

潜在性を考慮したコンシューマ分析の提案

星野寛登^{†1} 杉村博^{†2} 松本一教^{†1}

この論文では従来からのアンケート調査によるデータ収集に加えて、人間の潜在的な特徴データを収集できるシステムを提案する。潜在的なデータを利用することで、従来のアンケート調査からは知ることができなかった特徴を分析することが可能となる。本論文ではシステムの概要提案と小規模な実験結果を示す。

Towards Feature Discovery Reflecting Consumer's Implicit Attitudes

HIROTO HOSHINO^{†1} HIROSHI SUGIMURA^{†2} KAZUNORI MATSUMOTO^{†1}

This paper proposes a new combined method that can measure consumer attitudes. Conventional approaches such as interviews, protocol analysis, and data mining mainly focus on attitudes that explicitly appear on the conscious level so that implicit ones could not be measured by those conventional approaches. Recent studies on cognitive science states implicit human attitudes to some extent subject conscious behaviors. With this investigation, the implicit association test, which is often called IAT, is developed as an effective tool to measure implicit part of human understanding of things. We in this paper propose an extended version of IAT that runs on Web environment. Based on this tool, we propose a new data mining that reflects implicit information of human. This paper shows an essential idea of this study and results of small experiments.

1. はじめに

最近の日本国内向け家電は趣味性が高い製品が増えてきており嗜好品と呼ぶべきレベルになっている。音楽や映像の鑑賞のための製品であっても、性能の平均レベルが向上してきたために、趣味性の部分で商品価値が左右されてしまうケースが増えてきている。これからのコンシューマ向け製品あるいはサービスの開発においては、コンシューマの嗜好を正しく分析して企画を立てることが重要となってくる。もちろん従来からも商品企画の重要性は認識されてきており、専門的な研究活動も展開されてきている。企画立案の際に大きな役割を果たすものとして消費者行動調査がある[1,2]、これらの手法は多くのデータを集めようとすれば、コスト増加が問題となる。安いコストで多くのデータを集める方法としてはアンケートの利用があり（質問紙調査法）、メールや Web により電子的に行うこともできる。このようなアンケート調査を行う場合、回答選択肢の設計が重要となってくるため、調査目的に応じて例えば、リッカート尺度や SD 尺度などが用いられている。その他にも第 2 章で概観するように方法が開発されてきている。ところで、人間の意思決定、行動[3]は全てがはっきりと意識されて制御されているものだろうか。最近の認知心理学的な研究[4,5,6]によれば、意識されない潜在的な部分の影響が大きいことが指摘されている。すなわち、人間の意識の潜在的な部分が顕在的な部分に影響を与えているという理論である。このような解釈はかなり以前からあったが、

潜在的な部分を簡単に測定する手法が存在しないために、実用上利用することは不可能であった。ところが 1990 年代の終わりごろに IAT (Implicit Association Test) [7,8,9] という手法が開発され、簡単に実行できる潜在的な意識の測定法 [8] として認められるようになってきた。この手法を用いたマーケティング手法の提案もなされているが、いまだに十分であるとは言い難いレベルである。そこで本研究では、電子的なアンケートに加えて IAT の概念を導入した手法を用いることにより、潜在的な意識にも焦点を当てた調査ができるシステムを提案する。

2. コンシューマの調査分析手法

コンシューマ（消費者）の行動を理解するための手法としては、アンケートの他にインタビュー、プロトコル分析、観察法、データマイニング等が代表的である。これらの概要を述べその問題点を指摘する。

- アンケート法：安価に大量の調査を実行できる。他人からの評価を意識することで回答が影響を受けることがある。従来手法では潜在的な意識は調べることができない。
- インタビュー法：面接者が対象者から直接聞き取りを行って調査する。詳細な情報を集めることができるが面接者の能力に大きく依存する。実施コスト大。
- プロトコル分析法：消費者の行動を言語化することによりその行動を分析する。顕在下の行動しか調べることができない。大量の調査には不向きである。
- データマイニング：POS データなど、データを用いることでその背後にある何らかの規則性を見出すことができる。大容量の高速コンピュータ環境が必要

^{†1} 神奈川工科大学情報学部情報工学科
Course of Information and Computer Sciences, Graduate School of Kanagawa Institute of Technology

^{†2} 神奈川工科大学情報工学科 スマートハウス研究センター
Research Smart House Research Center, Kanagawa Institute of Technology

