

比較バンディットを用いたクラウドソーシングにおける品質・コストトレードオフの自動調整

Controlling the Quality-Cost trade-off in Crowdsourcing by Dueling Bandits

石島 正和^{1*} 小宮山 純平² 馬場 雪乃³
Masakazu Ishihata¹ Junpei Komiyama² Yukino Baba³

¹ 北海道大学 ² 東京大学 ³ 京都大学
¹ Hokkaido University ² Tokyo University ³ Kyoto University

Abstract:

We propose a new evaluation process for artifacts created by crowdsourcing workers. The propose method employs a dueling bandits algorithm for control quality-cost trade-off of crowdsourcing. We empirically shows that our proposed method reduces the cost of crowdsourcing without reducing the quality of the obtained artifacts.

1 はじめに

クラウドソーシングとは、インターネットを介して不特定多数の労働者 (Worker) に対して仕事 (Task) を依頼し、その成果物 (Artifact) 取得するプロセスのことである。一般的にクラウドソーシングの Worker の能力にはばらつきがあり、まともに Artifact を製作しないで手抜きをする Worker が存在する場合がある。そのため、同一の Task を複数の Worker に依頼し、複数の Artifact を統合することで、質の高い最終成果物を得る試みがなされる。例えば、Task が画像中に人物が写っているかを答える 2 値分類ならば、Artifact は Yes/No の 2 値変数であり、複数の Artifact は多数決により統合でき、1 人に Task を依頼するよりも質の高い最終成果物を得ることができる。しかし、Task によっては複数の Artifact を単純に統合できない場合がある [5]。例えば、Task が与えられた英文を和文に翻訳する問題であるとき、Artifact は日本語の翻訳文に対応し、複数の翻訳文を 1 つに統合することは容易ではない。このような場合、Artifact を統合する代わりに、得られた複数の Artifact から最も優れたものを 1 つ選び、最終的な出力にする方法が考えられる。しかし、Task によっては Artifact の良し悪しを定量的に測ることが困難な場合がある。例えば、Task がイベントロゴのデザインであるとき、Artifact はイベントロゴに対応し、それらを統合することは困難であり、かつ、それらの良し悪しを定量的に評価することも難しい。

結果の統合が困難である Task に対してクラウドソーシングを行い、品質の良い最終成果物を得るための枠組みとして、Two Stage Model (TSM) [1] が提案されている。このモデルでは、Creator と呼ばれる Artifact を生成するための Worker と、Evaluator と呼ばれる Artifact を評価する Worker が存在する。まず、複数の Creator に Task を依頼し、複数の Artifact を得る。そして、複数の Evaluator にそれらの Artifact を定量的に評価してもらい、得られた絶対評価を統合することで最終成果物を決定する。しかし、各 Artifact に絶対評価を与える仕事はコストが高く、また、Evaluator 毎に評価の基準が異なる可能性がある。

Two Stage Pairwise Model (TSPM) [8] では、Evaluator は 1 つ 1 つの Artifact に定量的な評価値を付与する代わりに、与えられた 2 つの Artifact のうち、どちらがより優れているかを回答する。TSPM は、通常の TSM と異なり、Evaluator の仕事は 2 つの Artifact のどちらが優れているかを答える 2 値分類問題であるため、1 つ 1 つの Artifact に定量的な評価値を与える TSM と比べて、Evaluator の仕事の難易度は低い。しかしその一方で、全ての Artifact の比較結果を得るには、各 Evaluator は全 Artifact の組を比較する必要がある。これにより、TSPM では最終成果物を得るために非常に高いコストが掛かることが予想される。

本稿では TSPM において、最終成果物の品質とそれを得るのに要するコストを自動調整する手法を提案する。具体的には、TSPM における品質・コストのトレードオフをコントロールする問題を比較バンディット問題に定式化することで、優秀な Creator の探索と、こ

*連絡先: 北海道大学 大学院情報理工学研究所
〒060-0814 北海道札幌市北区北 1 4 条西 9 丁目
E-mail: ishihata.masakazu@ist.hokudai.ac.jp

