

会話型インタフェースを用いたオンラインアンケートの回答態度改善手法の検討

A Study of Methods for Improving Response Attitudes in Online Surveys

Using Conversational Interfaces

平良 繁幸[†] 松田 裕貴[†] 福光 嘉伸[†] 諏訪 博彦[†] 安本 慶一[†]
Shigeyuki Taira Yuki Matsuda Yoshinobu Fukumitsu Hirohiko Suwa Keiichi Yasumoto

1. はじめに

オンラインアンケートは、従来の紙媒体でのアンケートと比較して、クラウドソーシングサービスとの相性が良く、紙媒体のアンケートより低コストで大規模なデータを取得することができる。そのため、オンラインアンケートは社会科学分野の調査研究や民間企業のマーケティング活動等に広く用いられている。しかし、アンケート調査においてユーザーが必ずしも適切な回答を行うとは限らない。例えば、回答の対価として報酬を付与する場合、ユーザーは回答の正確さよりも回答時間の短縮を優先し、可能な限り迅速に報酬を獲得しようとするように行動することがある。他にも、質問の内容が正しく伝わっていなかった場合や、ユーザーの集中力が低下している場合には、ユーザーの悪意の有無に関わらず、結果的に不適切な回答が発生することが考えられる。ここでいう不適切な回答とは、ユーザーの本心とは異なる回答のことを示す。

このような問題に対して Simon ら [1] は、人間の認知的資源には限りがあり、それによって生じる、アンケート調査において回答者が回答要求に対する努力を最小化しようとする傾向を努力の最小化 (Satisficing) と定義した。この Satisficing について、Maniaci ら [2] は、オンラインアンケート調査における Satisficing を、後述する様々な手法を用いて調査し、Satisficing によって生じる不適切な回答はデータの質に悪影響を及ぼすことを示している。また、三浦ら [3] の調査では、1800 人を対象としたオンラインアンケートを 2 つの調査会社で実施した結果、それぞれ全体の 51.2 %、83.8 % が不適切な回答をしていたと報告されており、Satisficing によって生じる不適切な回答は、調査結果の解釈を困難なものとするという点で、極めて望ましくない。よって、オンラインアンケート調査において、Satisficing を検出し適切な処理を施すことは、より真実に近い知見を得るために極めて重要であると考えられる。

本稿では、スマートフォン上でのオンラインアンケートを対象として Satisficing をリアルタイムに検知する方法と、会話型インタフェースを用いた行動変容のため

の介入方法を提案する。提案手法では、後上ら [4, 5] の開発したプラグインを拡張したものをを用いて、回答中のユーザーの画面操作をリアルタイムに記録し、得られる特徴量を用いた機械学習による Satisficing 検出を行う。さらに、アンケートを対話形式で実施し、得られた回答にテキストマイニングを適用することでユーザーの性格特性を推定する。最終的には、Satisficing と分類されたユーザーに対して、推定した性格特性に適した介入を実施し、ユーザーの行動変容を実現する。

2. 関連研究

Satisficing の検出を目的とした手法として、Oppenheimer ら [6] の IMC (Instructional Manipulation Check) や、Maniaci ら [2] の ARS (Attentive Responding Scale)・DQS (Directed Question Scale) がある。これらの手法は検出用の質問をオリジナルの質問票に追加して回答の正確性や一貫性から Satisficing を検出するものであり、Satisficing 関連の研究で一般的に使用されている。しかし、追加する必要がある質問は回答者を試すような内容であるため、挿入することで疑われているように感じるなど、回答者の心理的負荷が増加する可能性があり、適切に回答している回答者のモチベーションを低下させ、かえって Satisficing を発生させる要因となり得る。また、単純に質問数が増加するという点でも、追加の質問を挿入することは望ましくない。そこで、後上ら [4, 5] は、検出用の質問を用いずに Satisficing を検出することを目的として、機械学習による検出手法を提案している。この手法では、専用のプラグインを用いて取得したユーザーのアンケート回答時の画面操作ログ (スクロール速度、選択肢の変更回数、テキストの変更回数等) を特徴量として既存の ARS, DQS といった検出方法との比較を行い、各特徴量の有用性を示している。しかし、この手法ではアンケートの回答が全て完了した後に検出を行なっているため、検出した Satisficing に対する処理としては、データから省くといった対応に限られる。このことから、これまでに提案された手法では、同一人物が一度しか実施することができないタスク等、データの母集団に限られている場合、Satisficing の影響によって最終的に得られるデータが不足してしまう恐れがある。また、検出された Satisficing 群を取り除いたユーザー群が、母集団の

