

## インターネットを利用した安全かつ効率的なアンケート調査

酒井 淳一† 大野 浩之†

† 東京工業大学大学院 情報理工学研究科

### 概要

アンケート調査は問題解決の一つの手段として有効である。しかし、従来のアンケート調査は調査実施者の負担が大きく、個人でアンケート調査を実施するのは困難であった。近年、インターネットを利用して、個人でもアンケート調査を実施する例が増えている。しかしこのアンケート調査は、調査実施者の負担は軽減したが、アンケート調査で収集されるデータの信頼性が低下した。本稿では、インターネット上で安全かつ効率的にアンケート調査を実施するシステムを提案する。本システムはアンケート調査に要するほとんどの処理を自動化し、調査実施者の負担を軽減した。また簡単な機構で回答者のプライバシーを保護し、二重回答を防止した。そして信頼性の高いデータを収集できることを示した。

## A safe and efficient questionnaire on the Internet.

Jun'ichi SAKAI† Hiroyuki OHNO†

† Graduateschool of Information Science and Engineering, Tokyo Institute of Technology.

### abstract

Questionnaire is effective for solving problem. However investigation by sending out questionnaires is hard work and it's too difficult for individual to investigate by sending out questionnaires. Today, it has been increasing for individual to investigate by sending out questionnaires on the Internet. While the method reduce work of investigator, a value of collected information decrease. In this paper, we describe a system which send out questionnaires on the Internet safely and efficiently. The system reduce work of investigator and is privacy-oriented and prevent from multiple answer. We show that the system provide investigator a high value of collected information.

## 1 はじめに

アンケート調査は、問題解決のために必要な情報を得る目的で実施される。アンケート調査は人から直接情報を得るので、非常に有益な情報が得られる。現在では、紙の調査票を用いたアンケート調査が普及している。しかしこの方法は、調査を実施する際のコストおよび負担が大きい。特に、回収率を向上させようとすると、調査票の回収に大きな負担がかかる。そのため、個人でアンケート調査を実施するのは困難である。

調査実施者の負担を軽減する手段として、インターネットを利用することが考えられる。近年、WWWや電子メールを利用して、個人でもアンケート調査を実施する例が増えている。しかしこれらのアンケート調査は、回答者がどんな属性を持っている人かわからず、回答率も算出できないことが多い。よって、収集されるデータの信頼性は逆に低下した。

本稿では、インターネットを利用したアンケート調査の問題を解決する新たなアンケート調査実施システムを提案する。本システムには、調査実施者の負担を軽減し、信頼性の高いデータを得るための機構を組み

込んだ。

## 2 インターネットを利用したアンケート調査

### 2.1 アンケート調査の実施例と問題点

インターネットを利用したアンケート調査の例としては、WWWを利用したものや電子メールを利用したものなどがある。以下にその特徴を述べる。なお、本稿では、アンケート調査で収集されるデータを「調査データ」と呼ぶ。

WWWを利用した方法では、調査実施者がWWW上に調査票を公開し、そのページにアクセスした回答者が質問に回答する。従来のアンケート調査に比べて調査実施者の負担が軽減される半面、アンケート調査を実施する度にHTMLコードとCGIプログラムを書く手間がかかる。また、回答者がどんな人であるかわからず、回答率もわからない。このため、信頼性の高い調査データを得ることはできない。二重回答も防止できないことが多い。

電子メールを利用した方法では、調査票の配布およ

び回収に電子メールを利用する。母集団の人数がわかるので、回答率を算出できる。しかし、調査実施者があらかじめ回答者の電子メールアドレスを取得しなければならない。また、回答者の情報は、電子メールアドレス以外はほとんどわからず、特定の母集団にアンケート調査を実施することは困難である。回答者は匿名で回答できないので、回答率が低くなることが予想される。調査データのフォーマットを回答者が容易に書き換えることができるため、調査実施者が調査データを解析する際の負担が大きい。

