

# マウスログ解析によるWEBアンケートの高情報化

47086681 飯塚 勝哉  
指導教員 岡本 孝司 教授

In Web survey, a simple form is needed because a complicated form imposes a burden on respondents and reduces responses and their reliability. Therefore, we need to develop the method which extracts more information from surveys in a simple form. The track of mouse was used for detecting an indecision behind a final reply to a query. We focused on the survey for ranking and evaluated differences between ranks through analysis of the track of mouse.

Key words: Ajax, questionnaire, Mouse, Eye track, Information visualization, Web survey

## 1 緒言

人々の意識やニーズが多様化している現在、意識調査の一手段としてアンケートが世界各地で盛んに行われている。特に、近年のWEB技術の発達やインターネットの普及を背景に、WEBアンケートが注目されている。WEBアンケートは、安価・迅速かつ世界同時調査の実現可能性があるほか、秘匿性が高いため、他の手法では尋ねにくい内容についても詳細な調査が可能となる。しかし、詳細な意識調査を行うため、アンケートフォームを複雑にすると、回答者の負担が増加し、回収率の低下や回答精度の悪化等、集計結果の信頼性を低下させてしまう。そこで本研究では、Ajaxを用いてアンケート回答時のマウス挙動のデータを取得・解析し、回答時の迷い等の心理状態を加味した付加情報を取得することで、簡易なアンケートフォームで詳細な意識調査を行う手法の開発を目指す。具体的には、順位付け型(PA)アンケートに着目し、回答結果から得られる各選択肢の順位だけでなく、本来ならば回答者への負担が大きいが対比較法等でしか得られない順位間の差までも推定することを目的とする。

ここで、本研究のようにマウス挙動のデータを解析し、付加情報を得ようという試みはすでに様々行われている。王子<sup>1)</sup>は複数選択型(MA)のWEBアンケート回答時のマウス挙動から、詳細な意識調査が行えることを示している。しかし、マウスの挙動はタスクによって異なると考えられる。そこで、まず、①PA質問回答時のマウス挙動と心理的要素の関連性を評価し、②その特徴を踏まえて調査分析手法の開発・検証を行う。認知プロセスの解析には視線の動きがしばしば利用されるため<sup>2)</sup>、視線とマウスの挙動との関係性を評価することで、①を行う。

アンケート回答時において、回答者は各選択肢に対して1から0の範囲のある数値(潜在得点)を持ち、それに応じて各選択肢の順位を決定する。そのため、本研究では、各選択肢の潜在得点の差を順位間の差とし、推定する。

## 2 予備実験

### 2.1 実験概要

PA質問回答時のマウスと視線の動きの関連性を評価するため、予備実験を行った。女子大学生20名を対象とし、PA質問とそれに対応する同項目のMA質問を1組として、食に関する質問16組からなるWEBアンケートを実施した。アンケートに回答してもらった際、Ajaxを利用して回答者のマウスログを取得し、また、視線計測機を用いて回答者(6人)の視線の動きも取得した。さらに、WEBアンケート回答終了後、ヒアリング調査を行い、同じ質問の各選択肢に対する潜在得点を評価してもらった。つまり、1

つの質問に対し、PA・MA・ヒアリングの3つの形式で調査した。使用器具は、視線計測機(TOBII製 Tobii X50)、パソコン(ブラウザ…IE6.0、画面解像度…1024×768、画面表示…最大表示)、マウスを用いた。マウスログ取得の周波数は40Hz、視線データ取得の周波数は50Hzである。また、各質問に回答する際のマウスの初期位置は、画面下部である。回答者1人あたりの所要時間は、事前準備に10分、WEBアンケートに10分(PA質問平均回答時間…26秒、MA質問平均回答時間…12秒、質問1組あたり平均所要時間約40秒)、ヒアリング調査に10分の計30分であった。また、今回はPA質問の回答結果および回答時のデータ(回答時間・マウスログ・視線の動き)、ヒアリングの回答結果のみに着目する。それらのデータのうち、PA質問の回答結果(順位)とヒアリングの回答結果の順位が一致し、整合性が取れているデータ(全データの51.3%)のみを用いて分析を行った。ヒアリング結果は各選択肢の潜在得点の正解値として扱う。

