

作業者の状況がモバイルクラウドソーシングに与える影響の評価

Effect of Worker Situation on Mobile Crowdsourcing Performance

池田 和史[▼] 帆足 啓一郎[▼]

Kazushi IKEDA Keiichiro HOASHI

モバイルクラウドソーシングでは、時間や場所といった条件を満たす作業者に対し、能動的にタスクを通知することへの需要が大きい。能動的な通知では、作業者は必ずしもタスクの実施に適した状況にない可能性があるが、作業者の状況がモバイルクラウドソーシングの性能に与える影響はこれまで十分に評価されていなかった。我々は、タスクの実施率、受託金額、作業品質は作業者の状況に依存するか？という研究課題を設定した。経済学や心理学の知見に基づき、作業者の多忙度、疲労度、同行者の有無がタスクの実施に影響を与えるという仮説を立てた。50人の被験者による3週間の実証実験の結果、作業者の状況が性能に有意に影響することが確認された。

Mobile crowdsourcing platforms have enabled 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. However, the effects of mobile worker situations to task performance is an uninvestigated area. In this paper, our research question is, “do worker situations affect task completion, price and quality on mobile crowdsourcing platforms?” We draw on economics and psychology research to examine whether worker situations such as busyness, fatigue and presence of companions affect their performance. Our three-week field experiment with 50 participants revealed that worker situations significantly affect their performance on mobile crowdsourcing platforms.

1. はじめに

近年、時間や場所といった条件を満たす作業者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。レストランのレコメンド[1]や、ニュース記事の素材となる写真やビデオの撮影[2]など、モバイルクラウドソーシングを活用した研究事例は増加している。一般的なクラウドソーシングでは、作業者は自身の都合で仕事をするのに対し、モバイル環境ではGigwalk[3]のように、条件を満たす作業者へのプッシュ通知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作業員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。しかし、モバイルクラウドソーシングにおいて、作業員の状況がタスクの実施に与える影響の評価は十分に

行われていなかった。

我々は、作業員の状況はタスクの実施率や受託金額、作業品質に影響を与えるか？という研究課題を設定した。経済学や心理学の知見に基づき、作業員の多忙度、疲労度、同行者の有無がタスクの実施に影響を与え得るという仮説を立てた。経済学における機会費用の考え方[4]によれば、作業員がタスク以外に優先度の高い行動の選択肢を有している（すなわち多忙な）場合、より高い金額を提示しなければ、タスクを実施しないと考えられる。心理学の分野では、作業員の精神的な疲労がタスクの品質に影響を与えることが報告されている[5]。また、同行者がいる場合、モバイル端末を操作することに対する心理的な障壁がある[6]との報告がある。

本稿では、これらの状況がタスクの実施に与える影響を評価するため、50人の被験者による3週間の実証実験を実施した。モバイル端末上で実施可能な時間制限のあるタスクを不定期に通知し、タスク通知時の被験者の状況（多忙度、疲労度、同行者の有無）をアンケートにより取得した。これにより、状況毎のタスクの実施率を取得した。状況と受託金額の関係を評価するため、タスク金額が一定の金額を中心とする一様分布となるよう設定した。被験者は一般に安価なタスクほど拒否する傾向があるが、多忙などの特定の状況下でより安価なタスクを拒否する傾向があれば、当該状況において平均実施金額が相対的に高くなることを観測できる。状況と作業品質の関係を評価するため、正解が既知であるタスクを用いて、状況ごとの精度を評価した。実験の結果、タスクの実施率は作業員の多忙度に最も大きく影響され、多忙度が高い場合、タスクの実施率は相対値で30.1%低下した($p < .001$)。受託金額は同行者の有無に最も大きく影響され、同行者がいる場合、受託金額が7.67%上昇した($p < .001$)。作業品質は疲労度に最も大きく影響され、精度はF値で37.4%低下した($p < .05$)。加えて、被験者に対するアンケートで、我々の仮説を裏付ける意見が得られた。

本稿の構成は次の通りである。始めに、2章でクラウドソーシングおよびモバイル関連の研究を整理する。次に、3章で既存の経済学、心理学の知見を元に仮説を設定する。4章で実験の設計を説明し、5章で実験結果を示す。6章で実験結果の議論と今後の課題を示す。

2. 関連研究

2.1 クラウドソーシングの性質

クラウドソーシングにおけるタスクの実施率と金額、品質に関する既存の研究を紹介する。作業員にとって最も高いモチベーションはお金を得ることといわれる[7]ように、タスクの実施率と金額は強く関連しており、最も単純にタスクの実施率を増やす方法は、より多くの金額を提示することとされる[8]。タスクの金額を増やすことなく、実施率を向上させるアプローチとして、自身の貢献がどれだけ世の中に役に立つのかを説明する手法[9]や、作業員の好奇心を引き出す手法[10]などが提案されている。

