

## ソーシャルネットワークを用いたクラウドソーシングの品質向上

## Improving Data Quality of CrowdSourcing with Social Networks

西 智樹\*<sup>1</sup> 小出 智士\*<sup>1</sup> 大野 宏司\*<sup>1</sup> 長屋 隆之\*<sup>1</sup>  
Tomoki Nishi Satoshi Koide Hiroshi Ohno Takayuki Nagaya

\*<sup>1</sup>(株)豊田中央研究所  
Toyota Central R&D Labs., Inc.

Quality of collected data is one of the largest issues in crowdsourcing services. Previous studies focused on how to improve the quality with machine learning and/or how to give incentives to workers. In these studies, however, few researchers have taken advantage of social networks between workers. In this paper, we propose a novel mechanism on which workers can entrust tasks to a worker friend of them recursively and receive payoffs when the friend carry out the tasks appropriately. We analytically calculate the quality of the data collected with this mechanism, when workers' network is assumed a tree structure. We demonstrate the effectiveness of this mechanism in the tree-structured network with a simulation.

## 1. はじめに

クラウドソーシングサービス（以降、クラウドソーシング）は、不特定多数に仕事を依頼することによりコンピュータでは難しい、人の判断や創造性が必要な仕事を効率的に処理するための画期的な仕組みとして注目を集めている [Howe 06]. クラウドソーシングは比較的安価に大量の仕事を依頼できる反面、処理結果の品質が処理する人（以降、ワーカ）によって大きく異なるという問題が指摘されており、処理結果の品質向上はクラウドソーシングにおける重要課題の一つである。従来の研究では、収集された処理結果からいかに精度の高い正解値や識別器を推定するかを対象としており、処理結果自体の品質向上については扱われて来なかった。またこれまでの多くの研究ではワーカ同士のつながりについては考慮されておらず、ワーカは自分自身の力だけでタスクをこなす必要があった。

そこで本研究では、ソーシャルネットワークの様に一部のワーカ同士が知り合いであることを仮定し、ワーカ間でタスクを再帰的に委託可能にすることにより、処理結果自体の品質向上を実現するクラウドソーシング (REC, Recursive Entrustment-based Crowdsourcing) を提案する。REC では一人のワーカが委託できる人数を一人に限定し、報酬をタスクに正解したワーカ及び正解したワーカにタスクを委託したワーカに支払うことで、知人の中で相対的に能力の低いワーカは自分でタスクを処理するよりも知り合いのワーカにタスクを委託する方が高い報酬が期待できる。そのため能力の低いワーカは能力の高いワーカにタスクを委託するようになり、従来に比べ能力の高いワーカの割合が増加した高い品質の処理結果が得られることが予想される。つまり REC は能力の低いワーカには能力の高いワーカの探索を、能力の高いワーカにはタスクの処理を依頼するクラウドソーシングと見ることが出来る。

本論文では、ワーカ間の知人ネットワークが、次数が一定（知人の数が一定）の木で表現される場合について、REC により得られるタスクの処理結果の品質（正解率）を解析的に示す。また具体例として、ワーカ全体の正解率の分布がベータ分布であると仮定した場合について、平均タスク正解率を算出し REC の有効性を検証する。

## 2. 再帰的委託に基づくクラウドソーシング REC

本章ではタスクを再帰的に委託可能なクラウドソーシングについて説明する。従来型のクラウドソーシングと REC の違いを図 1 に示す。REC は（1）タスクの発見、（2）ワーカ間でのタスクの委託、（3）タスクの処理の 3 つにより構成されており、従来のクラウドソーシングに比べワーカ間でのタスクの委託が追加されたクラウドソーシングである。REC では一人のワーカが委託できる人数を一人に限定し、報酬として (a) 正解したワーカに  $R_a$ , (b) 正解したワーカにタスクを委託したワーカに  $R_b$ , ( $R_a \geq R_b$ ) が支払われるとする。この場合には知人にタスクを委託するよりも、タスクを処理した方が報酬の期待値が高い場合にのみタスクを処理することが合理的な選択となる。そのため全体に対して相対的に能力が低いワーカはタスクを委託する確率が高く、委託する場合には知人の中で最も能力の高い知人に委託するのが合理的な選択となる。これらのことから能力の低いワーカがタスクを処理する確率が減少し、かつ委託が繰り返される度にワーカの能力は単調増加するため REC は従来型のクラウドソーシングに比べ品質の高い処理結果が得られることが期待できる。