### 2.2 視線との比較によるマウス挙動の評価

PA質問回答時のマウスの動きと視線の動きの例を図1に示す。赤い点がマウスの挙動、青い点が視線の動きを表す。図1内の画面左上を原点とし、画面横方向をX軸、縦方向をY軸と設定する。図1より、マウスは主にチェックボックス上をY軸方向に動く。問題文の文章上をX軸方向に動く場合を除き、視線は主に選択肢の文章上をY軸方向に動く。よって、マウスと視線のY軸方向の変位に着目し、関係性を評価する。

図2にマウスと視線のY座標の時系列グラフを示した。赤い線がマウスのY座標値、青い線が視線のY座標値の時系列変化を表す。図2の各色の点は、各選択肢の位置(Y座標)と各選択肢が選択された時刻を示している。図2より、マウスの挙動は「急激な動き」と「停滞」の2種類に分かれるが、マウスが「急激な動き」をする場合、視線も同様、もしくは少しタイミングがずれて動いていることがわかる。一方、マウスが「停滞」している場合も、基本的にはマウスの「停滞」箇所と視線の位置は一致するが、図2の①のように、問題文を読んでいる際は、視線は画面上部の問題文のあたりに滞在するのに対し、マウスは別の位置に滞在することが多い。また、図2の②のように、マウスを「停滞」させたまま、視線は他の場所を歩き来しているケースもある。特に、マウスがチェック直前にある選択肢のチェックボックス上に「停滞」している際は、視線は他の選択肢間を歩き来しやすいという傾向が見られた。このように、マウスが「停滞」している場合は視線の位置と一致しないケースもあるが、基本的にはチェックを開始してからはマウスと視線の動きは類似していることが確認できた。この類似性を評価するため、双方のグラフ

の積分値からのずれの平均値を算出した。積分値を用いることで、マウスと視線の動きの少しのタイミング差も評価できると考え、以下の式から算出した。以下の式(1)で、 $\bar{d}$  はマウスと視線の y 座標の時間当たりのずれ平均値、 $y_m$  はマウスの y 座標、 $y_e$  は視線の y 座標、T はチェック開始から回答終了までの時間を表す。

$$\bar{d} = \frac{\int y_m dt - \int y_e dt}{T} \quad [\text{px}] \quad (1)$$

上記の式から算出した y 座標のずれ平均値  $\bar{d}$  の全体の平均は 25[px]であった。これは、各選択肢間の距離 50[px]の半分程度であり、小さいと言える。よって、マウスと視線の動きは類似しており、マウス挙動の分析から心理状態を加味した付加情報を取得できうる。

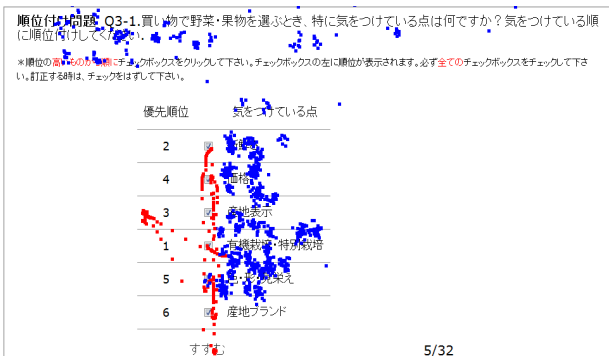


Fig. 1 Track of mouse and eye

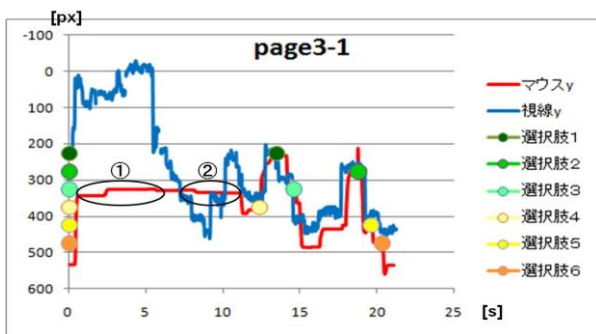


Fig. 2 Time series graph of mouse and eye Y-coordinate

### 3 実験

予備実験を踏まえて、マウス挙動データを用いて、順位間の差を推定するモデルを構築・検証するため、中部電力株式会社(以下中電と呼ぶ)における社内アンケート実験や東京大学大学院新領域創成科学研究科における学内アンケート実験を行った。今回は、中電における社内アンケート実験について示す。