タスクの品質管理に関しては、複数の作業員の回答が一致した場合にのみ報酬を支払う手法[11]や、作業員を訓練するためにフィードバックを与える手法[12]などが提案されている。品質の評価では、回答が既知であるタスクを混入させる手法[13]が広く用いられる。また、タスクの品質はタスクの金額とは関連がないと言われている[8]。

▼ 正会員 KDDI 総合研究所 kz-ikeda@kddi-research.jp

▼ 正会員 KDDI 総合研究所 hoashi@kddi-research.jp

2.2 モバイルクラウドソーシングと通知

モバイル環境を想定したクラウドソーシングの研究は近年注目を集めており[14]、様々なタスクが扱われている。シンプルなタスクとしては、モバイル端末を汎用なインターフェースとみなし、書き起こしを依頼するものがある[15]。より高度で状況に依存するタスクとして、特定の場所でレストランをRecommendするもの[1]や、ニュース記事の作成を効率化するために現地に居合わせた作業者に写真やビデオを撮影させるもの[2]などが提案されている。場所や時間を特定したタスクにおいては、条件を満たす作業者に対して、能動的に通知を行いたい場合も多いが、多忙や疲労といった作業者の内面的な状況を考慮してタスクを依頼するような提案は行われていない。

HCI分野ではモバイル端末向けの情報提示方法に関する研究が行われており、端末から取得可能なユーザの状況に基づいて、通知に適したタイミングを予測する手法が提案されている[16]。一般的に、通知は煩わしいものとなりがちであるが、通知される情報がユーザにとって関連性の高いものであれば、望ましく感じるという調査がある[17]。また、煩わしくても能動的に情報を取得するよりも通知されることを望むという調査も報告されている[18]。

3. タスクに影響を与える心理的な要因

3.1 研究課題

既存の通知に関する研究では、情報の閲覧率を向上させるための方式が多く提案されている。閲覧自体が目的となる広告や通知には有効であるが、タスクの実施率や受託金額、作業品質といったクラウドソーシングにおける性能を向上する観点では不十分といえる。作業者は現在地から作業場所までの距離を考慮して、タスクの金額が見合うかどうかを判断する[19]といった報告はあるが、作業者の内面的な状況がタスクの実施に与える影響は評価されていない。上記のことから、我々は次の研究課題を設定した。

課題: モバイルクラウドソーシングにおいて、タスクの実施率、受託金額、作業品質は作業者の状況に影響されるか？

3.2 仮説

経済学や心理学の分野の研究では、人の経済的な価値観や心理的な状態といった内面的な要因が作業の実施や効率性に与える影響について言及されている。これらの知見を元に、モバイルクラウドソーシングに影響を与え得る状況の仮説を立てる。

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[4]。一般に機会費用は収入に基づいて算出される[20]、旅行のような人の様々な活動について値付けを行うことができる[21]。機会費用がクラウドソーシングに適用可能かどうかを評価した研究報告は確認されていないが、作業者が多忙である場合、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、作業者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

[5]。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。例えば、Campagneらの研究では、運転手の疲労と速度の分散や車線のはみ出しに関連があることを報告している[22]。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、作業者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、作業者が疲労を感じている場合、タスクの作業品質は低下する。

同行者の存在がモバイル端末を使用する上で心理的な障壁になることが報告されている。Misraらは、テーブルにモバイル端末を置いた状態でカップルに会話をさせる実験を通して、モバイル端末の存在が「親近感」や「共感」を低下させることを確認した[6]。クラウドソーシング作業も同行者が存在する場合に、タスクの実施を躊躇する可能性が考えられる。また、同行者とのコミュニケーションとタスクを同時に実行することは、作業の品質を低下させる可能性もある。学生を対象とした実験において、授業中にノートパソコンを利用させることが、学生の集中力と試験の成績を低下させたと報告されている[23]。これらの知見を元に、我々は次の仮説を設定した。

仮説3: モバイルクラウドソーシングにおいて、作業者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。

4. 実験設計

作業者の状況がタスク実施に影響を与えるかどうかを検証するために被験者実験を実施した。50人の被験者が3週間(2016年8月8日~28日)の実験に参加した。被験者にはクラウドソーシングを実施するためのスマートフォンアプリを配布した。アプリは、一日に複数回タスクを通知する。被験者は自身の状況に応じて、タスクを実施してもよいし、無視してもよい。作業者がタスクを実施した場合、その時の状況として、多忙度、疲労度、同行者の有無をアンケートによって取得した。また、タスクを通知したが、実施しなかった時間帯についても、定期的にアンケートを通知し、当該時間の状況を振り返って取得した。タスクの金額は毎回異なり、設定した平均金額を中心とする一様分布とした。タスクには、正解が既知のデータが含まれており、作業者の回答精度を評価できるようにした。これらのデータを元に、我々は状況ごとのタスクの実施率、平均受託金額、作業品質を評価した。