れまでに得られた比較結果の活用をコントロールする。

本稿の構成は以下のとおりである。第 2 節では、本稿で扱うクラウドソーシングで得られた Artifact の評価モデルである TSPM を定式化し、Copeland 勝者の Artifact を最終成果物とする Copeland TSPM (CTSPM) を提案する。第 3 節では、CTSPM の品質とコストのトレードオフをコントロールする問題を比較バンディット問題として定式化し、比較バンディットアルゴリズムを用いて CTSPM を解く方法を提案する。第 4 節では、提案手法を用いることで、最終成果物の品質をあまり低下させることなく、コストを削減できることを実データを用いて実験的に示す。第 5 節では、提案手法に関連する研究をいくつか述べ、第 6 節ではまとめと今後の課題を述べる。

2 成果物の評価モデル

2.1 Two Stage Pairwise Model

ここではまず、本稿で扱う Two Stage Pairwise Model (TSPM) [8] を定式化する。Two Stage Model では、実際に目的の Task を解く Worker である Creator と、Creator の成果物である Artifact を評価する Worker である Evaluator が存在する。TSPM では Evaluator は Artifact 1 つ 1 つに評価値を与えるのではなく、2 つの Artifact が与えられた時に、どちらの方がより優れているかを返すことで、Artifact を評価する。

Task 数を ℓ 、Creator 数を m 、Evaluator 数を n とする。Task $t \in [\ell]$ を Creator $i \in [m]$ に依頼し、得られた Artifact を $a_{t,i}$ とし、 $A_{*,i} = \{a_{t,i} \mid t \in [\ell]\}$ 、 $A_{t,*} = \{a_{t,i} \mid i \in [m]\}$ とする。Artifact $a_{t,i}, a_{t,j} \in A_{t,*}$ の比較を Evaluator $k \in [n]$ に依頼し、得られた比較結果を $w_{t,i,j,k} \in \{0,1\}$ とする。ここで $w_{t,i,j,k} = 1$ は Task t において Creator i は Creator j に Evaluator k の評価の元で勝利したことを意味し、0 ならば i は j に敗北したこと意味する。TSPM は全対比較結果 $W = \{w_{t,i,j,k} \mid t \in [\ell], i, j \in [m], i \neq j, k \in [n]\}$ より最終成果物 $A_{final} \in \times_{t \in [\ell]} A_{t,*}$ を決定する問題である。

TSPM において全対比較結果 W より最終成果物 A_{final} を決定する方法は複数考えられる。しかし仮に、全対比較結果 W と最終成果物 A_{final} の品質が独立である場合、 W から品質の高い A_{final} を決定することは不可能である。 $q(A_{final})$ を最終成果物 A_{final} の観測することができない真の品質とし、 $E[q(A_{final})]$ をその期待値とする。つまり TSPM は W から $E[q(A_{final})]$ が高い A_{final} を決定する問題である。本稿では、 W と $E[q(A_{final})]$ にいくつかの仮定を導入することで、最も期待品質の高い成果物を選ぶ手法を提案する。

2.2 Copeland TSPM

本稿ではまず、 W と $E[q(A_{final})]$ に関して、いくつかの仮定を導入し、その仮定の元で $E[q(A_{final})]$ を最大化する Copeland TSPM (CTSPM) を提案する。まず、 W が従う分布に関して以下を仮定する。

仮定 1. 対称行列 $P_k^* \in [0,1]^{m \times m}$ を Evaluator k の**選好**と呼び、 $p_{k,i,j}^*$ を P_k^* の i 行 j 列目の要素とする。このとき、一対比較結果 $w_{t,i,j,k}$ は $p_{k,i,j}^*$ をパラメータするベルヌーイ分布に従う。

つまり、Evaluator k の判断における Creator i, j の勝率は $p_{k,i,j}^*$ であり、 $p_{i,j,k}^*$ は Task t に依存せず一定である。次に、選好 P_k^* と各 Creator の成果物 $A_{*,i}$ の期待品質 $E[q(A_{*,i})]$ に関して以下を仮定する。

仮定 2. P^* と L_i^* をそれぞれ以下とする。

$$P^* = \frac{1}{n} \sum_{k \in [n]} P_k^*, \quad (1)$$

$$L_i^* = \{j \in [m] \mid p_{k,i,j}^* < 1/2\}. \quad (2)$$

P^* を**真の選好**、 L_i^* を**真の負け数**と呼ぶ。このとき、 L_i^* と $E[q(A_{*,i})]$ の間には以下の関係が成り立つ。

$$L_i^* < L_j^* \implies E[q(A_{*,i})] > E[q(A_{*,j})] \quad (\forall i, j \in [m]) \quad (3)$$