#### 3.1 実験概要

中電の従業員 504 名を対象とした社内アンケートに本研究のシステムを適用し、Ajax を利用してアンケート回答時の回答者のマウスログを取得した(取得周波数は 40Hz)。1 人あたりの問題数は 24 問であり、そのうち PA 質問は 4 問である。また、各選択肢の潜在得点の評価は、PA 質問直後に同問題のスライドバー型質問を用意し、ス

ライドバーを動かして 0~100 点の範囲の連続値で回答してもらい、調査した。以後、スライドバーの値を各選択肢の潜在得点の正解値として扱う。アンケートページは予備実験と同じデザインを用い、選択肢の表示順が回答に与える影響をなくすため、表示順を回答ごとに変えた。各質問回答時のマウスの初期位置は、画面最下部である。

回答者 1 人あたりの平均回答時間は 48[s]であった。今回分析対象としたデータは、PA+スライドバー型の質問に対する 1008 名分のデータの中から以下のデータを除いた回答データ 478 名分(48%)である(1 問あたり 80 名分のデータである)。

- ・スライドバー型質問の回答がないデータ(233 名分, 23%)
- ・PA 質問の回答結果(順位)とスライドバーの回答結果の順位が一致しないデータ(262 名分, 26%)
- ・PA 質問に完答していないデータ(23 名分, 2.3%)
- ・回答時間が平均の 4 倍以上あるデータ(12 名分, 1.2%)

#### 3.2 従来手法の問題点

PA 質問において、従来は、集計値を出す際は順位間の差を一定とみなし、1 位を 100 点、最下位を 0 点として順位に応じ等間隔に得点配分した値(離散値)を各順位の選択肢の得点(順位得点)としていた。しかし、実際は、各順位間の差は一定ではなく、各選択肢の持つ潜在得点も連続値であり、そこに調査の質を低下させる恐れがある。ここで、スライドバーの値から各順位間の差を質問別に集計して算出し、図 3 に示す。図 3 より、各順位間の差は一定ではなく、そのばらつきも質問ごとに異なる。

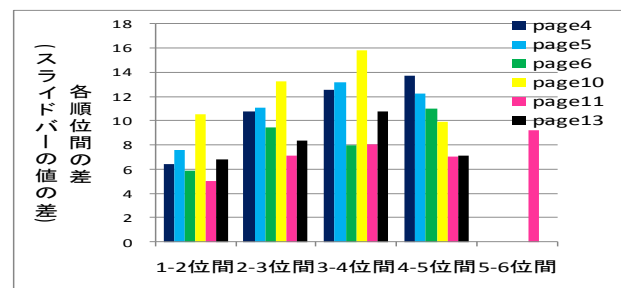


Fig. 3 differences between ranks

アンケートでは選択肢別の集計値が利用されることが多いため、選択肢別の集計値に対する順位間の差の影響を考える。図 4 に、質問ごとに従来手法で集計した各選択肢の順位得点の平均値と正解値との関係を示す。正解値の方が全体的に高得点になっているが、これは、今回の社内アンケート実験の特徴であり、スライドバー型質問で 0~100 点すべてを使って回答せず、全体的に高い点(いい評価)をつけているためであると考えられる。集計値に対する順位間の差の影響のみを考えるため、正解値の最低点を 0 点、最高点を 100 点に換算した時の正解値と順位得点の関係を図 5 に示す。図 5 より、順位得点は正解値とほぼ一致し、順位間の差を一定とした離散化による調査の質の低下は小さい。しかし、これは 478 名分の順位得点と正解値を比較した結果である。アンケート調査において

