

P2P 技術を利用したアンケート集計システムの設計と開発

三船真哲* 山田嶺* 澤津健吾* 小坂慶和* 綿貫理明* 飯田周作*

*専修大学ネットワーク情報学部

mail: *{n130203, n130048, n130058, n130068, watanuki, iida}@isc.senshu-u.ac.jp

インターネットを利用したアンケート集計は、街頭でのアンケート集計方法に比べ、手軽に行うことができ、配布、回答などのプロセスを全て電子的に扱うことができる。しかし、作成者が回答者にアンケートの実施自体を通知する事は難しい。集計専用のサイトを定義することでこの問題は解決できるが、サーバの負荷など、新たな問題が発生する。本稿ではこれら諸問題を解決するべく、P2P 技術を用いた集計システムを提案する。

Design and Development of the Questionnaire Processing System Using P2P Technology

Masatetsu Mifune* Ryo Yamada* Kengo Sawatsu*

Yoshikazu Kosaka* Osaaki Watanuki* Shusaku Iida*

*Network And Information, Senshu University

mail: *{n130203, n130048, n130058, n130068, watanuki, iida}@isc.senshu-u.ac.jp

Web-based questionnaire processing can be easily performed compared with the conventional paper-based methods, normally conducted in a street or by postal service. Furthermore, all processes such as distribution of the questionnaire and a reply can be completed electronically. However, the new method also has some problems. It is difficult for the originator to inform of the common respondents to start the questionnaire. It could be solved by defining the site, but a new problem arises, such as the load of the server. In this paper, we propose the new questionnaire processing system using P2P(Peer to Peer) technology to give a possible solution to the above problems.

1 はじめに

集計は、社会的重要度の高い調査手段の一つである。国勢調査、企業の意識調査、視聴率の調査など、集計によって得られたデータ

は、政治や経済、そしてマスコミや個人にまで、良くも悪くも大きな影響を与えることができる。集計は昔から頻繁に行われているものであるが、年代が進むにつれ、新たな集計

方法が登場し、その手段が多様化してきている。以前は街頭アンケートや店舗内での調査など、人手によって行われるものだけであったものが、コンピュータとネットワークを通じて、誰でも簡単にアンケートを実施することができる時代となった。

しかし既存の集計方法は、そのモデルにおいて様々な問題点がある。従来通りの人手の集計には人件費などのコストの問題と、集計結果の信頼度の問題、アンケートの実施場所が限定され、遠隔地のユーザが回答できない、もしくは実施自体を認識しない問題などがある。また、インターネットを通じたアンケート集計では、データの再利用性の問題、作成者がアンケートの実施を回答者に通知する手段が少ないなどの問題がある。

本稿では、現在のインターネットを用いたシステムによる集計方法の問題点を指摘し、その上で、それら諸問題を解決する P2P モデルのアンケート集計システムを提案し、負荷の分散とシステムの汎用性を中心に目指した新たな集計方法を示す。

2 既存の集計モデルの諸問題

計算機を用いた集計モデルを設計するにあたり、既存の集計手段の検討を行った。

2-1 アンケート集計の実施方法の多様化

冒頭でも述べた通り、アンケート集計は古くから行われているが、現在では計算機を利用し、回答者の回答データをコンピュータに入力するだけで瞬時に集計結果が得られる方法がある。これは、人件費や集計作業の労力など、既存の集計方法の問題であったコストの削減を実現する。さらに、インターネッ

トの普及によってアンケート収集のプロセスである作成・配布・回答・集計が全て電子的に執り行えるようになると、オンラインであればいつでもどこでもアンケート収集に携わる事が可能となり、アンケートの実施場所の限定、時間の制約の問題を解消することができるようになった。

紙媒体でのアンケート収集が、電子的な形態に代わることにより、コストと場所の問題を解決することが可能となる。

2-2 近年の集計モデルの問題点

アンケート収集の全プロセスがインターネットを通じて実現可能となった今、アナログのアンケート収集での問題点であった「時と場所の制限」は解決できた。しかし、コンピュータで集計を実現するにあたって、新たな問題が発生する。

