

モバイルクラウドソーシングにおけるワーカコンテキストの影響性評価

Effect of Situational Factors on Mobile Crowdsourcing

池田 和史*¹
Kazushi Ikeda

帆足 啓一郎*¹
Keichiro Hoashi

*¹ KDDI 総合研究所
KDDI Research

Mobile crowdsourcing platforms have enabled crowd workers to accept jobs wherever/whenever they are, and also provides opportunity for task requesters to order time/location specific tasks to workers. Since workers on mobile platforms are working on the go, the situation of the workers is expected to influence their performance. In this paper, we draw on economics and psychology research to examine whether situational factors affect task completion, price and quality on mobile crowdsourcing platforms. Our three-week between-subjects field experiment revealed that the performance of workers significantly differ in their situations.

1. 概要

近年、時間や場所といった条件を満たすワーカにモバイル端末を用いてタスクを実施させるモバイルクラウドソーシングに注目が集まっている。一般的なクラウドソーシングでは、ワーカは自身の都合で仕事をするのに対し、モバイル環境では条件を満たすワーカへタスクを能動的に通知する機能を有するものが多い。これにより、必ずしも作業に適した状況にないワーカもタスクを閲覧する点で、一般的なクラウドソーシングとは異なる。既存のモバイルクラウドソーシングにおけるモチベーションの研究 [Teodoro 2014] では、ワーカへのアンケートに基づいて、休日は時間的な余裕があるため、タスクを実施しやすいなど、状況がタスクの実施率に影響を与える可能性を報告しているが、定性的な評価に留まっており、定量的な評価は十分に行われていなかった。本稿では、**モバイルクラウドソーシングにおいて、ワーカの状況は性能に影響を与えるか?**という研究課題を設定し、経済学、心理学の知見を元に性能に影響を与えるワーカの状況の仮説を形成した。50 人の被験者による 3 週間の実証実験の結果、ワーカの状況は性能に有意に影響することが確認された。

2. 仮説形成

一般的にクラウドソーシングにおける性能の評価尺度として、タスクの実施率、報酬、作業品質などが用いられる [Mason 2009]。クラウドソーシングは報酬を伴う人の行動に関するものであるため、その理解には経済学や心理学の知見が参照されることが多い [Mason 2009]。本稿においても、同様のアプローチを用いてモバイルクラウドソーシングの性能に影響を与えるワーカの状況の仮説を形成する。

経済学における機会費用の考え方によれば、人は作業で得られる報酬と、その他の行動の選択肢を比較して優先度付けを行っている [Roberts 2007]。機会費用がクラウドソーシングに適用可能かを評価した研究報告は確認されていないが、ワーカが多忙である場合、タスクが実施されるには、多くの選択肢よりも優先される必要があるため、その機会費用は大きいと考えられる。そこで我々は次の仮説を設定した。

仮説 1: モバイルクラウドソーシングにおいて、ワーカが多忙と感じている場合、タスクの実施率は低下し、受託金額は上昇する。

心理学の分野では、人は精神的な疲労がある場合、モチベーションが低下し、注意が散漫になることが報告されている [Boksem 2006]。既存の研究では車の運転のような継続的な作業における疲労の影響を評価したものが多い。モバイルクラウドソーシングでは、特定の時間、場所で作業を行うような一度限りのタスクが想定されるが、我々はこのような環境であっても、ワーカの疲労が作業品質に影響するという仮説を設定した。

仮説 2: モバイルクラウドソーシングにおいて、ワーカが疲労を感じている場合、タスクの作業品質は低下する。

心理学の分野では、同行者がいる状況でモバイル端末を使用することが親近感や共感を低下させることが報告されている [Misra 2016]。また、学生を対象とした実験において、授業中のノートパソコンの利用が、集中力を低下させるとの報告がある [Hembrook 2003]。これらの知見から、クラウドソーシングにおいて、同行者の存在がタスクの実施を躊躇させたり、同行者とのコミュニケーションがタスクに対する集中力を低下させる可能性が考えられる。

仮説 3: モバイルクラウドソーシングにおいて、ワーカに同行者がいる場合、タスクの実施率は低下、受託金額は上昇し、作業品質は低下する。

3. 実験設計

設定した仮説を検証するため、被験者実験を実施した。タスクとしてワーカの現在地における携帯電話の通信品質調査を依頼した。ワーカが多忙度、疲労度がそれぞれ高い場合と低い場合、同行者がいる場合といない場合の性能を比較した。性能指標として実施率、受託金額、作業品質を評価した。