となる。すでに存在しているデータを用いることもできる。従来は潜在的な部分を対象とした手法はほとんど存在しない。

- ・ IAT 法：潜在的な意識を調べることを目的とする。ブランドに対する無意識のイメージを明らかにするなど利用されている。

このように様々な手法が開発されているが、大量のデータを比較的安価かつ容易に対象とすることができ、同時に人間の潜在的な部分も対象とすることができる手法開発が望まれている。

2.1 IAT による潜在意識の分析

IAT では「良い悪い」や「好き嫌い」などの人の感情とのつながりが深いかどうか測定できる。典型的なシステム例を図 1 に示す。この図のように、「花・虫」というターゲット概念に対して、「ポジティブ・ネガティブ」といった連想の強さを測定するものである。連想の強さを外部から測定する基準としては、その反応（マウスをクリックした時間）の速さを用いることが一般的である。



図 1 IAT で用いる質問画面の例

2.2 データマイニングとデータ分析手法

データマイニング[10]とは、大規模なデータから有効な知識を抽出する技術である。通常のデータの扱い方からは想像が及びにくい新たな知識発見の獲得が可能である。データマイニングの手法の 1 つにクラスタリング (図 2) がある。

クラスタリングとは、分類対象の集合を、内的結合と外的分離が達成されるような部分集合に分割することである。大きく分けて最短距離法などの階層的な手法と、k-means 法などの分割最適化手法があり、データの集まりをデータ間の類似度にしたがって、いくつかのグループにわけるとをいう k-means 法によるクラスタリングの例を下記に示す。k-means 法は特定のクラスタ数における各クラスタサイズを最小化にすることができるため、クラスタの境界が明確に決定するという利点がある。

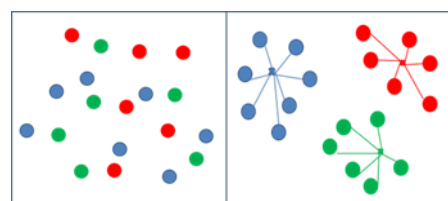


図 2 クラスタリングの概念図

因子分析とは、統計学的なデータ分析手法であり、アンケートデータの分析や心理学分野の分析手法として広く活用されている。観測結果 (変数) が真に観測したいものであるとは限らないため、真に観測すべき結果を潜在変数として仮定し、得られた観測変数の関係を分析することで潜在変数を見出すことになる。

3. システム提案と実験

先に述べたような従来手法の問題点を解決するためのシステムを提案する。その概要を図 3 に示す。

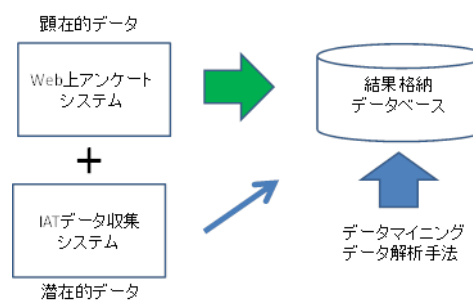


図 3 提案システムの概念構成

この図から分かるように本システムは 2 つのデータ収集部分から構成されている。顕在的なデータは通常と同様にして、Web 上でのアンケートシステムを用いて収集される。一方潜在的なデータは、Web を用いることは同様であるが、IAT の概念を取り入れて Web 上で指示されえる操作に対する反応時間を計測することで獲得される。これら両方の手法で獲得されたデータがデータベースに格納され、データマイニング等の手法を用いて分析される。

このシステムの評価を行うために実験を行ったのでその結果を以下に説明する。

3.1 従来手法による実験

大学生 20~25 歳の男女 5 人を対象にアンケートを行った。JAVA 言語で作成した 2 択アンケート YES (選択する) / NO (選択しない) で各設問の反応時間も同時に調査を行った。JAVA 言語で作成した GUI を図 4 に示す。



図4 JAVA言語によるアンケート

計測ありの場合は設問ごとの回答までに掛かった時間を取得できる。また、標準偏差を求め、ばらつきぐらいを確認した。5人とも標準偏差 1.0 以下のため、データを同じ視点で見ることができた。得られた各5人の平均クリック時間(秒)と標準偏差を表1に、各設問でのクリック速度グラフを図5に示す。

事前実験の結果から、設問ごとの回答までにかかる計測時間を「計測あり/計測なし」で見たとき、「計測あり」の方が回答者から得られる情報として、新たな、有益な情報となりうる情報であることは明らかである。また、回答者たちに対して、実験後に改めてアンケートの感想・評価を問い、問題点や不具合となる要素がないか確認した。その結果、作成した2択アンケートにおいて問題点や不具合となる要素がないことを確認し、次の実験へと移ることとした。

