

マイクロタスク型クラウドソーシングにおける 良質な訓練用タスクの自動生成

根本 千代之介

計算機ネットワークの発達により、クラウドソーシングが注目を集めるようになってきた。クラウドソーシングとは不特定多数の群衆(クラウド)に仕事(タスク)を委託することで問題を解決するアプローチである。タスクを処理する個々の人はワーカと呼ばれる。

クラウドソーシングにおける問題の一つは、タスク結果の品質が必ずしも高くはないということである。タスク結果の品質を下げる要因の一つとして、タスクに対する誤解が挙げられる。タスクに対する誤解とは、ワーカがタスク作成者の意図を取り違え、タスク作成者の意図から外れた回答をしてしまうことである。

誤解による品質低下を防ぐ手法として、ワーカに事前に訓練を行わせてからタスクを処理させる手法がすでに存在する。この手法では、事前にワーカに訓練用のタスクの集まり(訓練用タスク列)を処理させ、その解答を確認させることで訓練を行い、その後通常のタスクを処理させる。しかし、訓練用タスクは本来処理する必要のないタスクであり、ワーカにとって余計な労力となる。したがって、限られたタスク数で、より多くのタスクを正しく処理できるようになる訓練用タスク列(良質な訓練用タスク列)の方が望ましい。

以上のような背景から、本論文では良質な訓練用タスク列を自動生成するための二つの手法を提案した。提案手法 1 はタスクの処理結果に対する自信のありなしを利用する手法である。具体的には、ワーカにタスクを処理させる際、回答に対する自信のありなしを「ある」と「ない」の 2 択で入力させ、「自信がない」とされたタスクから訓練用タスクを選択する。提案手法 2 は「自信がない」と回答されたタスクから、訓練用タスク列の中に様々なパターンのタスクが含まれるようにタスクを選択する手法である。具体的には、「自信がない」とされた全てのタスクを順に 3 人のワーカに処理させ、誰か 1 人でも正しく回答できなかったタスクは新たなパターンのタスクと考え、訓練用タスクに採用する。

評価実験では、二つの提案手法と、全タスク中から無作為に訓練用タスクを選択する手法で合計三つの訓練用タスク列を作成し、各訓練用タスク列について品質を比較した。

実験の結果、「自信がない」とされたタスクを訓練用タスクに採用することで、無作為に作成したものより高い品質を持つ訓練用タスク列を作成できることが判明した。一方、提案手法 1 と提案手法 2 では品質に顕著な差が見受けられなかった。

(指導教員 森嶋厚行)