つまり、真の選好 P^* の元で負け数 L_i^* が最も小さい Creator の Artifact が最も高い期待品質を持つと仮定する。ここで

$$c^* = \arg \min_{i \in [m]} L_i^* \quad (4)$$

を真の選好 P^* の**Copeland 勝者**と呼ぶ。真の選好 P^* は各 Evaluator の選好 $p_{i,j,k}^*$ を k に関して平均化した選好であり、真の選好において勝率の高い Creator の Artifact は高い期待品質を持つ。真に優秀な Creator はどの Evaluator の選好でも上位に来ると思われるため、平均化しても上位に来ることが期待できる。また能力の低い Evaluator やランダムに振る舞う Evaluator が存在しても、その数が少なければ、その影響は平均化によって小さくなると期待できる。

上記の仮定より、コペランド勝者 c^* が既知であるとき、 $A_{final} = A_{c^*}$ とすれば TSPM の最終結果の期待品質を最大化できる。しかし、実際には P^* を知ることはできないため、真のコペランド勝者 c^* も知ることはできない。そこでここでは、全対比較結果 W よりコペランド勝者 c^* の推定値 \hat{c} を計算し、 $A_{final} = A_{\hat{c}}$ とする。この手法を Copeland TSPM と呼ぶ。 \hat{c} を経

験コープランド勝者と呼び、以下のように計算する。

$$\hat{c} = \arg \min_{i \in [m]} \hat{L}_i, \quad (5)$$

$$\hat{L}_i = |\{j \in [m] \mid \hat{p}_{i,j} < 1/2\}|, \quad (6)$$

$$\hat{p}_{i,j} = \frac{1}{\ell n} \sum_{t \in [\ell]} \sum_{k \in [n]} w_{t,i,j,k}. \quad (7)$$

$\hat{p}_{i,j}$ は Creator i の Creator j に対する経験勝率であり、 \hat{L}_i は i の経験負け数である。CTSPM のアルゴリズムは以下のように書ける。

1. 各 Task $t \in [\ell]$ に対して 2.-3. を行い全対比較結果 W を得る。
2. 各 Creator $i \in [m]$ に Task t を依頼し、Artifact $a_{t,i}$ を得る。
3. 各 Evaluator $k \in [n]$ に $a_{t,i}, a_{t,j}$ ($\forall i, j \in [m], i \neq j$) の比較を依頼し、一対比較結果 $w_{t,i,j,k}$ を得る。
4. 全対比較結果 W より経験コープランド勝者 \hat{c} を式 (5) に従い計算し、 $A_{final} = A_{*,\hat{c}}$ を出力する。

次に CTSPM を実行するために必要なコスト (費用) を定式化する。Creator 1 人に Task 1 つを依頼し、Artifact 1 つ得るコストを C_C とする。また、Evaluator 1 人に Artifact 2 つの比較を依頼し、一対比較結果 $w_{t,i,j,k}$ を得るコストを C_E とする。このとき、CTSPM を実行するために必要なコストは以下である。

$$C_{CTSPM} = \ell \left(mC_C + \frac{m(m-1)}{2} nC_E \right) \quad (8)$$

一般的に Creator と Evaluator では、前者の方が高度な仕事を行うため、 $C_C > C_E$ である。本稿では期待品質 $E[q(A_{final})]$ をできるだけ低下させることなく、このコスト C_{CTSPM} を削減することを目的とする。

2.3 Online CTSPM

CTSPM では全対比較結果 W をクラウドソーシングによって得た後に、経験コープランド勝者 \hat{c} を計算する。全対比較結果 W を得るには $O(mC_C + m^2nC_E)$ のコストを要する。そこで本稿では、全対比較結果 W を得る代わりに、コープランド勝者 c^* を推定するのに有用な W の部分集合を得る方法を考える。これを実現するために、本稿では CTSPM を online 化した以下の Online CTSPM を考える。

1. 何らかの方法で Task t , Creator i, j , Evaluator k を選ぶ。
2. Task t を Creator i, j に依頼し、Artifact $a_{t,i}, a_{t,j}$ を得る。