[†] 奈良先端科学技術大学院大学, Nara Institute of Science and Technology

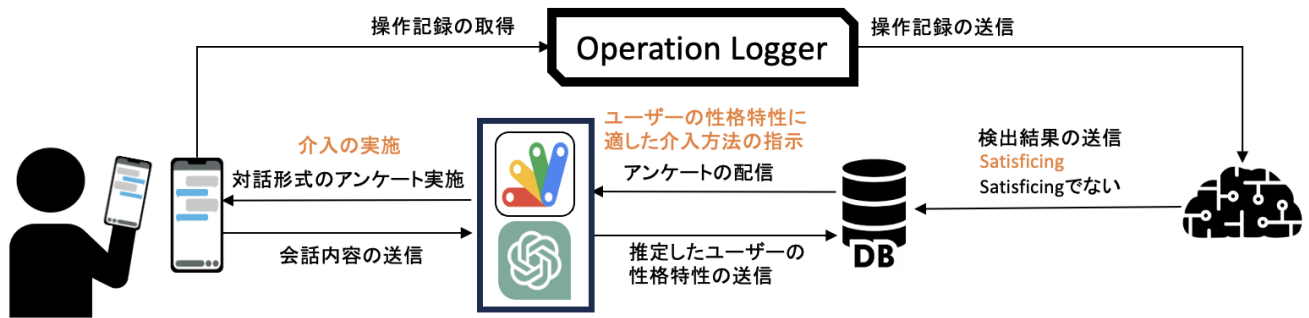


図 1: システム構成

特徴を有しているかという点で疑問が残る。

タスクに対するユーザのエンゲージメントや作業の質を向上させるための研究や、行動変容を促すことを目的とした介入方法に関する研究も多数行われている。Sihang ら [7] は、クラウドソーシング上でのマイクロタスクを対象とし、従来の Web インタフェースの代わりに会話型インタフェースを用いることで、ユーザーのエンゲージメントを向上させる手法を提案している。Jeffrey ら [8] は、長時間の作業中に適度な休憩を与えることで、ユーザーのエンゲージメントを大幅に改善できることを明らかにした。Peng ら [9] は、クラウドソーシングのマイクロタスクにおいて、少量の娯楽をタスクの間に提供することで作業の品質を維持しながら、ユーザーのエンゲージメントを大幅に改善できることを明らかにした。また、大山ら [10] は、参加型センシングを対象として、ボタンのタップと携帯電話を振る行為の 2 種類によって貢献の意思表示を行うことで速く回答を行う不適当な行為を抑制する方法を提案した。Zhang ら [11] は、ユーザに対する情報提示によりユーザの歩数の増加を促す行動変容において、提示する情報の対話スタイル（情報の粒度や婉曲表現の度合い）によって行動変容の効果が変化することを明らかにしている。これらの既存研究では、エンゲージメントを向上させる点に留まっており、一度 Satisficing と推定されたユーザーを改善させるような介入はされていない。よって、リアルタイムに Satisficing の検出を行い、介入によって行動変容を促すことで、適切な回答を得る重要性は高いと考えられる。

3. 提案手法

本研究の目的は、スマートフォン上でオンラインアンケートを実施する際に、回答者の Satisficing 傾向をリアルタイムで検出するとともに、ユーザーの性格特性ごとに適した介入方法によって行動変容を促すことにより、回答態度を改善させることである。また、本研究では、ユーザーの性格タイプ推定にテキストマイニングを用いる。テキストマイニングとは、入力されたテキストデー

タから分析に有益である情報を抽出し、感情分析やクラスタリングを行う手法であり、本研究では、ユーザーが対話形式のアンケートに回答する過程で取得できる会話の内容から、ユーザーの性格特性を推定するために用いる。

3.1 提案手法の概要

提案手法の概要およびワークフローについて、以下に詳細を述べる。

1. オンラインアンケートを作成し、会話型インタフェースを介して対話形式でユーザーからの回答取得を行う。
2. この際、アンケート回答システム上では、ユーザーの回答時の画面操作ログを定期的に収集する。また、会話型インタフェースとユーザーの会話内容に対しテキストマイニングを行い、ユーザーの性格特性を推定する。
3. 得られた画面操作ログを入力とし、機械学習モデルによって Satisficing を検出する。
4. Satisficing と分類されたユーザーに対し、推定された性格特性を考慮した介入を実施し、行動変容を促す。

