 少なくとも一つの問題に対して答え合わせを行い、その問題に正解していなければ報酬を減らすと伝える。

**ベイズ自白剤 (Bayesian Truth Serum (BTS))** 問題に対する他ワーカの成績を予測するよう依頼し、それが実際の結果に近い場合はボーナス受給の可能性ありと伝える。

### 5.2.3 ハイブリッドインセンティブ

ハイブリッドインセンティブは、社会的インセンティブと金銭的インセンティブを組み合わせたものである。

**連帯 (Solidarity)** チーム対抗戦のメンバになっていること、チームが勝てばボーナスを受け取れることを伝える。

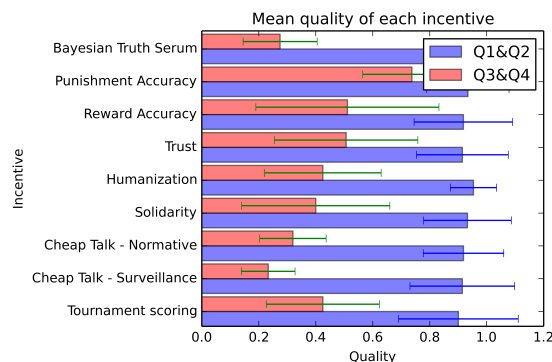


図 2: 各インセンティブに対する平均作業品質

## 5.3 実験結果

図 2 に、与えたインセンティブごとの平均作業品質を示す。横軸が平均作業品質を表す。ここで品質とは、発見された単語数をパズルに含まれる単語数で割った値で計算され、値が大きいほど良い。また、選択回答方式 (Q1&Q2) の結果と記述回答方式 (Q3&Q4) の結果を分けて示す。選択回答方式では、インセンティブを問わず 0.9 以上の品質を得ており、インセンティブ間に有意な差は見られないが、記述回答方式では、9 個のインセンティブ間で品質に差が見られる。「精度に応じた罰則」において品質が高く、一方、「ベイズ自白剤」、「監視の意識付け」において品質が低くなった。

## 6. シミュレーション

Mechanical Turk 実験の結果を元に、多腕バンディットアルゴリズムの適用可能性を検証する目的で、シミュレーションを行った。まず、各インセンティブに対する品質の分布を、Mechanical Turk 実験の結果を元に切断正規分布として表現する。つぎに、1 HIT を依頼する際、9 個のインセンティブに対する品質分布に従って品質を決定し、それを当該インセンティブを用いた場合に得られる品質と考える。この操作を 1,000 回

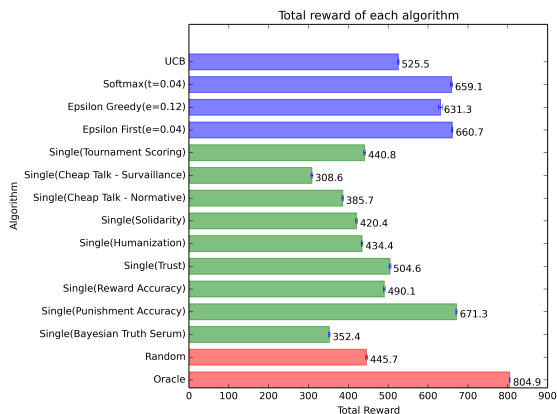


図 3: シミュレーションによるアルゴリズム評価結果

繰り返すことで、1,000 個の HIT を依頼した場合の総利得を求め、比較をした。総利得とは得られる品質の単純和である。図 3 に、4. 章で示した 4 つの多腕バンディットアルゴリズムの総利得を示す。横軸が総利得に対応する。なお、Single() と表記したものは、括弧内のインセンティブを選択し続けるアルゴリズムであり、Random はランダムに 9 個のインセンティブを選択するアルゴリズムである。また、Oracle と表記したものは、常に最善の選択をする、つまり、1 HIT に対して得られる 9 個の品質の値で最大のものを選択し続ける場合を表す。Oracle は比較目的で提示するもので、実現は困難である。

結果を見ると、 $\epsilon$ -first と Softmax では、事前知識がなくても、最善の「精度に応じた罰則」を単独で用いた場合と同等の性能を得ている。また、 $\epsilon$ -greedy もこれらに近い性能を得ている。一方、UCB の性能は劣っている。Kuleshov らは、多腕バンディットアルゴリズムに関して、 $\epsilon$ -greedy や Softmax などの単純なヒューリスティクスを用いた探索が、多くの場合で複雑なアルゴリズムより良い結果を出すことを示している [Kuleshov 14]。ワードパズルタスクに対するインセンティブ設計に関しても、同様の傾向が見られると言える。