以上の事柄から考察すると、インターネットを利用したアンケート調査は、紙による従来のアンケート調査に比べて調査実施者の負担が軽減される。また調査票を配布する時間および場所の制限を受けないという利点もある。しかし、信頼性の高い調査データを得るのは困難である。

## 2.2 本研究の目的

前節で述べた問題を解決するために、本研究は調査実施者の負担を軽減し、信頼性の高い調査データを得ることを目標とする。

現在、インターネットを利用したアンケート調査に関する研究は、匿名回答の実現に関する研究に偏りすぎている。この実現により、回答者のプライバシーを保護できるが、現在インターネット上では、ある特定の母集団に対してアンケート調査を実施することすらできていない。これでは、回答率の算出も困難であり、信頼性の高い調査データを得ることができない。そこで本研究は、特定の母集団に対してアンケート調査を実施すること一番の目的とする。

匿名回答の実現は、筆者らも非常に重要であると考える。しかし、匿名性を維持すると二重回答を防止できず、逆に二重回答を防止するためには匿名性が失われるという問題点がある。現在ではその問題を解決した技術がいくつかあるが [2]、その技術が実際にインターネット上での選挙やアンケート調査に利用された例は聞かない。その理由は、認証や暗号技術を駆使した複雑なものが多く、調査実施者や回答者に受け入れられにくいからである。そこで本研究では、簡単な仕組みで二重回答を防止し、匿名回答を可能にする機構を設計および実装した。

また、本研究ではアンケート調査の信頼性を高めるために、回答率の向上をめざす。回答率が低い理由は「個人の秘密を知られたくない」、「面倒である」が上位を占める [3]。そこで、本システムは回答者のプライバシーを保護し、WWWブラウザを使用してマウス

操作のみで簡単に回答できるようにした。

最後に、本研究は調査実施者の負担の軽減をめざす。従来のアンケート調査は調査実施のコストおよび負担が大きかった。個人が手軽にアンケート調査を実施できる必要である。

筆者らは、まず単一の組織内でアンケート調査を実施するシステムを設計・実装した。次に大規模な範囲でアンケート調査を実施するためにシステムを拡張した。本稿では、拡張前のシステムを「第一段階システム」、拡張後のシステムを「第二段階システム」と呼ぶ。現在は第一段階システムの実装を終え、第二段階システムを開発中である。

## 3 第一段階システムの設計と実装

### 3.1 第一段階システムの概要

第一段階システムはアンケートクライアントとアンケートサーバの二つのプログラム、回答者データベースと回答データベースの二つのデータベースから構成される。回答者データベースには、アンケート調査に回答してもよい人の年齢・性別などの属性が登録されている。回答データベースには各回答者からの調査データが「調査データファイル」という形式で格納されている。調査データファイルはアンケート調査ごとに作成され、回答者からの調査データが書き込まれる。

第一段階システムの特徴は調査実施者と回答者間の情報交換は全て第三者を介して行なわれることである。第一段階システムでは、調査実施者と回答者の間にアンケートサーバを仲介させ、アンケートサーバがアンケート調査に要するほとんどの処理を行なう(図1)。

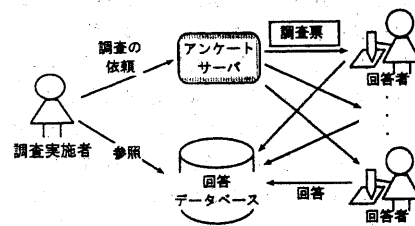


図1: システムの概要

### 3.2 回答者インターフェース

回答者に対するインターフェースとなる調査票を「楽々調査票」と名付けた(図2)。楽々調査票はHTML

で記述され、その URL が回答者に電子メールで届けられる。楽々調査票により、回答者は WWW ブラウザ上でマウスボタンをクリックするだけで回答できる。さらに楽々調査票により、各回答者からの調査データのフォーマットを統一できるので、調査実施者が調査データを解析する際の負担も軽減される。