表1 5人の平均クリック時間(秒)

計測	A	B	C	D	E
なし	0	0	0	0	0
あり	1.35	1.09	1.33	1.24	1.71
標準偏差	0.45	0.18	0.49	0.44	0.81

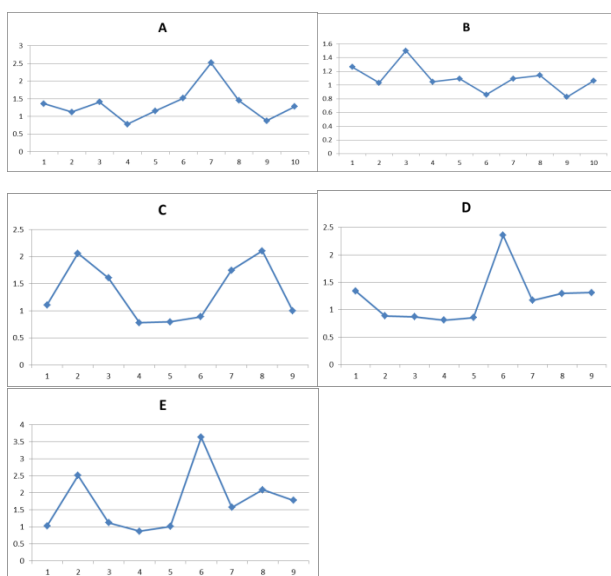


図5 回答者のクリック速度グラフ

3.2 実験1

従来手法による顕在的データ獲得と分析を行った。大学生20~25歳の男女16人を対象に2択アンケートを行った。実験1では、個人の好み、つまり質問された内容に対する個人の特性や意思決定の特徴の違いを検証する。アンケートの設問文を表2に示す。

表2 設問文

設問番号	設問文
Q1	音楽は聞きますか?
Q2	車は持っていますか?
Q3	ブランド物は好きですか?
Q4	漫画は好きですか?
Q5	アニメは好きですか?
Q6	ケーキは好きですか?
Q7	タバコは吸いますか?
Q8	あなたは暑がりですか?
Q9	あなたは寒がりですか?
Q10	ゲームは好きですか?

この実験で集めたデータをクラス数3, 5, 7に設定してk-means法を用いてクラスタリングを行った。その結果を図6に示す。クラス数3の場合、Q3の設問が他の設問と離れていることが見て取れる。クラス数5の場合、Q3, Q4, Q6で、クラス数7ではQ1, Q2, Q4, Q5が孤立する形となった。これらの結果から、設問Q3にクラスタリングした際に孤立する要因があると考えられる。

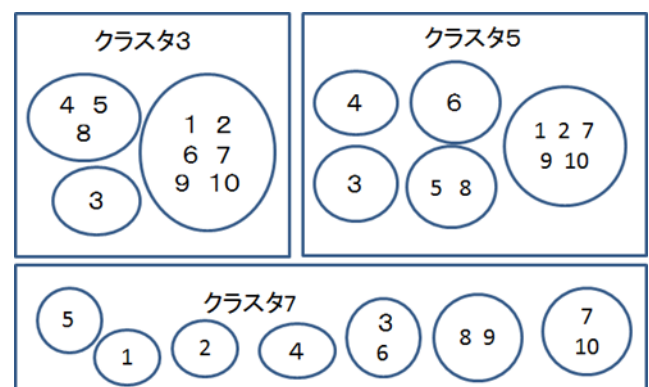


図6 クラス数3, 5, 7の結果

図7は実験1での計測時間を因子分析した結果である。大学生の14番, 15番の学部生が独立していることがわかる。残りの14人の学部生は、似たような結果を出していることから、14番, 15番の学部生が顕著に孤立しているといえる。

表3 設問文

設問番号	設問文
Q1	パソコンはよく使いますか？
Q2	プログラミングは好きですか？
Q3	勉強は好きですか？
Q4	数学は好きですか？
Q5	英語は好きですか？
Q6	携帯はよく使っていますか？
Q7	Twitterは使っていますか？
Q8	Facebookは使っていますか？
Q9	テレビはよく使いますか？
Q10	PS3は持っていますか？