3. Evaluator k に $a_{t,i}, a_{t,j}$ の比較を依頼し、一対比較結果 $w_{t,i,j,k}$ を得る。

4. 必要ならば 1.-3. を s 回繰り返す。

5. s 個の比較結果を元に経験コープランド勝者 \hat{c} を計算し、 $A_{final} = A_{*,\hat{c}}$ を出力する。

上記の 1.-3. を 1 ラウンド呼び、 s をラウンド数という。このとき、OCTSPM のコストを C_{OCTSPM} とすれば、以下の C_{OCTSPM} の上限が得られる。

$$C_{OCTSPM} \leq s(2C_C + C_E) + \ell C_C \quad (9)$$

ここで ℓC_C は推定コープランド勝者 \hat{c} を計算した後に、 \hat{c} に各 Task を依頼し、最終成果物 $A_{*,\hat{c}}$ を得るためのコストである。実際にはラウンド中に同一の Task を同じ Creator に複数回依頼する可能性があり、その場合は過去の Artifact を再利用するため、実際のコスト C_{OTSPM} は上記の上限より小さくなる。 C_{CTSPM} は Creator 数 m と Evaluator 数 n に依存するが、 C_{OCTSPM} は m, n には依存しない。ラウンド数 s のときのコストの削減率 $R \in \mathbb{R}$ は以下である。

$$R = \frac{C_{OCTSPM}}{C_{CTSPM}} \quad (10)$$

つまり、CTSPM に対してコストを R に削減したい場合、上式より OCTSPM が行えるラウンド数 s を逆算することができる。一方で、OCTSPM の最終成果物の期待品質は、各ラウンドでの Creator i, j の選択に強く依存する。仮に真のコープランド勝者 c^* の候補を正しく見積もることができ、 i, j をその中から選べれば OCTSPM の期待品質は CTSPM に近づく。本稿では、OCTSPM を比較バンディット問題として定式化することで、CTSPM の期待品質とコストのトレードオフを自動的にコントロールする。

3 Dueling Bandits for OCTSPM

本稿では OCTSPM を比較バンディット問題として定式化することで、クラウドソーシングで得られる最終成果物の期待品質と必要なコストのトレードオフを自動調整する。真のコープランド勝者 c^* が既知であるとき、 $A_{final} = A_{*,c^*}$ とすることが期待品質を最大化する。しかし実際には c^* を知ることはできないため、 c^* を推定するために複数の Creator を試す (探索する) 必要がある。しかし、Creator を試すにはコストが必要であるため、無駄な探索は避ける必要がある。

$t(s'), i(s'), j(s'), k(s')$ ($s' \in [s]$) をそれぞれ、OTSPM の s' 回目のラウンドで選ばれた Task, Creator, Evaluator とする。すると、OTSPM を実行するアル

ゴリズムは、 $t(s'), i(s'), j(s'), k(s')$ を過去の比較結果 $\{w_{t(s''), i(s''), j(s''), k(s'')} \mid s'' \in [s']\}$ より決定するアルゴリズムである。仮定 1 より $w_{t, i, j, k}$ は t に依存しないため、 t はどのように選んでも良い。また、仮定 2 より、各 Evaluator の選好 P_k^* を平均化した真の選好 P^* 上のコープランド勝者 c^* を考えるため、 k を一様分布に従ってサンプルすることで P_k^* を周辺化する。すると OCTSPM は $i(s'), j(s')$ のみを決定する問題になり、これは一般的な比較バンディット問題と等価である。

比較バンディット問題 [9] は、複数の選択肢からそれらの 2 対比較結果を得ることで、もっとも優れた選択肢を推定する手法である。比較バンディット問題では、通常のコストと探索とコストのトレードオフをコントロールすることで効率よく優れた選択肢を推定する。具体的には、比較バンディット問題では以下で定義される Regret を最小化することで探索とコストのトレードオフを調整する。

$$\text{Regret}(s) = \sum_{s' \in [s]} \left(L_{i(s')}^* + L_{j(s')}^* - 2L_{c^*}^* \right)$$

比較バンディットアルゴリズムが出力するラウンド s' でのコープランド勝者の推定値を $\hat{c}(s')$ とする。仮にアルゴリズムがこの推定値に自信を持っているとき、アルゴリズムは $i(s') = j(s') = \hat{c}(s')$ とする。 $i(s') = j(s')$ であるとき、Task t を依頼する Creator は 1 人であるため、そのラウンドに必要なコストは C_C だけでよく、2 人に依頼した場合と比べて $C_C + C_E$ だけコストを削減できる。つまり比較バンディットアルゴリズムを用いた OCTSPM では、有望な Creator i, j のみを探索することによるコスト削減だけでなく、推定に自信が有るときに無駄な探索を避けることによるコスト削減も見込める。

