

スキルの異なるワーカを組み合わせた 書誌誤同定判定クラウドソーシングの品質向上

あべ松 幸彦

近年、総合目録の構築等において、同一の資料に対応する複数の書誌データを機械的に同定する機械的書誌同定が行われている。しかし、機械的書誌同定を行った結果、異なる資料に対応する書誌データを誤って同定してしまう誤同定が発生している。この誤同定を発見するために、クラウドソーシングを用いて書誌誤同定判定を行うためのマイクロタスクを設計し、不特定多数の群集（ワーカ）に処理を行ってもらった。先行研究が行われた。

しかし、先行研究の実験結果には、書誌誤同定判定タスクを処理したワーカが、書誌に関する知識を持っていないために間違っただ判定をしているタスクが多く見られた。このような間違っただ判定を減らし、タスク処理結果の品質を向上させるため、本論文では、書誌に関する知識を持たない一般のワーカだけでなく、書誌に詳しいワーカに書誌誤同定判定タスクの処理を依頼することを提案する。

問題は、書誌に詳しいワーカは一般のワーカに比べて人数に限りがあることである。したがって、書誌に詳しいワーカに処理を依頼するタスクを選定する必要がある。しかし、書誌に詳しいワーカに処理してもらったタスクを選定する基準は自明ではなく、またそれらの基準をどのように用いれば、効率的に誤同定を発見できるかも自明ではない。

本論文では、依頼するタスクを選定する複数の基準の比較実験を行った結果を報告する。具体的には、下記の基準を利用してタスクの選定を行った。(1) 少なくとも一人の一般のワーカが書誌誤同定判定に対して「自信がない」と回答した場合、(2) 二人以上の一般のワーカが「自信がない」と回答した場合、(3) 二人以上の一般のワーカが「自信がある」と回答したが同定・誤同定の判定がワーカによって異なる場合、(4) 自信の有無に関わらず、二人以上の一般のワーカによる同定・誤同定の判定が異なる場合。

比較実験としては、あらかじめ行ったクラウドソーシングタスクの結果を利用して、書誌に詳しいワーカが各基準に基づき依頼されたタスクを処理するシミュレーションを行った。次に、そのシミュレーションの結果を、(a) 処理結果の精度と再現率、(b) 依頼したタスク数とシミュレーションによって得られた処理結果に含まれる誤った判定の数、(c) タスク全体にかかるコストの数値で比較した。その結果、いずれの基準においても処理結果のF値は0.9以上であるが、依頼したタスク数が増加すると処理結果に含まれる誤った判定の数が減少することを確認した。各基準を比較すると、特別に優秀な結果を出す基準はないものの、それらの基準を利用する方法を工夫すると、F値を維持したまま、タスク全体にかかるコストを押さえることができる事がわかった。

(指導教員 森嶋厚行)