以下の節では REC のタスクの品質について、ワーカ間のネットワークが木で表現される場合で検証する。

## 2.1 問題設定

本節では今回用いる記号と取り扱う問題を定義する。ワーカの知人ネットワークは、すべてのワーカに等しく  $m$  人の知人がいるが共通の知人はいない、次数  $m$  の木で表現されるとする。ワーカの集合を  $W = \{w | w \in \{1, \dots, \infty\}\}$  とし、ワーカ  $w$  の知人の集合を  $V_w = \{v | v \in W\}$ ,  $|V_w| = m$ ,  $V_w \cap V_{u \neq w} = \phi$  とする。ワーカ  $w$  のタスク正解率を  $X_w \in [0, 1]$ , ワーカ全体の正解率の確率分布を  $f(x)$  と定義する。各ワーカはタスクを処理する ( $a_1$ ), 知人にタスクを委託する ( $a_2$ ) の 2 種類の行動のどちらか一方を選択することができ、委託された知人も同様であるとする。各ワーカは自身と知人のタスク正解率は正確に知っているが、知人以外のワーカの正解率は知らないとする。ワーカ  $w$  は  $\forall v \in V_w, X_w > X_v$  の場合には処理し、それ以外の場合には  $v = \arg_{v \in V_w} \max(X_v - X_w)$  を満たす知人  $v$  に委託すると仮定する。これはタスクに正解したワーカと、正解

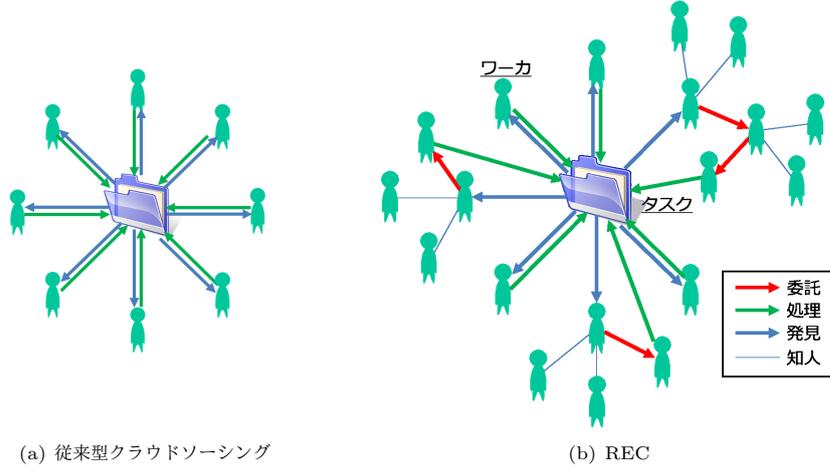


図 1: 従来型クラウドソーシングと REC

したワーカーにタスクを委託したワーカーに同額の報酬を払う場合には合理的な行動選択となっている。

## 2.2 タスク正解率分布

本節では REC のタスク正解率の分布を解析的に示す。  $n$  回目に委託された正解率  $x$  のワーカーが処理する確率は、

$$p(a_1, x, n) = p(a_1|x)p(x|n)p(n) \quad (1)$$

で表される。ここで  $p(a_1|x)$  は正解率  $x$  のワーカーが処理する確率、  $p(x|n)$  は  $n$  回目に委託されたワーカーの正解率が  $x$  である確率、  $p(n)$  は委託が  $n$  回以上繰り返される確率。ワーカーの正解率の確率分布が連続関数  $f(x)$  で表される時、  $p(a_1|x)$ 、  $p(x|n)$ 、  $p(n)$  はそれぞれ下記で表現できる。但しスペースがない場合には  $F(x)$ 、  $f(x)$  などの  $(x)$  は省略する。

$$p(a_1|x) = \left( \int_0^x f(t)dt \right)^m \equiv F^m(x), \quad (2)$$

$$p(x|n) = \begin{cases} f(x) & (n=0) \\ nmf(x)F^{n-1}(x) & (n>0) \end{cases}, \quad (3)$$

$$p(n) = \begin{cases} 1 & (n=0) \\ \frac{m}{(m+1)^n} & (n>0) \end{cases}, \quad (4)$$

$$p(a_1, x, n) = \begin{cases} F^m(x)f(x) & (n=0) \\ \frac{m^2}{m+1} \frac{1}{(n-1)!} F^{(n+1)m-1} f & (n>0) \end{cases}. \quad (5)$$