比較バンディットアルゴリズムが **一貫性** を持つとは、任意の選好 P^* とある定数 $\alpha > 0$ に対して、アルゴリズムが達成する Regret が $R(s) = o(s^\alpha)$ を満たすことである。これは、各ラウンド $s' \in [s]$ において、仮説「 $c^* \neq \hat{c}(s')$ 」を有意水準 $1/s'$ で棄却できることに対応する。つまり一貫性を持つ比較バンディットアルゴリズムは、探索のどの時点においても現時点の推定値について統計的な保証を持つ。

4 実験

本稿では CTSPM と比較バンディットアルゴリズムを用いた OCTSPM を実データを用いて比較をする。ここではまず、実験に用いる実データと比較バンディットアルゴリズムについて述べる。そして CTSPM と OTSPM を比較し、どの程度のコスト削減と品質低下が起こっているかを確認する。

4.1 データ・セット

本稿では、2 つの実データを用いて提案手法を評価する。各データのタスク内容、Task 数 ℓ 、Creator 数 n 、Evaluator 数 m はそれぞれ表 1 の通りである。実際にクラウドソーシングによって全対比較結果 W を得るには、膨大なコストがかかるため、表 1 のデータ・セットは W 全体ではなく、 W の一部のみである。各データ・セットに含まれる Artifact の数、比較された Airtifact のペア数、実際に得られた一対比較結果数はそれぞれ表 2 の通りである。

また、このデータ・セットでは、各 Artifact を信頼できる方法で評価した結果を含んでおり、その評価結果を用いることで各手法の最終成果物 A_{final} を定量的に評価可能である。各 Artifact は 30 人の Worker により 5 段階評価されており、 $q_{t,i}$ を Artifact $a_{t,i}$ のエキスパート評価結果の平均とし、 $q_i = \frac{1}{\ell} \sum_{t \in [\ell]} q_{t,i}$ とする。 q_i を Creator i の平均品質と呼ぶ。なお、5 段階評価を行った Worker 集合は、一対評価に参加した Evaluator 集合とは異なる。本稿では観測できない真の品質 $q(A_{*,i})$ の代わりにこの q_i を用いて提案手法を評価する。

表 1: 実験に用いる実データのタスク内容、Task 数 ℓ 、Creator 数 n 、Evaluator 数 m

| dataset | タスク内容 | ℓ | m | n |
|-------------|-------|--------|-----|-----|
| description | 画像説明 | 20 | 20 | 187 |
| translation | 英日翻訳 | 20 | 17 | 68 |

表 2: Artifact 数、比較ペア数、比較結果数

| dataset | # artifacts | # pairs | # comparisons |
|-------------|-------------|---------|---------------|
| description | 200 | 940 | 16,314 |
| translation | 190 | 825 | 15,980 |

4.2 実験設定

本稿では 4.1 で述べたデータ・セットを用いて CTSPM と比較バンディットを用いた OCTSPM の結果を比較する。具体的には、両手法を各データ・セットに適用し、OCTSPM が CTSPM に対してどの程度のコスト削減率 R でどの程度の品質の最終成果物を得られるかを確認する。ここで実験データは全対比較結果ではなくその一部であるため、場合によっては提案手法が要求する一対比較結果 $w_{t, i, j, k}$ を得られない可能性がある。そこで本稿では、比較バンディットアルゴリズムがラウンド s' で Creator $i(s'), j(s')$ を指定したときに、データが存在するように t, k を選ぶことでその問題を解決する。

本稿では OTSPM を比較バンディット問題として定式化したため、任意の比較バンディットアルゴリズムを採用可能である。本稿では以下で説明する Random, CCB, ECW-RMED の3種類のアロリズムを用いて実験を行い、それぞれの結果を CTSPM と比較する。

- Random は各ラウンドで一様分布に従い Creator を選択する。データを集めた後は、最尤推定により Copeland 勝者を推定する。この手法は一貫性を持たない。
- CCB (Copeland Confidence Bound) [11] はバンディット問題においてよく知られている UCB (Upper Confidence Bound、信頼上界) アルゴリズムを一对比較向けに改良したものである。CCB は勝率 $p_{i,j}^*$ の信頼区間を考え、信頼区間の上界を用いて各選択肢の過大評価されたコープランド数を推定する。CCB は各ラウンドでこの推定値が高い選択肢を選択する。
- ECW-RMED [6] はバンディット問題において知られている MED (最小経験ダイバージェンス) アルゴリズム [4] を一对比較向けに改良したものである。ECW-RMED は勝率 $p_{i,j}^*$ の経験推定が正しいという仮定において、それぞれの選択肢がコープランド勝者である尤度を計算し、尤度が $1/t$ 以上の選択肢 (コープランド勝者である可能性が一定以上である選択肢) をリストアップし、順番に比較していく。このアルゴリズムはラウンド数が十分大きいときの Regret が CCB と比較して小さいことが知られている。