作成者がアンケートの実施を回答者に知らせることは、街頭での調査などでもコストがかかるものであったが、インターネットを利用することでより困難になったと言える。現在、インターネット上には数十億ものサイトが存在する。その中で、作成者が自分たちのアンケートの実施を回答者にアピールする手段が少ないのである。大抵、多くの回答者を獲得することができるのは、ポータルサイトなど、アクセス数の多いサイトに留まっており、個人で運営しているサイトの集計などは、その存在すら認識されないことが多い。結果、作成者の多くはコミュニティサイトでの宣伝など、アンケート収集において多大な労力を消費しなければならない。

実施者のサイトの認知度が高いものと低いものとの間でサンプル数の差が発生する。

誰もが場所を選ばずアンケート収集を行える今、インターネット上に存在する膨大なアンケートを、回答者が検索することは非常に困難である。

以上の問題を解決する、アンケート集計専門のポータルサイトが既に存在する[1]。この「e-リサ+(プラス)システム」は、不特定多数のアンケート作成者が同一のサイトでアンケートを行うことで、上記の問題を解決している。また、作成者が簡単にアンケートを作成・公開・集計できるようなシステムになっているため、サイト公開の技術を知らない作成者でも手軽にアンケート収集を行うことができる。集計専用のポータルサイトとして位置づけ、簡単に誰でもアンケートが作成できるシステムにすることで、これまでの問題はほぼ解決できる。

しかし、上記のシステムには、システムにかかる負荷の問題が発生する。アナログの集計における人手のコスト同様、集計専門のポータルサイトのサーバには、多大な計算量の負荷が発生するのである。このシステムではアンケート収集の全プロセスがサーバによって処理されるため、不特定多数のアンケートがサーバに集計処理の要求をした場合、サーバがダウンしてしまう可能性がある。

3 提案モデルの検討

2の問題を受け、ポータルサイトのようなメリットを持ちつつ、負荷の分散を実現するモデルを検討した。本稿で提案する集計モデルは、上記の負荷の問題を P2P と呼ばれる通信モデルで解決する。

提案するモデルは、既存のインターネットを利用した集計システムのように Web 上でアンケート収集を行わず、代わりにアンケート集計専用のアプリケーションをユーザの各マシンにインストールする。

アプリケーションとして動作することにより、回答者はアンケート集計専門のポータルサイト同様、実施されているアンケートを容易に見見することができる。また、集計のプロセスを各作成者のマシンで実行させることで、不特定多数のアンケート集計によるサーバの負荷を大幅に低下させることができる。以下に問題解決のための技術及び具体的な集計モデルを示す。

アプリケーションとして動作することにより、回答者はアンケート集計専門のポータルサイト同様、実施されているアンケートを容易に見見することができる。また、集計のプロセスを各作成者のマシンで実行させることで、不特定多数のアンケート集計によるサーバの負荷を大幅に低下させることができる。以下に問題解決のための技術及び具体的な集計モデルを示す。

3-1 P2P 技術の採用

サーバに負荷がかかる理由は、その通信モデルにある。クライアントは自身で処理を行わず、サーバへ処理を全て任せる。サーバは、全てのクライアントの要求に対して結果を返すべく計算を行わなければならない。このようなクライアント・サーバモデルでは、サーバに負荷がかかることは自明である。今回通信モデルとして採用する P2P は、クライアント・サーバモデルのようにサーバにリソースを集中させず、エンドユーザ間で直接通信を行うモデルであり、一つひとつの P2P アプリケーションがクライアントとサーバの両方の機能を兼ね備えている。1つのアプリケーションをピア (Peer) と呼び、この言葉には「対等な関係」という意味がある。また、文献[2]にあるように、P2P にはサーバを全く介することなく通信を行うことが