REC のタスク正解率の確率分布  $p(a_1, x)$  は式 (5) を  $n$  に関して周辺化することにより導出できる。

$$\begin{aligned} p(a_1, x) &= \lim_{n \rightarrow \infty} \sum_{i=0}^n p(a_1, x, i) \\ &= \lim_{n \rightarrow \infty} F^m f \left( 1 + \frac{m^2}{m+1} \sum_{i=0}^n \frac{1}{i!} F^{(i+1)m-1} \right) \\ &= F^m f \left( 1 + \frac{m^2}{m+1} F^{m-1} \exp F^m \right). \end{aligned} \quad (6)$$

このことから正解率の期待値  $E_x$  及び初期に依頼したワーカーの数を  $N$  とした場合に最終的にタスクに関わった人数  $\tilde{N}$  は

$$\begin{aligned} E_x &= \int_0^1 x F^m f \left( 1 + \frac{m^2}{m+1} F^{m-1} \exp F^m \right) dx \\ &= \frac{1}{m+1} \left( 1 - \int_0^1 F^{m+1} + m(1-F^m) \exp F^m dx \right) \\ &> 1 - \frac{1}{m+1} \int_0^1 (F^{m+1} + mF^{2m}) dx, \end{aligned} \quad (7)$$

$$\begin{aligned} \tilde{N} &= \lim_{n \rightarrow \infty} \sum_{i=0}^n Np(i) \\ &= \left( 1 + \frac{m(e-1)}{m+1} \right) N < eN, \end{aligned} \quad (8)$$

で表される ( $e$  は自然対数)。式 (7) から正解率の期待値の下限値は知人の数  $m$  に対して単調増加し、  $m = \infty$  で正解率の最大値に収束することが分かる。また式 (8) から REC を導入することによりタスクに関わる人数は従来法と比べ  $e$  倍未満であることが分かる。

またワーカーのタスク正解率の期待値  $E_0$  からの REC によるタスク正解率の期待値  $E_x$  の増加量は

$$\begin{aligned} E_x - E_0 &= E_x - \int_0^1 x f dx \\ &> \int_0^1 F \left( 1 - \frac{1}{m+1} F^m (1 - mF^{m-1}) \right) dx \\ &> 0 \quad (\because F \leq 0), \end{aligned}$$

となり、REC によりワーカーのタスク正解率の期待値は常に増加することが分かる。

## 3. 実験

前章では知人ネットワークの次数が一定である場合の REC により得られる正解率の期待値を理論的に導いた。本章では REC と従来型のクラウドソーシング (従来法) とのタスク正解率を前章で導いた式に基づき具体的に計算し、REC の有効性を検証する。

表 1: パラメータの条件

パラメータ	サンプリング数
$m \in \{2, 6, 10\}$	3
$0.1 \leq E \leq 0.9$	9
$0 < V < \min\left(\frac{E^2(1-E)}{1+E}, \frac{E(1-E)^2}{2-E}\right)$	10

### 3.1 実験設定

ベータ分布は  $x \in [0, 1]$  で定義され比較的計算が容易な確率分布である (式 (9)),  $a, b$  は正の実数). そこで本実験では, ワーカのタスク正解率の分布としてベータ分布を仮定し, ベータ分布の平均  $E$ , 分散  $V$  及びネットワークの次数  $m$  を変化させた場合の REC と従来法のタスク正解率の平均値の関係を調べる. ベータ分布のパラメータ  $a, b$  はベータ分布の平均, 分散を用いて式 (10), 式 (11) のように表される.

$$\text{Beta}(x|a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}, \quad (9)$$

$$a = \frac{E^2(1-E) - EV}{V}, \quad (10)$$

$$b = \frac{(1-E)^2E - (1-E)V}{V}. \quad (11)$$

ベータ分布は単峰性になる  $1 < a, 1 < b$  の範囲と, 二峰性となる  $1 > a, 1 > b$  の範囲とで分布の形が大きく異なる. そこで本実験では単峰性の範囲について実験を行う (但し二峰性となる範囲に関しても同様に実験を行うことができる). ベータ分布が単峰性になる範囲は式 (10), 式 (11) により分散  $V$  の範囲として書き表すことができる (式 (12)).

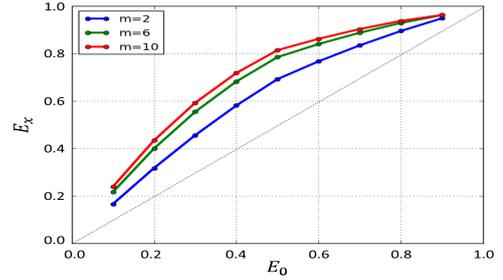
$$0 < V < \min\left(\frac{E^2(1-E)}{1+E}, \frac{E(1-E)^2}{2-E}\right) \quad (12)$$

下記の実験では知人ネットワークの次数が  $m \in \{2, 6, 10\}$  の場合について, それぞれベータ分布の平均 (ワーカのタスク正解率の平均)  $E$  を  $[0.1, 0.9]$  の範囲で 0.1 間隔 9 個, 分散  $V$  を式 (12) の範囲で等間隔 10 個, 計 90 個のパラメータについて計算し, REC と従来法のタスク正解率の期待値の比較を行う (表 1).