本稿では 4.1 で述べたデータ・セットに対し、CTSPM と上記の3つの手法を用いた OCTSPM を適用する。なお、OCTSPM は独立に 100 回施行し、その平均の性能を示す。

4.3 仮説の検証

CTSPM は、仮説 1, 2 より、真の選好 P^* のコープランド勝者 c^* が最も高い期待品質を達成すると仮定する。ここではまず、この仮説 1 が実データにおいて成り立っているかを確認する。データ全体を使って式 (6) より経験負け数 \hat{L}_i を計算し、その Creator の平均品質 q_i との関係を確認する。図 1 は、 \hat{L}_i と q_i の関係をプロットしたものである。どちらのデータ・セットにおいても、経験負け数 \hat{L}_i と平均品質 q_i は極めて強い比例関係があるが分かる。また、経験コープランド勝者 \hat{c} の平均品質 $q_{\hat{c}}$ が最も高い値となっていることが分かる。このことから、CTSPM の仮定 1, 2 は妥当であると考えられる。

4.4 実験結果

図 2 は、コスト削減率 R とその時点での経験コープランド勝者 \hat{c} の経験負け数 $\hat{L}_{\hat{c}}$ の関係である。データ全体を使って評価した場合、両データとも $\hat{L}_{\hat{c}} = 0$ である。一方で、Description データは 2 位以降が接戦であるのに対して、Translation データでは 1 位と 2 位以降の間で品質に大きな差がある。これより、Translation データの方が、コープランド勝者の推定が容易であると期待できる。グラフの青線が CTSPM での推定コープランド勝者の負け数であり、他の線が各手法を用いた OCTSPM のコスト削減率 R とその時の経験コープランド勝者の負け数である。図より、Translation データでは CTSPM の 3 割程度のコストで同等の負け数のコープランド勝者を発見できている。Description データでは、CTSPM と同等の性能を達成するのに 6 割程度のコストを要している。どちらの例でもコストの削減が行えていることが確認できる。

図 3 は、コスト削減率 R とその時点での最終成果物の平均品質 $q_{\hat{c}}$ の関係である。概ね負け数の推定と同様の関係となっていることが分かる。実験結果より、比較バンディットアルゴリズムを用いた OCTSPM により、コストの削減を達成しつつ、CTSPM と同等の品質を達成できることが確認できる。

5 関連研究

通常のクラウドソーシングにバンディット問題で一般的に利用される UCB を導入した手法が提案されている [3]。この手法では、2 値識別器を学習するための教師データをクラウドソーシングによって獲得する際に、教師ラベル (Artifact) を生成するエキスパート (Creator) を UCB を用いて決定する。エキスパート毎のコストが異なる場合に拡張した手法も提案されている [10]。これらの手法では、最終成果物である真のラベルは、複数の Creator の Artifact の多数決によって推定する。一方、本稿で提案した Copeland TSPM は、Artifact は単純に統合できないタスクを想定しており、Artifact の比較結果から最終成果物を推定する。本稿では Task 間の差は考慮しなかったが、Task に対する Worker の割当を探索・活用を繰り返しながら決める手法も提案されている [2]。また、Worker への報酬金額が参加 Worker 数に影響することに着目し、より多くの作業結果を集めることを目的として、報酬金額を UCB を用いて決定する手法も提案されている [7]。

TSPM における最終成果物の推定方法として、各 Creator と各 Evaluator の能力と各 Artifact の品質を同時に推定し、最も推定された品質の高い Artifact を出力する方法が提案されている [8]。この手法は、既存

のランキングアルゴリズムを全対比較結果 W から学習できるように拡張した手法である。この手法では W は与えられると仮定するのに対し、本稿の提案手法は、Copeland 勝者を推定するのに必要な W の部分集合を能動的に獲得することが異なる。