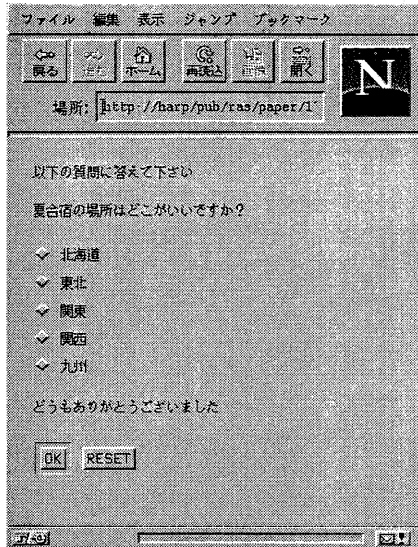


図 2: 楽々調査票

### 3.3 調査実施者インターフェース

調査実施者に対するインターフェースはアンケートクライアントである。アンケートクライアントはコマンドライン・電子メール・WWW から利用できる。アンケートクライアントはアンケートサーバにアンケート調査を依頼し、アンケートサーバから調査データファイルの URL を受信する。

調査実施者がアンケートクライアントに与える情報は、自分の名前、調査対象の定義、「楽々調査票定義コード」のみであり、調査実施者が回答者を選出する必要はない。これは調査実施者の負担を軽減するだけでなく、回答者のプライバシー保護にもなる。

楽々調査票定義コードは、楽々調査票を作成するために必要な情報を定められた文法で記述されたものである。調査実施者は楽々調査票そのものを作成する必要はなく、楽々調査票定義コードを記述すればよい。

調査データファイルの URL は、アンケートクライアントをコマンドラインから利用した場合は端末上に

表示され、電子メールから利用した場合は電子メールで送信され、WWW から利用した場合は WWW ブラウザ上に表示される。調査実施者は WWW ブラウザを利用して調査データファイルを参照する (図 3)。

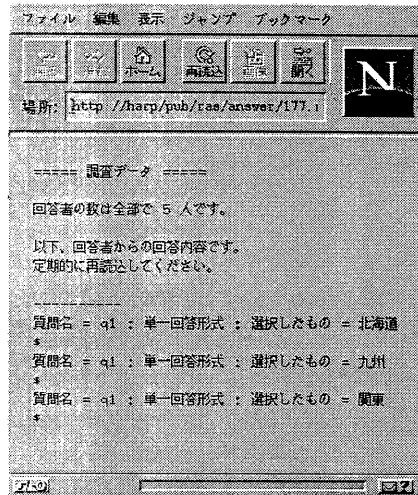


図 3: 調査データ

### 3.4 アンケートサーバ

アンケートクライアントからアンケート調査を依頼されたアンケートサーバは、まず回答者データベースを参照して、調査対象の定義に合う回答者を抽出する。次に楽々調査票定義コードから楽々調査票を作成し、同時に回答処理プログラムを生成する。この二つは回答者ごとに作成される。回答処理プログラムは楽々調査票のフォームデータを処理する CGI プログラムである。そして、回答者に楽々調査票の URL を電子メールで送信し、アンケートクライアントに調査データファイルの URL を送信する。最後に、調査データファイルを作成し、回答者の数を調査データファイルに書き込む (図 4)。

回答者が楽々調査票の質問に対して回答すると、その回答は回答処理プログラムへ送られる。そして回答処理プログラムは、後述する機構により、その回答が二重回答でないことを確認し、調査データを調査データファイルに書き込む (図 5)。調査データファイルに書き込む情報は、回答者からの回答のみで、回答者の個人情報は一切含まれない。したがって調査データからは回答者を判別できない。