は、サンプル数が少ない場合もあり、例えば、本実験の場合、管理職のみの問題で回答者数が20~30名以下の設問もあった。このように、少ないデータ数で詳細な情報を推定しなければならない。そのような場合でも精度の高い調査を行えるよう、順位付け回答のデータ数を少なくした際、同様の結果が出るか調べる。そこで、母集団に対するデータ数(%)と、質問ごとの集計値と正解値との平均点数差との関係を図6に示す。図6より、データ数が40%をきるとずれが平均して5点以上になり、問題によっては7.8点ずれるものもある。これらの誤差は、突出した選択肢がなく、どの選択肢も同様に選択される”page5”や”page11”のような問題においては、各選択肢間の集計値の差があまりない分、集計結果に大きく影響してしまう。

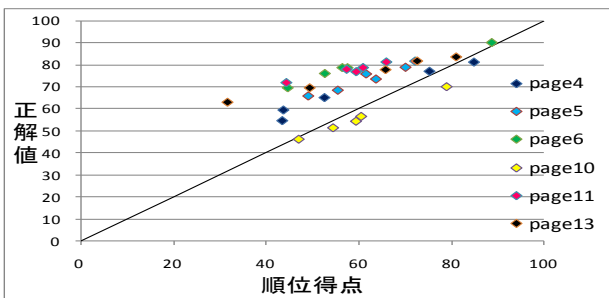


Fig. 4 Correlation between Bar and PA score (case1)

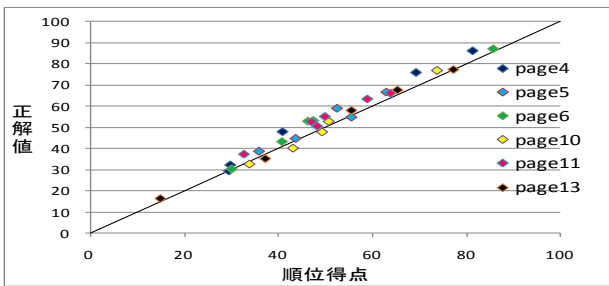


Fig. 5 Correlation between Bar and PA score (case2)

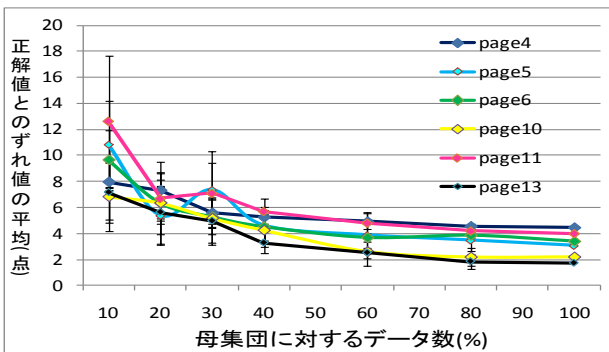


Fig. 6 Correlation between Error and the number of data

### 3.3 順位間の差推定モデルの構築

従来手法では上記のような問題が生じるため、本研究では、順位間の差を推定するモデルを構築し、そのモデルを使って、選択肢別集計値の精度向上を図る。

PA 質問のタスクの性質上、ある順位  $i$  位の選択肢を選

んだ後  $i+1$  位を選ぶまでの時間の長さ  $T_{i,i+1}$  (回答時間間隔) は、 $i+1$  位の選択肢を選ぶ際の「迷い」を反映していると考えられる。しかし、それが  $i$  位の選択肢とどの選択肢間で迷ったかは特定しづらい。そこで本研究では、回答時間間隔  $T_{i,i+1}$  に着目し、ある特定の順位間の差を求めるのではなく、各選択肢の潜在得点をそれぞれ独自に算出した。その後、最低点を0点、最高点を100点に換算することで、それぞれの結果を全体に反映し、各順位間の差を反映した潜在得点算出モデルを構築した。各選択肢の潜在得点は順位得点と同じ点数を基準に、その選択肢を選ぶ際の「迷いの強さ」に応じて、図8のように算出される。選ぶのに「迷った」分、その選択肢の潜在得点は下がると考えられるので、本研究では潜在得点関数を1次直線で設定し(ある値以降は一定値)、また、各選択肢の潜在得点の最大値と最小値は、前後の順位の順位得点の中間値とした。