6 おわりに

本稿ではクラウドソーシングにおける Two Stage Pairwise Model の期待品質とコストのトレードオフを自動的にコントロールする手法を提案した。提案手法は Corpland TSPM を online 化した OCTSPM を比較バンディット問題として定式化することで、比較バンディットアルゴリズムを用いて最終成果物の平均品質とそれを得るためのコストの自動調整を実現した。また実験により、提案手法が実データにおいて CTSPM と同等程度の品質を、より少ないコストで達成できることを確認した。

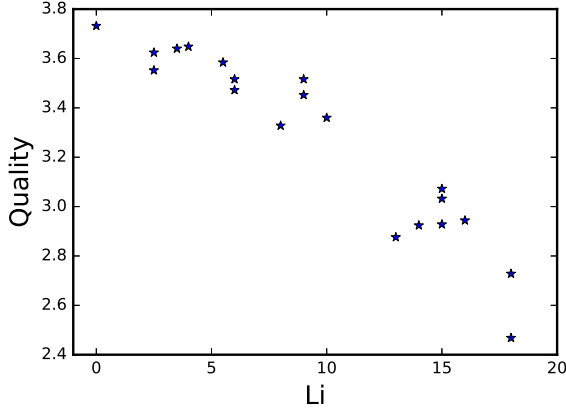
今後の課題として、提案法を Task の違いや、Evaluator の能力を考慮できるように拡張することが挙げられる。本稿の提案手法は Creator の能力のみを考慮しており、Evaluator の能力や Task の難易度を考慮していない。将来的には提案手法を Task 毎に選好が異なるモデルに拡張し、Evaluator の能力も同時に考慮できるように拡張したい。

謝辞

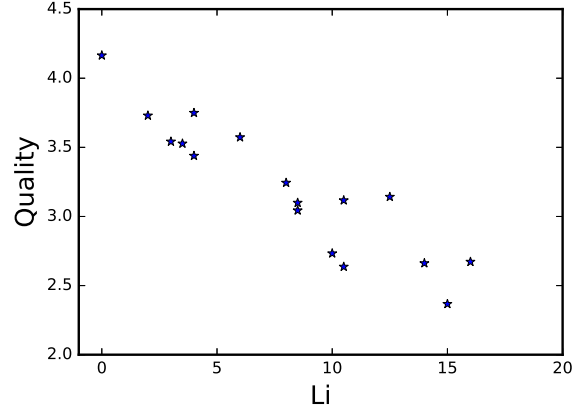
本研究の一部は JSPS 科研費基盤 (S) 15H05711 の助成によります。

参考文献

- [1] Yukino Baba and Hisashi Kashima. Statistical quality estimation for general crowdsourcing tasks. In *The 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2013, Chicago, IL, USA, August 11-14, 2013*, pages 554–562, 2013.
- [2] Xi Chen, Qihang Lin, and Dengyong Zhou. Optimistic knowledge gradient policy for optimal budget allocation in crowdsourcing. In *Proceedings of the 30th International Conference on Machine Learning, ICML 2013*, pages 64–72, 2013.
- [3] Pinar Donmez, Jaime G. Carbonell, and Jeff G. Schneider. Efficiently learning the accuracy of labeling sources for selective sampling. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris, France, June 28 - July 1, 2009*, pages 259–268, 2009.
- [4] Junya Honda and Akimichi Takemura. An asymptotically optimal policy for finite support models in the multiarmed bandit problem. *Machine Learning*, 85(3):361–391, 2011.
- [5] Panagiotis G. Ipeirotis. Analyzing the Amazon Mechanical Turk marketplace. *ACM Crossroads*, 17(2):16–21, 2010.
- [6] Junpei Komiyama, Junya Honda, and Hiroshi Nakagawa. Copeland Dueling Bandit Problem: Regret Lower Bound, Optimal Algorithm, and Computationally Efficient Algorithm. In *Proceedings of the 33rd International Conference on Machine Learning, ICML 2016, New York City, NY, USA, June 19-24, 2016*, pages 1235–1244, 2016.
- [7] Adish Singla and Andreas Krause. Truthful incentives in crowdsourcing tasks using regret minimization mechanisms. In *Proceedings of the 22nd international conference on World Wide Web, WWW 2013*, pages 1167–1178, 2013.
- [8] Takeru Sunahase, Yukino Baba, and Hisashi Kashima. Pairwise HITS: Quality Estimation from Pairwise Comparisons in Creator-Evaluator Crowdsourcing Process. In *21st AAAI Conference on Artificial Intelligence (AAAI)*, 2017.
- [9] Yisong Yue, Josef Broder, Robert Kleinberg, and Thorsten Joachims. The k-armed dueling bandits problem. *J. Comput. Syst. Sci.*, 78(5):1538–1556, 2012.
- [10] Yaling Zheng, Stephen Scott, and Kun Deng. Active learning from multiple noisy labelers with varied costs. In *2010 IEEE International Conference on Data Mining, ICDM 2010*, pages 639–648, 2010.
- [11] Masrour Zoghi, Zohar S. Karnin, Shimon Whiteson, and Maarten de Rijke. Copeland dueling bandits. In *Advances in Neural Information Processing Systems 28 (NIPS2015)*, pages 307–315, 2015.