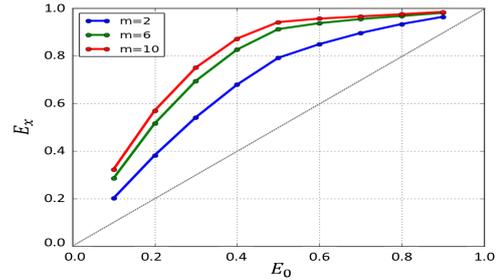
### 3.2 実験結果

知人ネットワークの次数が  $m \in \{2, 6, 10\}$  の場合について, ワーカのタスク正解率の期待値と REC におけるタスク正解率の期待値を図 2 に示す. 図 2(a), 図 2(b), 図 2(c) はそれぞれ, ワーカのタスク正解率の分散  $V$  を式 (12) の範囲で変化させた時の REC のタスク正解率の平均期待値, 最大期待値, 最小期待値を表している. 図 2(a) から従来法に比べ REC によりタスクの正解率の平均期待値が次数によらず上昇していることが分かった. 特に次数が 10 の場合にはワーカのタスク正解率の期待値が 0.2 から 0.7 の範囲では平均期待値が 0.2 以上向上することが分かった. また平均期待値は次数が増加するほど上昇することを確認した. さらに図 2(b) からワーカのタスク正解率の期待値が 0.4 以上では REC により平均期待値が最大で 0.85 以上にまで向上する可能性があることが分かった. しかしながら図 2(c) から従来法に比べ正解率の期待値がほとんど向上しない場合があることが明らかになった.

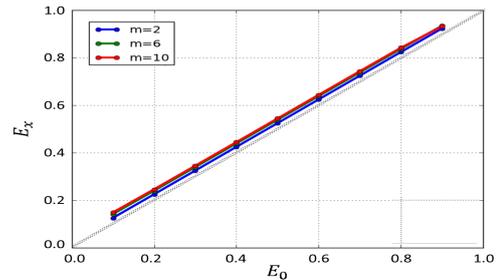
知人ネットワークの次数が 10 の場合のワーカの正解率の分散と REC によるタスク正解率の平均期待値の増加量の関係を図 3 に示す. この図からワーカのタスク正解率の期待値が小さ



(a) 平均値



(b) 最大値



(c) 最小値

図 2: ワーカのタスク正解率の期待値と REC によるタスク正解率の期待値の関係

いほど REC による平均期待値の増加量が多いことが分かる. また正解率の分散が大きいほど REC による品質向上が期待できることが分かり, 図 2(c) のような状況はワーカの正解率にほとんどばらつきが無い場合に起きることが明らかになった.

## 4. 関連研究

クラウドソーシングにより収集されたデータの品質向上に関する研究は数多く行われている. Sheng らはラベル付けタスクにおいて, 1 つのデータに対して複数のワーカがラベル付けを行った結果から真のラベルを推定することによりラベルの推定精度を向上させる手法を提案した [Sheng 08]. 櫻井らは真の作業結果と真の主観的な自信を申告することが最適戦略となるメカニズムについて提案し, 主観的な自信を用いることにより品質が向上することを示した [櫻井 12]. Whitehill らはワーカの能力に加え, 問題の難易度をモデル化することで収集した結果から品質向上が可能であることを示した [Whitehill 09]. これらの研究では収集された結果からいかに精度の高いラベルを推定するかを対象としており, 収集される結果自体の品質向上については扱われていなかった.