4.1 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を募集した(N=50, 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のスマートフォンまたはiPhoneを保有する被験者を集めた。それ以外に属性を偏らせる要因はなく、募集時も属性によるフィルタリングは実施しなかった。

4.2 タスクとアプリケーション

ログの取得や通知条件を柔軟に設定できるように、実験には我々が独自に開発したクラウドソーシングアプリケーションを利用した。アプリはインストールされたモバイル端末に対し、指定された時刻にタスクを通知する機能を有する。タ

表 1 実験設定の概要

Table 1 Summary of experimental settings

タスク種別	通信品質調査
所要時間	216 秒 / タスク (中央値)
通知頻度	27 回 / 日 (7AM-8:30PM)
報酬金額	10~50 円 / タスク (平均 30 円, 5 円刻みの一様分布)
評価指標	実施率, 平均受託金額, タスク品質
被験者数	50 人
実験期間	3 週間 (2016 年 8 月 8 日~8 月 28 日)

スクは通信品質に関する調査として、現在地における体感的な通信品質をアンケートで回答させた。通信品質の調査は、モバイルクラウドソーシングでは一般的なタスクとされている [24], [25]。アプリは不定期にタスクを通知し、作業者はタスクの金額を確認した後、タスクを実施するか否かを選択することができる。各タスクについて、作業者がタスクを完了したか否かをアプリのログ情報から把握することを可能とした。

実験設定の概要を表1に示す。タスクの所要時間は中央値で216秒、通知は一日27回、約30分間隔で不定期に行われた。通知からタスクを受託するまでの期限は5分とし、タスクの受託後、実施完了の期限は15分に設定された。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。被験者は50人、実験期間は3週間(21日)であった。

通信品質評価タスクでは、作業者は3分間のストーリーミング動画を視聴し、視聴環境における動画品質に関するアンケートに答えるというタスクとした。クラウドソーシングを利用したストーリーミング動画のQoE (Quality of Experience) 評価はよく知られたタスクの1つである [26]。アンケートでは、通信品質に関する質問と作業者の現在の状況に関する質問を実施した。表2に質問項目を示す。通信品質に関する質問では、通信品質の劣化により動画が停止した回数と、動画にノイズが発生した回数を報告させた。加えて、通信品質に影響を与える現在の状況に関する質問として、移動手段を選択形式で取得し、現在地をフリーテキストで入力させた。作業者の状況については、多忙度、疲労度、同行者に関する質問を行うことで取得した。

作業品質を評価するため、正解が既知であるタスクを混在させる方法を用いた。作業者に提示される3分間の動画は全て同じ内容であるが、25%の確率で意図的にノイズを混入させた動画を提示した。動画中の異なる時間に約15秒のノイズを発生させた5種類の動画を用意し、ランダムで提示した。これらの動画を提示された作業者が表2のアンケートの1-2でノイズありと回答すれば、正しくタスクを実施したといえる。一方、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再生方法とした。通信品質が低下した場合、動画の再生がバッファに追いつき、再生が停止し、以降の部分は再生されないようにした。これにより、通信品質が悪い場合、1-1で「(f)動画が停止して終了」が選択されるため、意図的にノイズを混入した動画と区別することができる。

表 2 アンケートタスクの質問項目

Table 2 Task description

<p>1. ビデオを視聴して次の設問に回答してください。</p> <p>1-1 動画が止まることは何回ありましたか。 (a) なし (b) 1回 (c) 2回 (d) 3回 (e) 4回以上 (f) 動画が停止して終了 (15秒経過しても再開しない)</p> <p>1-2 画像が乱れることは何回ありましたか。 (a) なし (b) 1回 (c) 2回 (d) 3回 (e) 4回以上</p>
<p>2. あなたの現在の状況について教えてください。</p> <p>2-1 現在の移動状況を教えてください。 A. <input type="radio"/> 移動中 <input type="radio"/> 滞在中 B. <input type="radio"/> 電車内 <input type="radio"/> 屋外 <input type="radio"/> 屋内</p> <p>2-2 現在地の詳細を教えてください(文章入力) e.g. 「○○駅の××線ホームにいます。」</p> <p>2-3 多忙度を教えてください。 (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</p> <p>2-4 疲労度を教えてください。 (a) とても疲れている (b) どちらかといえば疲れている (c) どちらかといえば疲れていない (d) 全く疲れていない</p> <p>2-5 同行者を教えてください。 (a) 自分一人 (b) 家族 (c) 知人・友人 (d) 仕事の関係者 (e) その他 (フリー記述)</p>