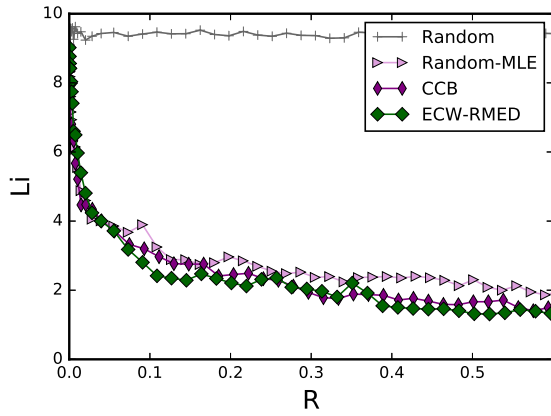


(a) Description

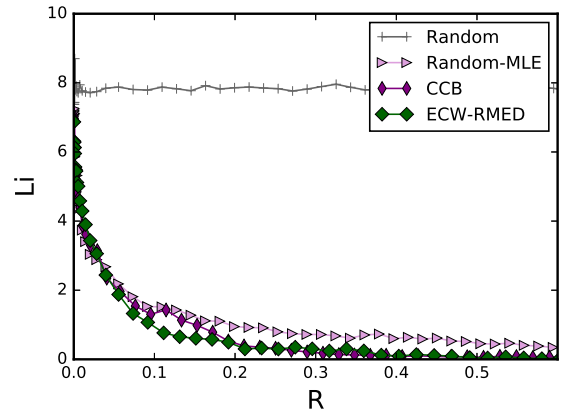


(b) Translation

図 1: 経験負け数 \hat{L}_i と Creator i の平均品質 q_i の関係.

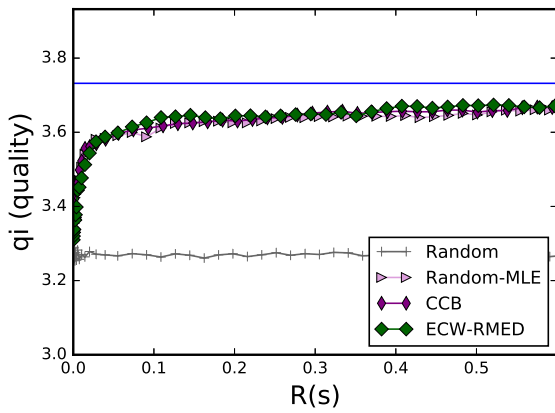


(a) Description

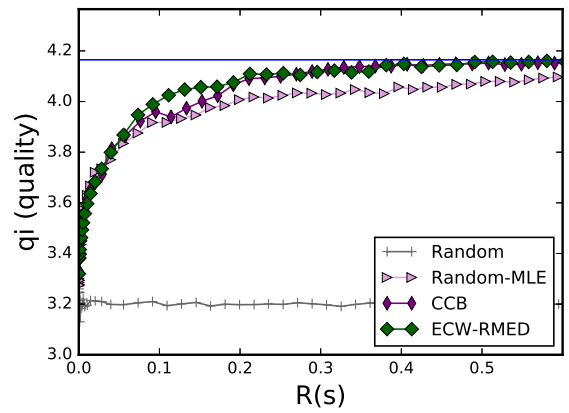


(b) Translation

図 2: コスト削減率 R と経験コープランド勝者 \hat{c} の経験負け数 $\hat{L}_{\hat{c}}$ の関係.



(a) Description



(b) Translation

図 3: コスト削減率 R と経験コープランド勝者 \hat{c} の平均品質 $q_{\hat{c}}$ の関係.