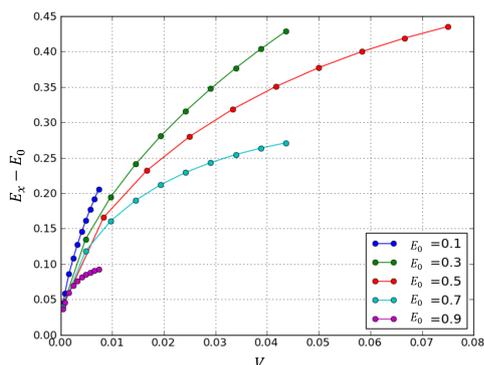


図 3: ワーカーの正解率の分散と REC によるタスク正解率の平均期待値の増加量の関係

ソーシャルネットワークを活用したクラウドソーシングもいくつか提案されている。例えば Pickard らは参加者の増加とタスクを処理する動機を両立させる報酬の与え方として Recursive Incentive Mechanism (以降, RIM) を提案した [Pickard 10]. この論文ではタスクを受け取った人はタスクを処理しかつ知人全員に紹介することが合理的な選択であるため, 効率的にタスク情報が伝播, 処理されることが示された。しかしながらワーカーが合理的である場合には全員がタスクを処理するため, 収集したデータの品質は向上せず, 関わるワーカーが大幅に増加するために膨大な確認作業が必要となるという問題があった。Naroditskiy らは RIM で集めた大量の処理結果に対しての確認作業を, 処理したワーカーにタスクを紹介したワーカーに課すことにより, 処理結果の品質を向上させる手法を提案した [Naroditskiy 12]. 彼らはタスク正解率が一定の場合についてこの仕組みが上手く機能するための, 確認作業の失敗に対してのペナルティ, 確認コスト, タスク正解率の関係を明らかにした。

## 5. 結論

本論文では, 再帰的にタスクを委託することが出来るクラウドソーシング REC を提案し, 次数  $m$  の木の場合の正解率の分布を解析的に示した。ワーカーのタスク正解率の分布としてベータ分布を仮定し, 従来型のクラウドソーシングと REC の平均正解率を比較した。その結果タスク委託機能を追加することにより従来法に比べ正解率の高い処理結果が得られることを確認した。

本論文では知人ネットワークとして木を仮定したが, 知人の数が一定であることや, コミュニティが存在しないなど実世界のソーシャルネットワークとは異なる点がある。そのため今後はより実世界の知人ネットワークに近い性質をもつ一般グラフや Watts らによって提案された WS モデル [Watts 98] において REC の有効性を検証していく予定である。また単純化のためワーカーが選択できる行動をタスクの処理とタスクの委託のみに限定していた。しかしながら実際にはある割合でタスクを委託されても処理しないというワーカーが存在することが知られている [Dodds 03]. そのためこのようなワーカーが存在する場合での REC の有効性についても今後検証していく予定である。さらには報酬の与え方に関しても, 簡単のためタスクに正解したワーカー及びそのワーカーにタスクを委託したワーカーに同

額の報酬を支払うことを仮定していた。しかしながら, これはワーカーがボランティアとして働いている場合を除いて, あまり現実的では無いため, より現実的な報酬を仮定した場合についても検証していく必要がある。本報告ではシミュレーションにより REC の有効性を検証するに留まっているため, 今後は実際の環境で実験を行い検証していく予定である。

## 参考文献

- [Dodds 03] Dodds, P., Muhamad, R., and Watts, D.: An experimental study of search in global social networks, *science*, Vol. 301, No. 5634, pp. 827–829 (2003)
- [Howe 06] Howe, J.: The Rise of Crowdsourcing, in *Wired Magazine* (2006)
- [Naroditskiy 12] Naroditskiy, V., Rahwan, I., Cebrian, M., and Jennings, N.: Verification in referral-based crowdsourcing, *PLoS ONE* (2012)
- [Pickard 10] Pickard, G., Rahwan, I., Pan, W., Cebrian, M., Crane, R., Madan, A., and Pentland, A.: Time Critical Social Mobilization: The DARPA Network Challenge Winning Strategy, *arXiv preprint arXiv:1008.3172* (2010)
- [Sheng 08] Sheng, V., Provost, F., and Ipeirotis, P.: Get another label? improving data quality and data mining using multiple, noisy labelers, in *Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 614–622 (2008)
- [Watts 98] Watts, D. and Strogatz, S.: The small world problem, *Collective Dynamics of Small-World Networks*, Vol. 393, pp. 440–442 (1998)
- [Whitehill 09] Whitehill, J., Ruvolo, P., Wu, T., Bergsma, J., and Movellan, J.: Whose vote should count more: Optimal integration of labels from labelers of unknown expertise, *Advances in Neural Information Processing Systems*, Vol. 22, pp. 2035–2043 (2009)
- [櫻井 12] 櫻井 裕子, 沖本 天太, 岡 雅晃, 兵藤 明彦, 篠田 正人, 横尾 真 クラウドソーシングにおける品質コントロールの一考察, 合同エージェントワークショップ&シンポジウム 2012(JAWS2012) (2012)