4.3 作業者の状況

多忙度、疲労度、同行者は表2の2-3~2-5に記載の質問と選択肢により取得する。多忙度は4種類の選択肢で取得するが、分析時は「(a) とても忙しい」、「(b) どちらかといえば忙しい」を多忙、「(c) どちらかといえば忙しくない」、「(d) 全く忙しくない」を非多忙として扱う。同様に疲労度も疲労と非疲労、同行者は同行者あり、なしの2値として扱う。

タスク実施率の評価には、タスクを実施しなかった場合についても状況を取得することが必要となる。そのため、3時間ごとに作業者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、同行者の有無を表2の2-3~2-5の質問を行うことで取得した。タスク実施率はこの定期アンケートをベースに分析する。

4.4 評価尺度

状況ごとにタスク実施率、平均受託金額、作業品質を評価した。タスク実施率の算出には定期アンケートの回答を利用した。平均受託金額と作業品質については、タスクを完了した場合のみの情報から取得できることから、タスク内で取得したアンケートの情報を用いる。タスクの実施率は次のように定義し、状況ごとにタスクの実施率を算出する。

$$\text{実施率} = \text{完了したタスク数} / \text{通知したタスク数}$$

平均受託金額は、完了したタスクのみを集計対象とし、状況ごとに受託金額の平均を算出する。タスクの金額は30円を中心とする一様分布に設定されており、特定の状況下の作業者が安価なタスクを断る傾向があれば、受託したタスクの平均金額は高くなる。

表3 作業者の状況とタスク実施率

Table 3 Task completion rate across worker situations

状況	実施率(%)	絶対/相対差(%)	p 値
多忙	25.7	7.74 / 30.1	.0000***
非多忙	33.4		
疲労	29.7	2.28 / 7.72	.7554
非疲労	30.1		
同行者あり	28.2	5.03 / 17.9	.0000***
同行者なし	33.2		

タスクの作業品質は、意図的にノイズが混入された動画の検出精度を評価することで、作業者の注意力を評価する。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。

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

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

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

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

上記の定量評価に加えて、実験終了後に、実験参加者に対してタスクの実施と状況の関係性に関するアンケートを実施し、意見を分析した。具体的には、「タスクを受けにくいのはどのような状況のときか」、「タスクに集中しにくいのはどのようなときか」、「安価なタスクを意図的に受託しなかったことはあるか」と、それぞれの理由についてアンケートを実施した。

5. 実験結果

50名の被験者が実験に参加した。被験者は一日27件のタスク通知を受け、21日間で567件のタスク通知を受けた。その中で、3件以上のタスクを完了した被験者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。

5.1 タスク実施率

始めに、作業者の状況とタスクの実施率の関係について評価をする。タスクの実施率はタスクを完了した時と完了していない時の両方の状況を取得する必要があるため、定期アンケートに基づいて分析を行った。定期アンケートの回答率は36%であった。定期アンケートとタスク内の両方で状況が取得できた場合について、多忙度、疲労度、同行者の有無の一致度を評価すると、いずれも90%以上であり、3時間程度であれば被験者は概ね状況を記憶していることが確認できた。

表3は状況ごとのタスク実施率と絶対/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。作業者が多忙の場合、非多忙の場合と比較して、タスク実施率は絶対値で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、タスク実施率は絶対値で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、多忙な作業者や同行者のいる作業者にタスク

表4 作業者の状況と平均受託金額

Table 4 Mean accepted task price across worker situations

状況	平均受託金額(円)	相対差(%)	p 値
多忙	32.8	5.95	.0050**
非多忙	30.9		
疲労	32.2	4.02	.0543
非疲労	31.0		
同行者あり	32.8	7.67	.0003***
同行者なし	30.5		

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。

5.2 受託金額

次に、作業者の状況とタスクの受託金額の関係について評価する。タスクの金額は30円を平均とする10円から50円の5円刻みの一様分布であるため、特定の状況において安価なタスクを断ると、当該の状況における平均受託金額が高くなる。タスクを完了した場合のみから分析可能なため、タスク内で取得した状況を利用した。表4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、多忙な作業者や同行者のいる作業者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。実施率と受託金額は共に、多忙度と同行者について有意差を観測した。この結果は、タスク金額と実施率に強い関連があるとするMasonらの研究結果とも合致する[8]。