3.1 被験者

人材派遣会社に登録されたモニタ調査員から被験者 50 名を募集した (58% が女性、年齢は 20-58 歳、平均年齢 37.5、標準偏差 9.1)。被験者は職業を有し、居住地、勤務地は日本各地に及ぶ。参加の条件として Android スマートフォンまたは iPhone を保有する被験者を募集した。募集条件に属性を偏らせる要因はなく、属性によるフィルタリング等は実施していない。

3.2 タスクとアプリケーション

被験者は本実験用のアプリケーションを各自のスマートフォンにインストールする。アプリは不定期にタスクを端末上に通知する。本実験では、約 30 分間隔で一日 27 回通知されるよう、設定した。ワーカはタスクの金額を確認した後、実施するか否かを選択できる。ワーカがタスクを完了したかは、アプリのログから取得可能とした。タスクはモバイルクラウドソーシングでは一般的な通信品質の調査を採用した[Faggiani 2014]。ワーカは 3 分間の動画を視聴し、体感的な通信品質を報告する。実験を通してタスクの所要時間は中央値で 216 秒であった。報酬は 30 円(1 タスク 4 分換算で時給 450 円)を中心とする 10 円~50 円(同、時給 150 円~750 円)の 5 円刻みの一様分布とした。ワーカが多忙などの特定の状況で安価なタスクをより断る傾向があれば、当該状況における平均受託金額の上昇を確認できる。実験期間は 3 週間(21 日)であった。

3.3 比較条件

ワーカの多忙、疲労、同行者の有無といった状況はアンケートによって取得した。タスク実施時に、動画の品質と合わせて状況を報告させた。多忙度は「(a) とても忙しい」、「(b) どちらかといえば忙しい」、「(c) どちらかといえば忙しくない」、「(d) 全く忙しくない」の 4 段階で取得し、(a)、(b)を多忙、(c)、(d)を非多忙とした。疲労についても同様に疲労と非疲労を取得した。同行者については、あり、なしを取得した。

実施率の評価には、タスクを実施しなかった場合についてもワーカの状況を取得する必要がある。そのため、3 時間ごとに別途アンケートを通知し、直近の 3 時間でタスク通知があった時刻を提示して、当該時刻における多忙度、疲労度、同行者の有無を回答させた。実施率はこの定期アンケートに基づいて算出する。

3.4 評価指標

状況ごとに実施率、平均受託金額、作業品質を評価した。タスクの実施率は「完了したタスク数 / 通知したタスク数」と定義する。平均受託金額は、完了したタスクの受託金額の平均を状況ごとに算出する。タスクの作業品質は、正解が既知であるタスクを混在させる方法で評価する。意図的にノイズを混入した動画を 25%の確率で提示し、当該動画をノイズがあると回答すれば、正しく作業を実施したものと見なす。実際に通信品質の低下によって、動画にノイズが発生するケースを区別するため、アプリでは動画全体をバッファリングすることで、ノイズが発生しないようにした。評価指標としては、Recall, Precision, F measure, Accuracy を用いる。ノイズのある動画を Positive とし、TP (True Positive), TN (True Negative), FP (False Positive), FN (False Negative)を用いて、それぞれ次のように定義する。

$$Recall = TP / (TN + TP)$$

$$Precision = TP / (FP + TP)$$

$$F\ measure = 2 / (1 / Recall + 1 / Precision)$$

$$Accuracy = (TP + FN) / (TP + TN + FP + FN)$$

4. 実験結果

実験に参加した 50 名の被験者のうち、3 件以上のタスクを完了した 47 名を分析対象とした。これらの被験者によって実施されたタスクの総件数は 3,596 件であった。状況ごとのタスクの実施率、平均受託金額、作業品質について報告する。

状況	実施率(%)	絶対/相対差(%)	p 値
多忙	25.7	7.74/30.1	.000***
非多忙	33.4		
疲労	29.7	2.28/7.72	.755
非疲労	30.1		
同行者あり	28.2	5.03/17.9	.000***
同行者なし	33.2		

表 1 ワーカの状況ごとの実施率

状況	平均受託金額(円)	相対差	p 値
多忙	32.8	5.95	.005**
非多忙	30.9		
疲労	32.2	4.02	.054
非疲労	31.0		
同行者あり	32.8	7.67	.000***
同行者なし	30.5		