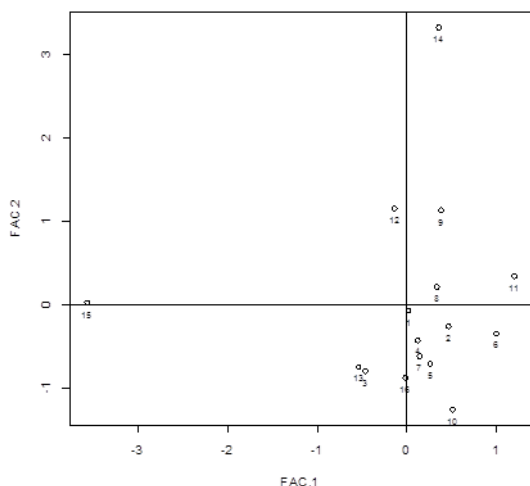


図7 計測時間の因子分析

次に紙媒体でのSD法による因子分析の結果を図8に示す。図8から読み取れるように、2極に分かれているように見受けられる。4番、11番、16番の学部生をグループ①としたとき、グループ①は似たような傾向にあるといえ、残りの学部生たち（グループ②とする）は似ていないといえる。しかし、よく見ると、8番、15番の学部生がグラフ軸のマイナス値を取っており、一概にグループ②に属しているとはいえない。

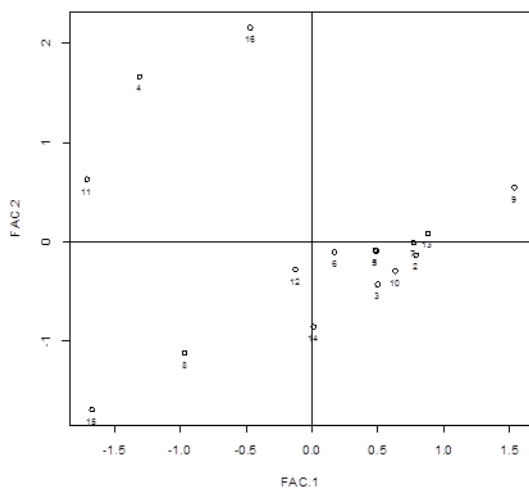


図8 SD法による因子分析

3.3 実験2

実験2では、回答者が情報系に属しているため、設問を情報系に関連するような内容にし、共通する状況での回答者全員の特徴があるかを検証した。アンケートの設問文を表3に示す。

実験1同様に、収集したデータをk-means法を用いてクラスタリングをした。その結果を図9に示す。クラスタ数3の場合、2個、3個、5個と分かれる結果となった。クラスタ数5の場合、Q7とQ3が孤立し、クラスタ数7の場合、Q2、Q3、Q4、Q6、Q7の設問が孤立した。

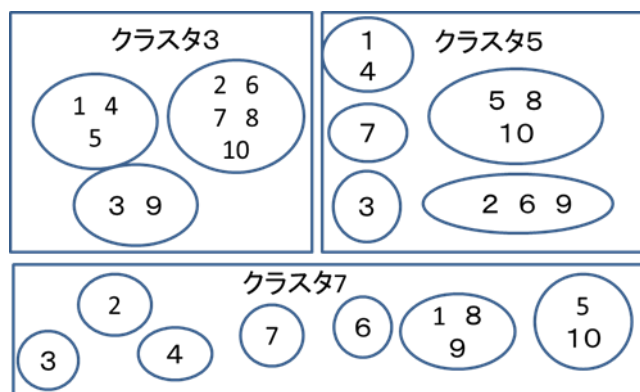


図9 クラスタ3, 5, 7の結果

実験2での因子分析の結果を図10に示す。実験2では今までとは異なる結果が出た。どの番号も全体的に散らばっており、容易にグループを付けることはできないことがわかった。しかし、6番と8番のような対極に位置する番号があることもわかった。対極に位置する番号には他にも13番と10番といった組み合わせがあるだろうとみて取れる。

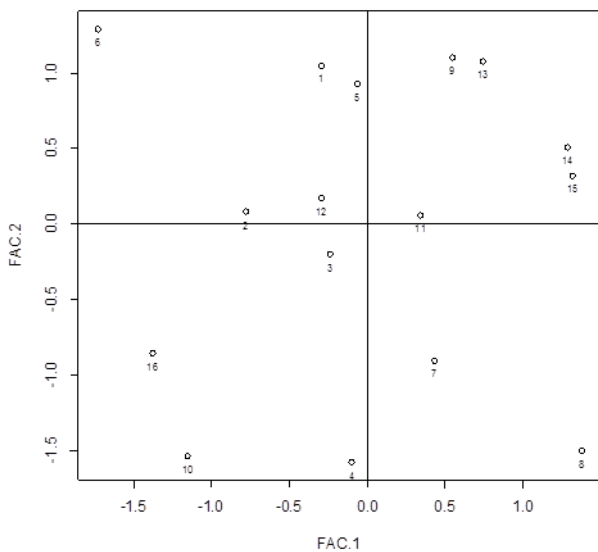


図 10 因子分析 紙媒体での SD 法

3.4 実験 3

学部生 20～25 歳の若者 16 人を対象に日常的な生活を中心にしたアンケートを行った。また実験 3 では、紙媒体である SD 法でのアンケートも行い、両者を分析し、潜在的な特徴が得られるかどうか検証した。その結果を表 4 に示す。