このとき、ある選択肢を選ぶ際の「迷い」を表すパラメータとして、回答時間間隔を基に以下の3つを用いる。

- 1、各選択肢を選ぶまでの回答時間間隔
- 2、条件付き時間間隔

予備実験より、チェック直前にマウスがチェックボックス上に滞在しているとき、視線は他の選択肢間を見ている。この特徴より、回答者は  $i$  位の選択肢をチェックする直前、他の選択肢と比較し、チェックしていいか再確認をしたり、または  $i$  位の選択肢チェック後他の選択肢をどうチェックしていくかを考えており、 $i$  位の選択肢を選ぶのに「迷った」他の時間に比べ、「迷い」の度合いが少ないと考えられる。そこで、この時間をカットした時間間隔を用いる。

- 3、移動量加味時間

予備実験より、マウスが「急激に動く」ときは視線も同様に動くことが分かった。つまり、マウスだけ適当に動かすような無駄な挙動は少なく、マウスが動く時は何かの意図があると考えられる。そのため、選択肢間を行き来するような挙動は、その選択肢間での「迷い」を反映しているが、マウスの動きの所要時間はとても短く、回答時間間隔にこの「迷い」が反映されていない。そこで、移動量の多さに応じて重みづけした回答時間間隔を用いる。

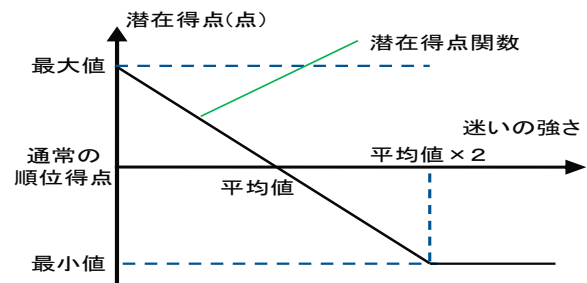


Fig. 7 Explanation of model

これら3つのパラメータには個人差がある。そこで、回答者ごとにある順位  $i$  位の選択肢を選ぶ際の各パラメータの平均値を算出して、平均値で各値を割り、個人差の影響をなくした。また、page11のみ選択肢数が6個であり、他の質問と回答の仕方が異なるため、分析から外した。

### 3.4 各パラメータおよびモデルの評価

各パラメータと提案モデルを用いて算出した集計値の精度を評価した。まず、各質問の2位, 3位, 4位(1位は100点, 5位は0点で固定されている)の潜在得点の平均値を順位ごとに算出する。その後、従来の順位得点、正解値についても同様に、順位ごとに得点の平均値(各順位が何点になるかを示す)を算出し、順位得点と正解値との誤差に比べ、モデル値と正解値との誤差がどの程度か評価した。評価方法は式(2)で算出した値で行い、式(2)の値が1より小さければ誤差が従来値より小さくなっている。

$$E = \frac{|\text{正解値} - \text{モデル値}|}{|\text{正解値} - \text{順位得点}|} \quad (2)$$

表1に、各パラメータと提案モデルを用いて算出した各順位のEの平均値(2位~4位分の平均)を質問ごとに示す。表1より、条件付き時間を用いた場合が一番誤差が小さくなっているため、今回は条件付き時間を用いて潜在得点の算出を行う。パラメータによる違いは今後さらなる検討が必要である。また、表1より、page6とpage10の質問における2位(もしくは3位)の潜在得点と正解値とのずれが大きいが、この原因については、モデルの特性上、各順位の潜在得点の初期値と可動範囲が決まっていることに起因する可能性があり、今後さらに調べる必要がある。今回は、page6とpage10以外の問題について着目し、条件付き時間間隔と提案モデルを用いて算出した集計値と正解値との誤差(E)と、母集団に対する利用したデータ数の割合の関係を図9に示す。図9より、母集団に対するデータ数が30~40%ぐらいまでなら、3つの質問において従来の順位得点より誤差が小さくなり、提案モデルで順位間の差を推定し、正解値に近い各順位の集計値を出せることがわかる。例として、表2に50%のデータ数から提案モデルにより算出した集計値の誤差Eを順位ごとに示す。page6とpage10以外の問題では、全ての順位において誤差が小さくなっている。その母集団の50%のデータ数から母集団の正解値を提案モデルで推定した結果の例を図10に示す。従来手法による推定結果と比べ、各選択肢別の集計値が正解値に近づいた。また、従来の手法では集計結果の順位が誤っていたところも、提案モデルにより修正された。その修正された点数は約3点であり、これは、ある順位の選択肢のとり得る値の範囲(20点)の3%の変化にあたり、一定の価値があると考えられる。