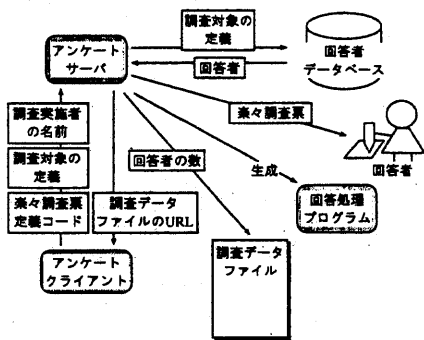


図 4: アンケートサーバの動作

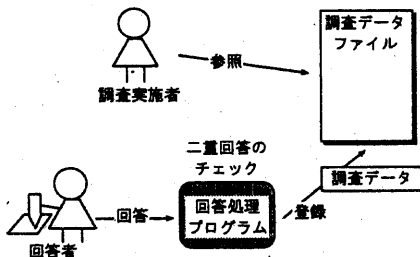


図 5: 回答処理プログラムの動作

### 3.5 回答者のプライバシー保護と二重回答の防止

第一段階システムでは、次の二つの機構で、回答者のプライバシーを保護する。

まず第一に、調査実施者が回答者の電子メールアドレスを知ることができない機構を用意した。これは第三者であるアンケートサーバが調査実施者と回答者を仲介することにより実現されている。回答者の選出はアンケートサーバが行なうので、調査実施者は回答者の電子メールアドレスを知る必要がなく、また知ることもできない。

第二に、第一段階システムは匿名回答を実現することで、回答者のプライバシーを保護した。匿名回答は回答処理プログラムにより実現されている。さらに本システムは簡単な仕組みで二重回答を防止した。第一段階システムは、回答者ごとに楽々調査票を作成し、楽々調査票ごとに回答処理プログラムを作成している。回答処理プログラムは、自分自身が一回実行されると、自分自身を消去し、自分が担当している楽々調査票を書き換える。この仕組みにより、同一回答者からの二重回答を防止する。また、回答者は楽々調査票の質問

に回答するだけでよく、鍵の交換やパスワードの入力などの作業は一切不要である。

### 3.6 楽々調査票の作成

楽々調査票を作成するには、HTML コードと CGI プログラムを書かなくてはならない。本システムは楽々調査票定義コードから、HTML コードと CGI プログラムを出力する機構を用意した。楽々調査票定義コードは楽々調査票記述言語で記述される。楽々調査票記述言語は、楽々調査票の作成に特化した言語として、筆者らが開発したものである。図 6 に図 2 の楽々調査票を作成するための楽々調査票定義コードを示す。

```

1: define q1
2: {
3:   form "tanitsu_kaitou";
4:   question "夏合宿の場所はどこがいいですか?";
5:   choice "北海道";
6:   choice "東北";
7:   choice "関東";
8:   choice "関西";
9:   choice "九州";
10: }
11:
12: main()
13: {
14:   print("以下の質問に答えて下さい");
15:   print(q1);
16:   print("どうもありがとうございました");
17: }
18: $

```

図 6: 楽々調査票定義コード

この機構により、調査実施者の負担が軽減されるだけでなく、調査実施者はさまざまな回答形式の質問を作成できる。選択肢の中から一つだけ選択する「単一回答形式」、選択肢の中から複数選択できる「複数回答形式」、選択肢に順位をつける「順位回答形式」などの質問を、楽々調査票記述言語を用いて、同じフォーマットで記述することができる。

## 4 第一段階システムの評価と考察

### 4.1 運用結果と考察

第一段階システムの評価のため、筆者らの研究室で試験運用した。今回の運用では、回答者データベースに研究室の構成員全員を登録した。

第一段階システムにより、調査実施者はアンケート調査を手軽に実施できた。その理由は以下の 3 つが考えられる。第一に、調査実施者は調査票の配布および

び回収をする必要がないこと、第二に、調査実施者は楽々調査票を作成する必要がないこと、第三に、調査実施者は調査対象者リストを得る必要がないことである。

今回の運用では、調査対象の定義として、「ネットワーク管理者」、「修士課程の学生」、「学部学生」などが指定できる。これにより、特定の母集団に対し、アンケート調査を実施できた。母集団の人数は調査データファイルに記されるので、回答率も算出できる。