表 4 2 択アンケート結果

	3 年生 (7 人)	4 年生 (9 人)	16 人
合計	874.2	1331.9	2206.1
平均	1.24	1.47	1.37
最大値	8.62	11.62	11.62
最小値	0.18	0.17	0.17

紙の調査結果より、YES/NO 問題において「YES」と答えていながら SD 法では低い点数が出ていたり、「NO」と答えていながら、高い点数が出ていたり 2 つのアンケート結果に誤差が生じている問題が複数みられた。このような結果から、2 つのアンケートの解答の差が大きくなる設問をピックアップした結果、YES/NO の設問 5 段階評価の解答に違いが出る実験者は 16 中 11 人いることがわかった。実験結果から、紙式の 5 段階評価のときは時間があり、考えてしまうために、顕在意識が表れたといえる。反対に、2 択式の YES/NO のとき選択数がしかなく、またマウスでクリックするだけなので時間はあまり掛からない。無意識での判断のために、潜在意識が強く表に出たと考えられる。そのために、同じアンケート内容でもアンケートの方式によって、解答に差が出るような事態になったと考えられる。

4. おわりに

この論文では従来からのアンケート調査によるデータ収集に加えて、人間の潜在的な特徴データを収集できるシステムを提案した。実験 1 では回答者の好みや特性といったものを抽出しようとした結果、時間での因子分析と SD 法の因子分析とで異なる結果を得られた。似ている傾向にある回答者をグループ化することで、少なからず、設問から潜在的な特徴となりうる要素をいくつか見つけることができた。実験 2 では実験 1 と同様に設問を変えてアンケートを行った。しかし、実験 1 の結果とは異なり、反応速度を計測したデータでの因子分析では、分析しても結果を得ることができなかった。実験 3 ではアンケートから得たデータと SD 法でのデータとで、異なる部分を発見した。しかし、十分な潜在的特徴を得ることはできなかった。潜在的なデータを利用することで、従来のアンケート調査からは知ることのできなかつた特徴を分析することが可能となることを小規模な実験で確かめた。ここでの実験は小規模であるので今後は実用規模に拡大して精密な実験を行う予定である。

参考文献

- 1) Allwood, C.M. and Selart, M. Decision making: social and creative dimensions, (2001).
- 2) 竹村和久. 行動意思決定論, 日本評論社, (2009).
- 3) Bosson, J.K., Swann Jr, W.B. and Pennebaker, J.W. Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited?. Journal of personality and social psychology, Vol.79, No. 4, pp. 631-643, (2000).
- 4) Greenwald, A.G. and Banaji, M.R. Implicit social cognition: attitudes, self-esteem, and stereotypes. Psychological review, Vol. 102, No. 1, pp. 4-27, (1995).
- 5) Nisbett, R.E. and Wilson, T.D. Telling more than we can know: Verbal reports on mental processes. Psychological Review; Psychological Review, Vol. 84, No. 3, pp. 231-259, (1977).
- 6) Greenwald, A.G., McGhee, D.E. and Schwartz, J.L.K. Measuring individual differences in implicit cognition: the implicit association test. Journal of personality and social psychology, Vol. 74, No. 6, pp. 1464-1480, (1998).
- 7) Banaji, M.R. The nature of remembering: Essays in honor of Robert G. Crowder. pp. 117-150, (2001).
- 8) Greenwald, A.G., Banaji, M.R., Rudman, L.A., Farnham, S.D., Nosek, B.A. and Mellott, D.S. A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. Psychological Review; Psychological Review, Vol. 109, No. 1, pp. 3-25, (2002).
- 9) Nosek, B.A., Greenwald, A.G. and Banaji, M.R. Understanding and using the Implicit Association Test: II. Method variables and construct validity. Personality and Social Psychology Bulletin. Vol. 31, No. 2, pp. 166-180, (2005).
- 10) 元田浩, 津本周作, 山口高平, 池沼正行. データマイニングの基礎, オーム社, (2006).