5.3 作業品質

作業者の状況と作業品質の関係について評価する。本タスクでは、作業者は3分の動画を視聴し、動画の品質に関するアンケートに回答する。作業品質を評価するため、意図的に品質を劣化させた動画の検出精度について、Precision, Recall, F値, Accuracyを前述のとおり定義した。評価結果を表5に示す。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。状況によって精度には差が見られた。F値を基準に見ると、多忙な場合や疲労している場合、同行者がいる場合は、それぞれ非多忙、非疲労、同行者なしの場合と比較して、いずれもF値は低くなった。

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

表5 作業者の状況とPrecision, Recall, F値およびAccuracy
Table 5 Recall, precision, F measure and accuracy of the tasks across the conditions

状況	Prec.	Recall	F 値	Acc.	p 値
多忙	0.567	0.188	0.282	0.742	.7193
非多忙	0.461	0.261	0.333	0.748	
疲労	0.478	0.184	0.265	0.730	.0379*
非疲労	0.500	0.287	0.364	0.761	
同行者あり	0.449	0.191	0.268	0.733	.0730
同行者なし	0.529	0.277	0.364	0.759	

表6 Q1:タスクを受けにくい, Q2:タスクに集中しにくい, について各状況の回答数と比率

Table 6 The number and ratio of situations where workers feel difficulties to Q1: accept tasks and Q2: concentrate on tasks

状況	Q1 回答数(%)	Q2 回答数(%)
多忙	21 (50%)	21 (50%)
疲労	4 (10%)	4 (10%)
同行者	10 (24%)	15 (34%)
その他	7 (17%)	2 (5%)

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作業を実施しようとしたり、同行者に気を取られて作業に集中できない、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。これは作業者が忙しいために、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。

5.4 被験者意見の分析

タスクの実施率と精度について、被験者の主観的な意見を収集するため、3週間の実験期間終了後にアンケートを実施した。アンケートでは、「Q1. 3つの状況(多忙, 疲労, 同行者)の中で最もタスクを受けられないと感じるのはどのようなときか」と「Q2 3つの状況の中で最もタスクの内容に集中しにくいのはどのようなときか」を選択させ、それぞれの理由について尋ねた。50人の被験者のうち、42人(84%)から回答を得た。表6にQ1, Q2の回答数とその比率を示す。

表6によれば、最もタスクを受けられないのは多忙時であり、二番目は同行者がいる時であった。これはタスクの受託率が多忙時に最も低く、同行者がいる時が2番目に低いという実験結果に合致している。多忙時にタスクを受けられない理由としては、「仕事中で携帯端末を開くことができない」という意見が最も多く、「忙しくて3分の時間を作るのが難しい」といった意見もあった。同行者がいる時にタスクを受けられない理由としては、「相手に失礼だから」という理由が最も多かった。これは我々が仮説形成時に参照したMisraらの調査[6]がモバイルクラウドソーシング上でも有効であることを示唆している。

表7 作業者の状況がタスク実施へ与える影響の仮説と実験結果の対応 (- は仮説設定なし, n.s.は有意差なし)

Table 7 Correspondence matrix of hypotheses and experimental results about the relationships between worker situations and task performance

状況	条件	実施率	受託金額	作業品質
多忙	仮説	低下	上昇	-
	結果	低下	上昇	n.s.
疲労	仮説	-	-	低下
	結果	n.s.	n.s.	低下
同行者	仮説	低下	上昇	低下
	結果	低下	上昇	n.s.

最もタスクに集中しにくいと回答されたのも同様に多忙時で、二番目は同行者がいる時となった。実験において実際にタスク精度が最も低下したのは、疲労時であったことから、作業者の感覚と実際の行動が異なったものと考えられる。多忙を選択した理由としては、「忙しくて動画を見続けることができない。」といった意見が多かった。同行者を選択した理由としては、「同行者に気を使って早く終わらせようとする」といった理由が多く、いずれも精度の低下に影響を及ぼすと考えられる。疲労を選択した作業者の絶対数は少なかったが、「眠くて最後まで集中して見られない」といった意見があった。これらのことから、疲労による作業品質の低下は作業者自身では気づきにくい可能性が示唆された。

6. 実験結果の議論と今後の課題

本稿における我々の研究課題は、モバイルクラウドソーシングにおけるタスク実施率、受託金額、作業品質は作業者の状況によって異なるか? というものであった。この研究課題に対し、経済学や心理学の知見から、タスクの実施に影響を与え得る3つの状況(多忙度, 疲労度, 同行者の有無)を仮説として設定した。これらの仮説と実験結果の対応を表7にまとめる。