しかし、アンケート調査が手軽に実施できるようになったことから、それほど重要でないアンケート調査が増えた。これにより逆に回答者の負担が増えた。そのため、回答者が回答しない場面も多かった。楽々調査票は回答する手間はそれほどかからないが、楽々調査票を見るために毎回WWWブラウザを起動しなければならない、それが回答者の負担になっていることがわかった。

アンケート調査の数が増えた理由は、アンケート調査を手軽に実施できるようになったこともあるが、第一段階システムの運用がまだ実験段階であり、運用実績も浅いことから、このシステムを利用したい人が多いからであると考えている。したがって、この問題は次に減少すると予測している。また、この問題を解決するために、回答者に送信する電子メールのテキスト部分にアンケート調査の概要を記述する予定である。これにより、回答者はWWWブラウザを起動する前にアンケート調査の内容がわかり、回答すべきかどうかの判断ができる。

## 4.2 拡張の必要性

第一段階システムは筆者らの研究室のみで運用されている。今後、大規模な範囲でアンケート調査を実施していく必要がある。

しかし、大規模な範囲で第一段階システムを運用するには問題があることがわかった。今回の運用では、アンケート調査は同一組織内で実施されており、アンケートサーバはその組織内で動作し、回答者データベースはその組織内に置かれていた。調査対象の範囲が複数の組織にまたがる場合は中立的な第三者機関を用意し、そこでアンケートサーバを起動させる必要がある。第一段階システムではアンケートサーバが回答者データベースを管理しており、アンケートサーバが第三者機関で動作する場合、回答者データベースは第三者機関に置かれる。しかし、第三者機関に回答者データベースを置くことは、回答者のプライバシー保護のためにも避ける必要がある。

したがって、第一段階システムは同一組織内でのアンケート調査には有効であるが、大規模なアンケート調査には問題がある。そこで、第一段階システムを拡張した第二段階システムを設計した。

## 5 第二段階システムの設計

### 5.1 第二段階システムの概要

第二段階システムは、第一段階システムのアンケートサーバを各組織で動作させる。回答者データベースはアンケートサーバが管理する。これにより回答者データベースは回答者の所属する組織に置かれる。第三者機関にはアンケートサーバを取りまとめるチーフサーバを置く(図7)。回答データベースはチーフサーバが管理する。第二段階システムは、第一段階システムの利点を全て引き継ぐ。

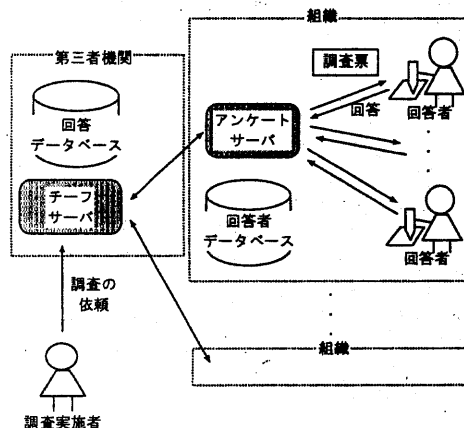


図7: 第二段階システムの概要

### 5.2 第二段階システムの動作

第二段階システムでは、調査実施者および回答者のインターフェースは第一段階システムと同じである。

アンケートクライアントは、通信する相手が調査実施者とチーフサーバになる以外、第一段階システムと同じ動作をする。

チーフサーバは、アンケートクライアントから受信した情報を各組織で動作するアンケートサーバに送信し、アンケートサーバおよび回答処理プログラムから受信した情報を回答データベースに登録する。チーフサーバとアンケートサーバとの情報交換には常に調査データファイルのURLを伴う。アンケートサーバお