## 7. 議論

Shaw らは「ベイズ自白剤」が作業品質の向上に有効であるという結果を得ていた [Shaw 11]\*2。それに対し、我々の実験では異なる結果が得られた。すなわち、Shaw らの研究で効果が認められていた「ベイズ自白剤」を与えた場合に品質が最も低くなり、一方、効果がないとされていた「精度に応じた罰則」を与えた場合に最も品質が高くなった。これは、タスクや、2011 年時点と 2016 年時点といったタスク依頼時期によって、品質向上に有効なインセンティブが異なる可能性を示している。つまり、クラウドソーシングにおいて普遍的に有効なインセンティブが存在しないことになり、多腕バンディットアルゴリズムを援用する価値があると言える。

また、シミュレーションの結果、既存の多腕バンディットのアルゴリズムを適用することで、単一インセンティブの中で最善のものを利用する場合と同等の性能を得た。理想値 (Oracle) と比べても 80% 程度の総利得を達成できており、インセンティブ設計への多腕バンディットの適用は有効であると言える。

\*2 Shaw らは「合意に応じた罰則 (Punishment Agreement)」も有効であったと述べているが、本実験では評価に含めていないため、これについては議論を省略する

さて、これまでは単一タスクに議論を限定してきたが、Mechanical Turk を見れば、依頼されるタスクは様々である。それらのタスク各々に対して、ゼロから学習を始めるのは効率が悪いように見える。例えば、類似タスクに関しては、これまでの学習結果を初期値として与えれば、アルゴリズムが早く収束し、より高い利得が得られると期待できる。しかし、インセンティブの観点でタスク間の類似性を判断することは容易でない。上述の実験では、選択回答方式と記述回答方式を一つのタスクとして扱ったが、前者と後者でインセンティブ間の効果差が異なっていた。依頼者が複数タスクを持つ場合のタスク類似性に基づく効率化の検討は今後の課題である。

## 8. おわりに

本研究では、クラウドソーシングにおけるインセンティブ設計の課題に対し、多腕バンディット問題としてのモデル化を示した。また、ワードパズルタスクに対する Mechanical Turk 実験から、9 個のインセンティブに対する作業品質データを獲得し、それを元にしたシミュレーションにより既存の多腕バンディットアルゴリズムを評価した。インセンティブの効果に関しては、既存研究と異なる結果を得た。これは、普遍的に有効なインセンティブが存在しないことを示しており、タスク毎のインセンティブ設計が必要と言える。また、多腕バンディットアルゴリズムが単一インセンティブの中で最善のものを利用する場合と同等の性能を持つことを確認した。これらより、インセンティブ設計への多腕バンディット適用の有効性を示した。

## 謝辞

本研究は、日本学術振興会科学研究費基盤研究 (S)(24220002, 平成 24 年度～28 年度) の補助を受けた。

## 参考文献

- [Hossain 12] Hossain, M.: Users' motivation to participate in online crowdsourcing platforms, in *Innovation Management and Technology Research (ICIMTR), 2012 International Conference on*, pp. 310–315 (2012)
- [Kuleshov 14] Kuleshov, V. and Precup, D.: Algorithms for multi-armed bandit problems, *arXiv preprint arXiv:1402.6028* (2014)
- [Mason 10] Mason, W. and Watts, D. J.: Financial incentives and the performance of crowds, *ACM SigKDD Explorations Newsletter*, Vol. 11, No. 2, pp. 100–108 (2010)
- [Rogstadius 11] Rogstadius, J., Kostakos, V., Kittur, A., Smus, B., Laredo, J., and Vukovic, M.: An Assessment of Intrinsic and Extrinsic Motivation on Task Performance in Crowdsourcing Markets., in *ICWSM* (2011)
- [Shaw 11] Shaw, A. D., Horton, J. J., and Chen, D. L.: Designing incentives for inexpert human raters, in *Proceedings of the ACM 2011 conference on Computer supported cooperative work*, pp. 275–284 (2011)
- [Tran-Thanh 12] Tran-Thanh, L., Stein, S., Rogers, A., and Jennings, N. R.: Efficient crowdsourcing of unknown experts using multi-armed bandits, in *European Conference on Artificial Intelligence*, pp. 768–773 (2012)