表7から、本稿で設定した仮説の多くが実験によって裏付けられたといえる。同行者がいる際の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。作業者の状況がタスクの実施率、受託金額、作業品質に与えた影響について、詳しく考察を行う。

タスクの実施率については、既存の情報通知に関する研究では、ユーザが通知を閲覧するかどうか、という観点でモバイル端末のログを解析することで、通知に適したタイミングを探すとといった研究が行われている。そのような目的に対しては、モバイル端末上のアプリの切り替えのタイミングなど、瞬間的な状況が重要となる。一方、本稿ではモバイルクラウドソーシングで3分程度の時間を要するタスクを実施させる場合を想定した。被験者の意見にもあったが、3分の作業を実施するには、多忙時や同行者がいる状況では、作業を完遂できない可能性があるため、実施を躊躇する傾向があることが確認された。

受託金額はタスクの実施率との関係性は強いとされる[8]のように、実施率に影響する多忙度、同行者が受託金額にも影響することが確認された。多忙時は他に実施すべき事柄があるため、機会費用の考えに基づいて、金額が小さい場合はタスクが実施されにくくなると考えられる。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。実験後のアンケートで「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と回答した作業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[5]が、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、タスクを作業者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。

本研究で得られた知見の活用方法について検討する。実験結果から、作業者の状況はタスクの実施率、受託金額、作業品質に影響を与えることが確認できた。実験では、作業者の状況をアンケートによって取得したが、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、作業者の状況を取得する手段が必要となる。センサデバイスや端末の通信状況からユーザの多忙度や疲労度、同行者を推定することも可能と考えられるが、現状では精度面での課題もある。

我々は、定期アンケートで取得した作業者の多忙度、疲労度、同行者を集計した。図1に平日および休日、時間帯ごとの、多忙、疲労、同行者あり、に該当する作業者の割合をそれぞれ示す。図1によれば、平日、休日共に午前中の方が多忙や疲労の状況にある作業者が少なく、同行者がいる作業者の割合も低いことから、タスクを依頼するのに適した状況にあるといえる。平日の昼間は同行者がいる割合が多くなるが、休日の昼間は比較的少ない、といった傾向も見られた。図1は本実験に参加した被験者全体の傾向をまとめたものであるが、勤務の曜日や出勤時間などの生活パターンは人ごとに異なる。作業者ごとの行動ログを分析することで、現在時刻にタスクの実施に適した状況にあると推測される人に対してタスクの通知を行うことで、タスクの実施率、受託金額、作業品質を高められる可能性がある。実際に図1の下段にタスクの受託金額を示しており、作業者の状況と相関があるように見える。平日の多忙な作業者の割合と受託金額の相関係数は0.58、休日の多忙な作業者の割合と受託金額の相関係数は0.50とやや相関があることが分かる。

最後に、本研究の制約について述べる。本稿におけるタスクの一般性に関して、モバイルクラウドソーシングにおいて、通信品質調査を行うタスクは一般的であり[24]、[25]、通信品質の調査にストーリーミングを用いることもよく知られている[26]。一方、詳細なタスク設計は議論の余地があると考えられる。本稿では、タスクの所要時間を3分に設定したが、これは短時間で完了するタスクの場合、金額にかかわらず作業者がタスクを実施してしまう可能性があったためである。モバイル端末の通知を確認し、タスクの金額を確認するまでに実時間で数秒から十秒程度を要する。現在の作業を止めるためのスイッチングコストも大きいので、タスク自体が1分で

あっても、切り替え等に要する時間が埋没コスト(sunk cost)として計算され、タスク金額にかかわらず、実施したほうが良いという判断に至る可能性があった。

作業者の状況については、アンケートによって入手したため、多忙や疲労は作業者の主観に基づくものである。従って、作業者が認識していない潜在的な疲労がタスクの品質に影響を与えるなどの可能性を排除しきれない。精神的な疲労を測定する既存研究として、脳波を測定する手法[5]やウェアラブルデバイスを用いて運転手の疲労を推定する手法[27]などが提案されている。しかし、前者は器具を常時装着することが困難であり、日常生活を対象とするモバイルクラウドソーシングには適さない。後者は運転という特定用途における疲労の測定に留まっている。作業者が認識している状況に限定されるものの、アンケートを利用する本稿の方法は、現状における現実的な方法であったと我々は考えている。

実験結果では、多忙度と同行者の有無に高い関連性が見られた。アンケートの設計において、「同行者がいて多忙」「同行者がいなくて多忙」「同行者がいて多忙ではない」「同行者がいなくて多忙ではない」といった詳細な条件を設定することで、より精緻な分析が可能になることが見込まれる。