表 2 ワーカの状況ごとの平均受託金額

4.1 実施率

表 1 は状況ごとの実施率と絶対/相対差およびカイ二乗検定による有意差検定の p 値を表す。ワーカが多忙の場合、非多忙の場合と比較して、実施率は絶対値で 7.74%、相対値で 30.1% 低下し、カイ二乗検定によって有意な差があることが確認された ($\chi^2(1, N=8213) = 57.8, p < .001, \phi = 0.084$)。疲労度については、有意な差は観測されなかった。同行者がいる場合、いない場合と比較して、実施率は絶対値で 5.03%、相対値で 17.9% 低下し、有意な差があることが確認された ($\chi^2(1, N=8213) = 22.2, p < .001, \phi = 0.052$)。これらのことから、多忙なワーカや同行者のいるワーカにタスクを依頼した場合、タスクが実施される確率は大きく低下することが分かる。

4.2 受託金額

表 2 に状況と平均受託金額、受託金額の相対差、t 検定による有意差検定の p 値を示す。多忙の場合、非多忙の場合と比較して、平均受託金額は 5.95% 増加し、t 検定によって有意に差があることが確認された ($t(2808) = 2.29, p < .001$)。疲労度が高い場合、低い場合と比較してタスク金額が 4.02% 増加したが、 $p=0.0543$ と有意差があると判定するには至らなかった。平均受託金額に関しては、同行者がいる場合といない場合の差が最も大きく、同行者がいる場合 7.67% 増加した ($t(3561) = 3.43, p < .001$)。これらのことから、多忙なワーカや同行者のいるワーカにタスクを依頼した場合、安価なタスクは実施されにくくなることが分かる。

4.3 作業品質

表 3 は状況ごとの Precision, Recall, F 値, Accuracy と Accuracy に対するカイ二乗検定の p 値を示す。想定されるタスクの利用用途は、通信品質が悪い場所を発見し、改善を試みることであるため、Recall や Precision, F 値といった指標が重視される。一方で、これらは互いにトレードオフの関係にあることから、有意差検定は Accuracy に対してカイ二乗検定を用いて実施した。検出対象の動画が全体の 25%と少ない点、またノイズの発見には画面を集中して視聴する必要があることから、F 値は全体的に 0.2 から 0.4 と低く、難しいタスクであったと言えるが、状況によって精度に差が見られた。

状況	Prec.	Recall	F 値	Acc.	p 値
多忙	0.567	0.188	0.282	0.742	32.8
非多忙	0.461	0.261	0.333	0.748	30.9
疲労	0.478	0.184	0.265	0.730	32.2
非疲労	0.500	0.287	0.364	0.761	31.0
同行者あり	0.449	0.191	0.268	0.733	32.8
同行者なし	0.528	0.277	0.364	0.759	30.5

表3 ワーカの状況ごとの Precision, Recall, F 値, Accuracy

疲労の場合、非疲労の場合と比較して Precision, Recall 共に低下した。これは、本稿で仮説として設定したように、疲労時は生産性が低下し、判断を誤りやすくなるなど、作業に集中できていないことが原因と考えられる。F 値で 0.364 から 0.265 (相対値で 37.4%)と 3 つの状況の中で最も低下し、Accuracy に対する有意差も確認された ($\chi^2(1, N=3596) = 4.31, p < .05, \phi = 0.035$)。同行者がいる場合も疲労時と同様に Precision, Recall が共に低下し、F 値も低下した。同行者がいる場合、なるべく短い時間でタスクを実施しようとしていたり、同行者に気を取られてタスクに集中できない、といったことが原因と考えられるが、Accuracy に対するカイ二乗検定の p 値は 0.067 であり、有意な差を示すには至らなかった。多忙な場合、非多忙な場合と比べて Recall は低く、Precision が高くなるという結果が得られた。これはワーカが忙しいために、明確にノイズと分かるケースのみを検出し、動画を細部まで見なかったことが原因と推測される。

5. 議論

本稿では、モバイルクラウドソーシングの性能に影響を与えるワーカの状況として、多忙、疲労、同行者の有無を仮説として設定した。仮説 1 の多忙は 4.1 節と 4.2 節、仮説 2 の疲労は 4.3 節から影響を裏付けられた。仮説 3 の同行者の影響は、実施率と受託金額については、有意差が確認できたが、作業品質についてはカイ二乗検定の p 値が 0.067 と 0.05 よりもわずかに大きく、有意差を確認するに至らなかった。ワーカの状況が性能に与える影響について、考察を行う。

実施率については、既存の情報通知に関する研究において、モバイル端末の操作状況が通知の閲覧率を高める[Okoshi 2015]という報告がある。一方、タスクを実施するには瞬間的な状況に加えて、ワーカの多忙度や同行者の有無といった、タスクを完遂できる状況にあるかも重要であることが確認された。