よび回答処理プログラムから受信した情報は調査データファイルに登録されるため、調査データファイルのURLも一緒に受信しなければ登録できない。

アンケートサーバは第一段階システムのアンケートサーバとほぼ同じ動作をする。回答者の数はチーフサーバに送信する。回答処理プログラムも第一段階システムとほぼ同じ動作をする。調査データはチーフサーバに送信する。調査データファイルはチーフサーバのみが書き込むことができるので、このような仕組みになる。

図8に第二段階システムの動作の流れを示す。

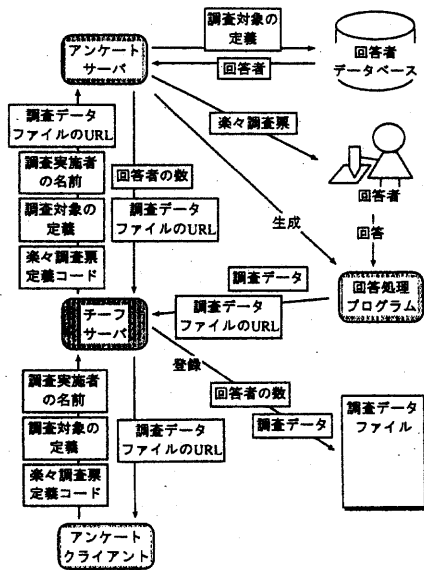


図8: 第二段階システムの動作

## 6 今後の展開

### 6.1 本システムの応用

本システムは筆者らの研究室内で、システム管理者間の意志決定にも利用されている [4]。本システムは、特にネットワーク障害が発生した場合などの、緊急時における対処方法の決定に重要な役割を果たす。問題を発見した管理者は、すぐにアンケート調査を実施し、調査データから対応策を決定する。この場合、回答者である管理者はすぐに質問に答える義務がある。本システムを利用すれば、これらのことが容易に行なえる。

しかし、緊急時に計算機を利用できない管理者もあ

り、WWWブラウザだけでなく電話やFAXからも楽々調査票を参照でき、その質問に回答できるようにする必要がある。また、公衆情報端末PICKLES[5]を利用して、Chipcardからも楽々調査票を参照できるようにする予定である。

### 6.2 今後の課題

現在用意している回答形式は少なく、楽々調査票の上でさまざまな質問を作成できるようにする必要がある。またグラフを用いた質問など、紙のアンケートでは実現しえなかった回答形式も開発していく。

現在は調査データのフォーマットがあらかじめ決められている。調査実施者が調査データのフォーマットを指定できるようにする予定である。

また現在では、調査対象の定義として、一つの属性しか指定できない。「学生」かつ「20歳以上」といった、複数の属性の積集合や和集合を指定できるようにする予定である。

## 7 おわりに

本稿では、インターネットを利用して安全かつ効率的にアンケート調査を実施するシステムについて述べた。第一段階システムで、信頼性の高いアンケート調査を実施できることを示した。また調査実施者および回答者の負担を軽減できることを示した。

次に、大規模なアンケート調査を実施できる第二段階システムを提案した。今後は第二段階システムの実装を進めていく。

## 謝辞

多くの助言を頂いた国立がんセンターがん情報研究部の中川晋一先生に感謝します。

## 参考文献

- [1] 辻 新六, 有馬 昌宏. アンケート調査の方法. 朝倉書店, 1987.
- [2] 法邑 茂夫, 堺 隆一, 笠原 正雄. 通信ネットワーク上における経路情報を秘匿する二, 三の方式. 信学技報, ISEC93-66, pp7-11. 電子情報通信学会, 1993.
- [3] 浜砂 敬郎. 統計調査環境の実証的研究. 産業統計研究社, 1990.
- [4] 大野 浩之. システム管理者が離散した状況下におけるシステム管理手法. 分散運用技術研究会予稿集, pp.13-18. 情報処理学会, May 1997.
- [5] 木本雅彦, 大野 浩之. 街角公衆情報端末計画 ~ PICKLESの概要 ~. 情報処理学会第52回(平成8年度前期)全国大会講演論文集(6), pp.381-382. 情報処理学会, 1996.