モバイルクラウドソーシングでは、現地調査のような、物理的な移動を伴うタスクも存在する。これらは、作業者のより詳細な状況(屋内にいるかどうかなど)に依存するため、本稿では問題を簡潔にする観点から、その場で実施可能なタスクを用いた。移動を伴うタスクもモバイルクラウドソーシングの主要なタスクの1つであるため、これらに影響を与える作業者の状況についての評価は今後の課題と考えている。

また、本稿では、タスクが1種類しかない場合について、評価を実施した。実際のモバイルクラウドソーシングでは、複数のタスクが同時に募集されている可能性があり、そのような場合に、作業者の状況がタスク選択にどのような影響を与えるかの評価は、今後の課題と考える。

7. まとめ

本稿では、多忙度、疲労度、同行者の有無といった作業者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。作業者の状況については、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、作業者が多忙な場合、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。被験者に対するアンケートにおいても、多忙時や同行者がいる場合は、タスクを受けにくいという意見が得られた。

実店舗での調査や特定の場所に関する調査といったモバイルクラウドソーシングは今後さらに普及することが予想され、なるべく多くのタスクを高品質に実施するニーズは大きいといえる。本稿の調査によって、休日/平日や時間帯によって作業者の状況に傾向があることも確認できた。これらの傾向の活用や作業者の行動ログを解析することによって、より多くのタスクを高品質に実施させることが可能となると考えられる。本稿では一例として日時や時間によって状況が異なることを示した。将来的には、センサ技術や端末ログからの状況取得の研究が進めば、それらで取得した状況を利

		7	8	9	10	11	12	13	14	15	16	17	18	19	20
平日	多忙	0.378	0.426	0.488	0.516	0.534	0.566	0.581	0.535	0.515	0.518	0.548	0.628	0.612	0.618
	疲労	0.394	0.384	0.401	0.447	0.486	0.544	0.548	0.568	0.560	0.599	0.635	0.688	0.678	0.718
	同行者	0.550	0.595	0.635	0.656	0.674	0.716	0.721	0.688	0.674	0.637	0.668	0.692	0.623	0.509
休日	多忙	0.215	0.252	0.289	0.252	0.291	0.333	0.351	0.387	0.367	0.356	0.406	0.471	0.526	0.444
	疲労	0.363	0.339	0.341	0.358	0.409	0.458	0.477	0.544	0.541	0.586	0.558	0.563	0.603	0.651
	同行者	0.549	0.566	0.610	0.614	0.615	0.678	0.690	0.668	0.700	0.671	0.685	0.697	0.744	0.683

		7	8	9	10	11	12	13	14	15	16	17	18	19	20
平日	受託金額	28.7	30.2	30.9	31.2	32.2	31.6	30.3	27.4	30.8	30.1	30.5	31.7	32.5	32.9
休日	受託金額	29.3	30.7	28.6	30.6	31.8	29.7	32.6	30.1	29.8	29.8	31.7	32.1	34.5	28.8

図 1 平日および休日の時間帯ごとの作業者の状況(多忙, 疲労, 同行者ありの作業割合)(上側)とタスクの受託金額(下側)
Figure 1: The ratio of workers in high busyness, high fatigue and with someone situations (upper table) and mean accepted task price (lower table) grouped by weekdays, weekends and time of day

用して、モバイルクラウドソーシングの効率を高めることが期待される。

今後の課題として、移動を伴うようなモバイルタスクに本稿の知見を適用するにあたって、影響を与える要因の調査や、複数のタスクが受託できる場合に、作業者の状況がタスクの選択にどのような影響を及ぼすかの評価を予定している。

[文献]

[1] Alt, F., Shirazi, A., Schmidt, A., Kramer, U. and Nawaz, Z. Location-based crowdsourcing: extending crowdsourcing to the real world. In Proc. NordiCHI 2010, 13-22.
 [2] Vaataja, H., Vainio, T., Sirkkunen, E. and Salo, K. Crowdsourced news reporting: supporting news content creation with mobile phones. In Proc. MobileHCI, 2011, 435-444.
 [3] Gigwalk, <http://www.gigwalk.com/>
 [4] Roberts, R. Getting the most out of life: The concept of opportunity cost. The Library of Economics and Liberty, 5, 2007.
 [5] Boksem, M. A., Meijman, T. F. and Lorist, M. M. Mental fatigue, motivation and action monitoring. Biological psychology, 72, 2, 2006, 123-132.
 [6] 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, 48, 2, 2016, 275-298.
 [7] Antin, J. and Shaw, A. Social desirability bias in reports of motivation for US and India workers on Mechanical Turk. In Proc. CSCW, 2011.
 [8] Mason, W. and Watts, D. J. Financial incentives and the "performance of crowds". In Proc. HCOMP, 2009, 77-85.
 [9] Beenen, G., Ling, K., Wang, X., Chang, K., Frankowski, D., Resnick, P. and Kraut, R. Using social psychology to motivate contributions to online communities. In Proc. CSCW, 2004, 212-221.
 [10] Law, E., Yin, M., Joslin Goh, K. C., Terry, M. and Gajos, K. Z. Curiosity killed the cat, but makes crowdwork better. In Proc. CHI, 2016.
 [11] Shaw, A. D., Horton, J. J. and Chen, D. L. Designing incentives for inexpert human raters. In Proc. CSCW 2011, 275-284.