受託金額については、多忙時は他に実施すべき事柄があるため、機会費用の考え方に基いて、金額が小さい場合はタスクが実施されにくくなったと考えられる。同行者との時間も同様に価値のあるものと捉えていると推測される。実験後の被験者へのアンケートにおいて、「安価なタスクを受託しなかったことがあるか」という質問を実施したところ、Yes: 50%, No: 50%と意見が別れた。Yes と回答した被験者の代表的な意見として、費用対効果が悪い、割に合わないといった意見が多く見られた。「忙しかったりする中でも 50 円なら優先するが、10 円だと他を優先することがあるため」など、機会費用を明確に意識した意見も見られた。一方で、No と回答した参加者の意見には、少しでも収入になるため、今対応可能かどうかで判断したため、といった意見があった。これらの意見からも、状況が安価なタスクを受けるかどうかに影響することが分かる。

作業品質については、既存の心理学の研究において、疲労度が高くなると、作業の精度が低下することは報告されているが[Boksem 2006]、自動車の運転のような同一の作業を繰り返し継続することによる疲労が対象とされていた。一方、モバイルクラウ

ドソーシングにおいて、タスクをワーカに通知するような環境において、タスクとは異なる要因で蓄積された疲労が作業品質に影響を与えるという知見は、我々の知る限り報告されていない。

本稿の実験では、ワーカの状況をアンケートによって取得した。実際にモバイルクラウドソーシングで利用するためには、ワーカのモバイル端末の使用状況やスケジューラからワーカの多忙度、疲労度、同行者の有無を推定することで、タスクに適した状況にあると判断されるワーカに優先的にタスクを割り当てることが可能と考えられる。

6. まとめ

本稿では、多忙度、疲労度、同行者の有無といったワーカの状況がタスクの実施率、受託金額、作業品質、といったモバイルクラウドソーシングの性能に与える影響について調査を行った。ワーカの状況については、経済学や心理学の知見を用いて、タスクの実施に影響を与え得る要因を仮説として設定した。設定した仮説を検証するため、50 人の被験者により、モバイルクラウドソーシング上で通信品質の調査タスクを実施した。3 週間の実験の結果、ワーカが多忙な場合、タスクの実施率は相対値で 30.1% 低下、同行者がいる場合、タスクの受託金額が 7.67% 上昇、疲労している場合、タスクの精度が F 値の相対値で 37.4% 低下することが確認された。

モバイルクラウドソーシングの普及に伴い、なるべく多くのタスクを高品質に実施するニーズの増加が予想され、本稿の知見は有用といえる。本実験では現在地で実施可能な単独のタスクを対象とした。今後の課題として、より多様なタスクや実環境への適用を考慮して、移動を伴うタスクや複数のタスクが選択できる場合に影響を与える要因の調査を計画している。

参考文献

- [Boksem 2006] Boksem, M. A., Meijman, T. F. and Lorist, M. M.: Mental fatigue, motivation and action monitoring, *Biological Psychology*, 2006.
- [Faggiani 2014] Faggiani, A., Gregori, E., Lenzini, L., Luconi, V. and Vecchio, A. Smartphone-based crowdsourcing for network monitoring: Opportunities, challenges, and a case study. *IEEE Communications Magazine*, 2014.
- [Hembrooke 2003] Hembrooke, H., and Geri G.: The laptop and the lecture: The effects of multitasking in learning environments, *Computing in Higher Education*, 2003.
- [Mason 2009] Mason, W. and Watts, D. J.: Financial incentives and the "performance of crowds", In *Proc. HCOMP*, 2009.
- [Misra 2016] Misra, S., Cheng, L., Genevie, J. and Yuan, M.: The iPhone effect: the quality of in-person social interactions in the presence of mobile devices, *Environment and Behavior*, 2016.
- [Okoshi 2015] Okoshi, T., Ramos, J., Nozaki, H., Nakazawa, J., Dey, A. K. and Tokuda, H. Attelia: Reducing user's cognitive load due to interruptive notifications on smart phones. In *Proc. PerCom*, 2015.
- [Roberts 2007] Roberts, R.: Getting the most out of life: The concept of opportunity cost, *The Library of Economics and Liberty*, 2007.
- [Teodoro 2014] Teodoro, R., Ozturk, P., Naaman, M., Mason, W. and Lindqvist, J.: The motivations and experiences of the on-demand mobile workforce, In *Proc. CSCW*, 2014.