3.2 システム設計

提案システムの構成を図 1 に示す。提案システムは、対話形式でのアンケート調査の実施や、ユーザーの性格特性に合わせた介入を行う会話型インタフェース、アンケート回答時の画面操作ログを取得する機能、ユーザーとの会話からテキストマイニングによってユーザーの性格特性を推定する機能、リアルタイムで Satisficing を検出する機能で構成される。

3.2.1 会話型インタフェース

対話形式でのアンケート調査を実施する会話型インタフェースの詳細について述べる。本研究では、会話型

表 1: 抽出する特徴量

特徴量	単位	独自追加
回答時間	s	-
前問までとの回答時間の変化	%	○
非操作時間	s	○
非操作時間が長すぎる回数	回	-
テキストの削除回数	回	-
連続同一回答数	問	-
文字数	文字	-
タップ回数	回	○
タップ間隔	s	○
前問までとのタップ回数の変化	%	○
前問までとのタップ間隔の変化	%	○

-: 後上ら [4, 5] による特徴量,
○: 新規に追加する特徴量

インタフェースとしてチャット UI とチャット AI である ChatGPT を用いる。具体的には、取得したオンラインアンケートの設問を、ChatGPT を介して対話形式に変換し、チャットアプリ内のボットからの問いかけとして出力する。ユーザーからの返答に対しても同様に ChatGPT を介した出力を行うことで、対話形式でのアンケート調査を実装する。

3.2.2 画面操作記録機能

画面操作を記録する機能について述べる。画面操作ログ取得には、後上ら [4, 5] が作成した Operation Logger を、チャットアプリに対応するよう拡張し用いる。

次に、画面操作データから抽出する特徴量について述べる。本研究では、後上ら [4, 5] が用いている特徴量を拡張し、リアルタイムに抽出可能な特徴量を追加する。表 1 に用いる特徴量とその単位を示す。特徴量のうち後上らの提案したものは“-”，新たに考案し追加するものは“○”を独自追加の列に示している。作業内容を見ず意図的に速く回答を行う場合では、同じ回答を内容を見ずに繰り返したり速く回答することがあるため、現状の連続同一回答数や問題間の回答時間などの特徴量が有用であると考えられる。注意散漫や疲労により集中が切れた場合では、前問から作業速度が下がることや作業が止まることがあるため、問題間の回答時間の変化や非操作時間などの特徴量が有用であると考えられる。

3.2.3 テキストマイニングによる性格特性推定機能

テキストマイニングを用いてユーザーの性格特性を推定する機能について述べる。本研究では、テキストマイニングによる推定を各設問の回答時に行うことで、ユーザーの性格特性を定常的に推定する。また、性格特性を推定する指標として Big5 尺度を採用する。Big5 尺度は、

人間のパーソナリティの性格を分析する理論の一つで、心理学の分野において広く扱われている。Big5 尺度の特徴として、人の性格を外向性、誠実性、協調性、開放性、神経症的傾向の 5 項目で評価し、典型的ではなく个性的に分析を行うものである。

3.2.4 不適切回答検出機能

アンケート回答時の不適切回答の検出を行う機能について述べる。本研究では、機械学習による分類を各設問の回答時に行うことで、Satisficing をリアルタイムで検出する。機械学習のアルゴリズムは、実験を行い取得した特徴量の傾向を基に最適なものを選択する。また、Satisficing になる前後の特徴量を比較することで、分類に有用な特徴量を選択する。

3.2.5 行動変容のための介入機能

Satisficing が検出された際に、回答中のユーザーに行動変容を促す機能について述べる。本研究では、行動変容の方法として会話型インタフェースによる介入を行う。ユーザーの性格特性をテキストマイニングによって推定し、性格特性ごとの適切な介入方法を ChatGPT にプロンプトとして入力することで、適切な回答をするよう行動変容を促すことが可能であると考えられる。

4. 実験計画

本章では、実際に行う実験の概要や計画について述べる。本稿の手法は、リアルタイムの介入を行うことでユーザーの回答中の行動変容を目的として提案される。以下に、目的を達成するための実験概要を示す。