[12] Yu, L., André, P., Kittur, A. and Kraut, R. A. Comparison of social, learning, and financial strategies on crowd engagement and output quality. In Proc. CSCW 2014, 967-978.
 [13] Le, J., Edmonds, A., Hester, V. and Biewald, L. Ensuring quality in crowdsourced search relevance evaluation: The effects of training question distribution. In Proc. CSE, 2010, 21-26.
 [14] Liu, Y., Lehdonvirta, V., Alexandrova, T. and Nakajima, T. Drawing on Mobile Crowds via Social Media. Case UbiAsk: Image Based Mobile Social Search Across Languages. Multimedia Systems, 18, 1, 2012, 53-56.
 [15] Gupta, A., Thies, W., Cutrell, E. and Balakrishnan, R. mClerk: enabling mobile crowdsourcing in developing regions. In Proc. CHI, 2012, 1843-1852.
 [16] 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, 96-104.
 [17] Czerwinski, M., Cutrell, E. and Horvitz, E. Instant messaging: Effects of relevance and timing. In Proc. HCI, 2000, 71-76.
 [18] Iqbal, S. T. and Horvitz, E. Notifications and awareness: A field study of alert usage and preferences. In Proc. CSCW, 2010, 27-30.
 [19] Musthag, M. and Ganesan, D. Labor dynamics in a mobile micro-task market. In Proc. CHI, 2013, 641-650.
 [20] Goldman, A. and Johansson, J. K. Determinants of search for lower prices: An empirical assessment of the economics of information theory. Journal of Consumer Research, 5, 3, 1978, 176-186.
 [21] Johnson, M. B. Travel time and the price of leisure. Economic Inquiry, 4, 2. 1966, 135-145.
 [22] Campagne, A., Pebayle, T. and Muzet, A. Correlation between driving errors and vigilance level: influence of the driver's age. Physiology and Behavior, 80, 2004, 515-524.
 [23] Hembrooke, H., and Geri G. The laptop and the lecture: The effects of multitasking in learning environments. Journal of computing in higher education, 15, 1, 2003, 46-64.
 [24] 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*, 52, 1, 2014, 106-113.
- [25] Kaup, F., Jomrich, F. and Hausheer, D. Demonstration of NetworkCoverage—A Mobile Network Performance Measurement App. In *Proc NetSys*, 2015.
- [26] Hoßfeld, T., Seufert, M., Hirth, M., Zinner, T., Tran-Gia, P. and Schatz, R. Quantification of YouTube QoE via crowdsourcing. In *Multimedia*, In *Proc. ISM*, 2011, 494-499.
- [27] Chen, L. B., Chang, W. J., Su, J. P., Ciou, J. Y., Ciou, Y. J., Kuo, C. C. and Li, K. S. M. A wearable-glasses-based drowsiness-fatigue-detection system for improving road safety. In *Proc. of IEEE Global Conference on Consumer Electronics*, 1-2, 2016.

池田 和史 Kazushi IKEDA

2008年大阪大学大学院博士前期課程修了。同年 KDDI (株) 入社, (株) KDDI 総合研究所所属。自然言語処理, 機械学習, ソーシャルメディア解析, クラウドソーシング等の研究に従事。2014年~2015年スタンフォード大学コンピュータサイエンス学科客員研究員。日本データベース学会, 電子情報通信学会, 情報処理学会, 各会員。博士 (情報科学)。

帆足 啓一郎 Keiichiro HOASHI

1995年, 早稲田大学理工学部情報学科卒業。1997年, 同大学院修士課程修了。同年, 国際電信電話 (株) (現 (株) KDDI 総合研究所) 入社。以来, マルチメディア情報検索等の研究に従事。現在, 同社知能メディア G・グループリーダー。2001年~2005年, 早稲田大学メディアネットワークセンター非常勤講師。FIT2004 ヤングリサーチャー賞受賞。経産省「始動 Next Innovator 2015」第1期選抜メンバー。日本データベース学会, 電子情報通信学会, 情報処理学会, ACM, 各会員。工学博士。