	Page4	Page5	Page6	Page10	Page13
回答時間間隔	0.66	1.084	1.327	1.252	0.516
条件付き時間	0.662	0.791	1.207	0.977	0.493
移動量加味時間	0.555	1.084	1.262	1.041	0.501

Table. 1 Evaluation of parameter

	Page4	Page5	Page6	Page10	Page13
2位	0.405	0.688	1.178	1.626	0.871
3位	0.910	0.995	1.543	0.854	0.577
4位	0.667	0.790	0.925	0.729	0.415

Table. 2 E of model score(fifty percent data)

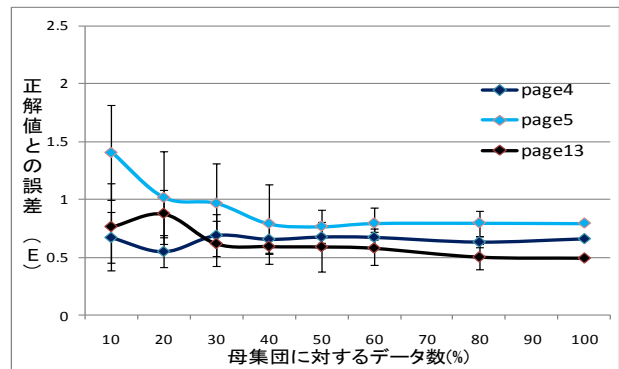


Fig. 8 the graph of model sensitivity

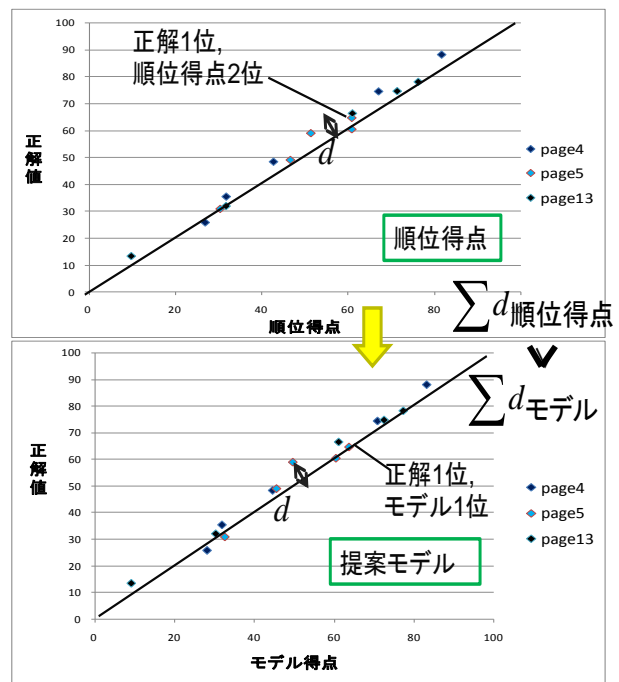


Fig. 9 The Effect of Model

## 4 結論

回答者の負担が少なく、かつ詳細な意識調査が可能なアンケート手法の開発を目的として、順位付け型質問に着目し、回答時のマウス挙動から順位間の差の情報を取得する手法を開発した。従来の離散化された順位得点による集計手法は、データ数の多い推定には影響はないが、データ数が減ると誤差が生じ、各選択肢の選択率が近いような問題では大きな影響を与える。本提案モデルにより、母集団の50%のデータ数までなら、順位間の差を推定し、集計値の誤差を減らすことができることを示した。今後は、潜在得点関数の最適化や他のパラメータサーベイ、モデルが有効ではない質問や回答の特徴を分析し、モデルを改良する必要がある。

### 文献

- 1) 王子富幸, “Web 環境下における意思決定に関する研究”, 東京大学大学院新領域創成科学研究科, 2008
- 2) R.Keith, “Psychological Bulletin”, 124, 372-42