- (1) テキストマイニングを用いて、Big5 に基づいた性格特性の推定が可能なモデルを開発するために、オンラインアンケートを作成し、学習データを取得する実験を行う。
- (2) 提案手法の Satisficing の検出・分類を行う分類モデルを開発するために、オンラインアンケートを作成し、学習データを取得する実験を行う。
- (3) オンラインアンケートをチャットアプリ上でユーザーに与え、画面操作データをシステムによってリアルタイムに取得し、特徴量を抽出する。会話型インタフェースとの会話記録を取得し、テキストマイニングによる性格特性の推定を実施する。
- (4) Satisficing である時と Satisficing でない時の特徴量の傾向の違いを分析することで、学習に用いる特徴量とデータの特徴性に適切なアルゴリズムを選択し、分類モデルの作成を行う。分類モデルの精度評価には、Accuracy, Precision, Recall, F1 Score を用い、

汎化性能の検証は、10-fold cross validation によって行う。不適切回答の検出方法としては、1つの設問が終わるごとに分類を行うことで逐次的な検知を行う。分類にかかる実行時間として、次の設問への回答が終わるまでに検出が出来ることを要件とする。

- (5) 分類モデル作成後に、介入によってどの程度の確率で Satisficing が改善されるかの実験を行い、行動変容を促す介入の手法について、効果を測定する。

5. まとめ

本研究はスマートフォンを用いたオンラインアンケートを対象とし、ユーザーの Satisficing 傾向をリアルタイムに検知する手法と、ユーザーの性格特性に適した介入によって行動変容を促す方法を組み合わせることで、回答の品質を向上するためのシステムを提案した。提案手法ではアンケート回答中の画面操作をリアルタイムに記録し、特徴量を抽出することでアンケート回答中に Satisficing を検出できるよう検討を行った。また、ユーザーの性格特性についてテキストマイニングを用いて推定し、性格特性を考慮した介入方法によって行動変容を促すよう検討を行った。

今後の予定としては、各システムおよびアンケートの開発を行い実験を実施し、提案手法の有効性を検証することが挙げられる。

謝辞

本研究の一部は、JST さきがけ (JPMJPR2039) の助成を受けて行われたものです。

参考文献

- [1] Herbert A Simon. Rational choice and the structure of the environment. *Psychological review*, Vol. 63, No. 2, p. 129, 1956.
- [2] Michael R Maniaci and Ronald D Rogge. Caring about carelessness: Participant inattention and its effects on research. *Journal of Research in Personality*, Vol. 48, pp. 61–83, 2014.
- [3] 三浦麻子, 小林哲郎. オンライン調査モニタの satisficing に関する実験的研究. *社会心理学研究*, Vol. 31, No. 1, pp. 1–12, 2015.
- [4] 後上正樹, 松田裕貴, 荒川豊, 安本慶一. オンラインアンケート回答時のスマートフォン画面操作状況に基づく不適切回答検出. 第 25 回一般社団法人情報処理学会シンポジウム・インタラクシオン 2021, pp. 11–20, 2021.
- [5] Masaki Gogami, Yuki Matsuda, Yutaka Arakawa, and Keiichi Yasumoto. Detection of Careless Responses in Online Surveys Using Answering Behavior on Smartphone. *IEEE Access*, Vol. 9, pp. 53205–53218, 2021.
- [6] Daniel M Oppenheimer, Tom Meyvis, and Nicolas Davidenko. Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of experimental social psychology*, Vol. 45, No. 4, pp. 867–872, 2009.
- [7] Sihang Qiu, Ujwal Gadiraju, and Alessandro Bozzone. Improving Worker Engagement Through Conversational Microtask Crowdsourcing. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI'20, pp. 1–12, 2020.
- [8] Jeffrey M Rzeszotarski, Ed Chi, Praveen Paritosh, and Peng Dai. Inserting micro-breaks into crowdsourcing workflows. In *The First AAAI Conference on Human Computation and Crowdsourcing*, HCOMP'13, pp. 62–63, 2013.
- [9] Peng Dai, Jeffrey M Rzeszotarski, Praveen Paritosh, and Ed H Chi. And Now for Something Completely Different: Improving Crowdsourcing Workflows with Micro-Diversions. In *Proceeding of The 18th ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW'15, pp. 628–638, 2015.
- [10] Kohei Oyama, Yuki Matsuda, Rio Yoshikawa, Yugo Nakamura, Hirohiko Suwa, and Keiichi Yasumoto. A Method for Expressing Intention for Suppressing Careless Responses in Participatory Sensing. In *18th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, MobiQuitous'21.
- [11] Zhihua Zhang, Juliana Miehle, Yuki Matsuda, Manato Fujimoto, Yutaka Arakawa, Keiichi Yasumoto, and Wolfgang Minker. Exploring the Impacts of Elaborateness and Indirectness in a Behavior Change Support System. *IEEE Access*, Vol. 9, pp. 74778–74788, 2021.