出来るピア型 P2P と、各ピアの情報のみをサーバに通知するハイブリッド型 P2P がある。本システムは、サーバにピアの情報のみを保存するハイブリッド P2P モデルを採用することで、サーバの負荷軽減を実現する。また、同モデルによって、セキュリティをも高めている（3-4 で解説）。

3-2 P2P の違法性解決

既存の P2P アプリケーションに対するユーザの利用用途は、文献[2]にあるように、ファイルの交換自体を目的としているものが多く、著作権を侵害する場合が少なくない。最近では、国内でも P2P を利用したファイル共有ソフトで逮捕者が出たことは記憶に新しい。本システムは、このような P2P 技術に関する違法性を、以下のような利用方法で解決することができる。

利用目的の限定

既存のファイル共有アプリケーションは、音楽ファイルや動画ファイル、ソフトウェアに至るまであらゆるデータを交換することができた。それに対して本アプリケーションは、データ共有に関してはアンケートファイルのみであるため、既存のアプリケーションのような利用をされることが無い。

データの暗号化・電子透かし

データファイルから直接バイナリデータなどのデータが抽出された場合や、アプリケーション内での画像表示・動画再生を違法にキャプチャーされた場合には、が無効化され、既存のファイル共有アプリケーションと同様の利用をされかねない。このため、保存されるデータファイルは暗号化を施すことにより、アプリケーション外からのデータ参

照を防止する。また、画像には電子透かし[3]、動画・音声には雑音の挿入など、不要なデータを組み込むことで違法なデータ共有を防ぐことができる。さらにテキストデータに関しては、選択範囲でのコピーを禁止する。これにより、全てのデータに対してコピーという行為を防止することが可能となり、回答者における、アンケートデータへの関与は閲覧までに留めることができる。

3-3 共有データの汎用性

プログラムのデータ交換に利用するデータ構造は、他アプリケーションとの親和性の観点から XML を採用する。実際のデータ自体は暗号化され、アプリケーション以外からのデータ参照は不可能である。本システムで実装できなかった複雑な集計を、他のアプリケーションで行いたい場合や、データを別の利用目的で使用したい場合には、データの再利用が必要である。そこで、データの参照が容易な XML を採用することで、再利用を実現する。データの内部構造は、データ上部にアカウント情報やファイル情報などのヘッダ情報、下部に実際のデータが定義される。

3-4 セキュリティ

集計専用のモデルであるが故に、不正なアンケートデータへの対処が必要となってくる。やりとりされるアンケート収集は、健全な利用目的であるべきだが、他人への誹謗中傷を記述したアンケートや、無意味な文字列が続くアンケートなど、不正なデータがネットワーク上に流れる可能性は否定できない。

対策としては、サーバにアカウント管理を徹底させることである。アカウントの情報はサーバが管理しているため、不正なデータま

たはアカウントの登録を抹消するなどの処置を取ることができる。また、アカウントの取得の際、認証機能を強化することによって、正しい目的意識を持ったユーザが利用する健全なシステムにすることができる。

3-5 システムが社会に浸透するための工夫

システムの機能を十分に考慮し、実装したとしても、実際に社会で利用されなければシステムの価値は無い。以下は少しでも多くの回答者と作成者を獲得するための、システム側からのアプローチである。

見返りの検討

多くの回答者は、自らアンケートに答えるために、アンケートを捜し求めることはない。大抵、回答者は何かしらの見返りを求めてアンケートに回答する。企業では、そのような見返しとして懸賞つきアンケートを実施し、多くの回答を得ている。

本システムでもこのような見返しを実現するため、アプリケーションに広告を表示し、得られる広告費を回答者に還元する回答ポイント制を考案した。具体的には、回答者がアンケートファイルを取得して回答する際、アンケートの分野に応じた広告を掲載する。ユーザには、アンケートに回答するたびにポイントを付与し、ポイントが増えるほど多くの金額をユーザが手に入れることができる、というようなものである。また、企業が本システムを利用するようであれば、従来の企業アンケートの方針をそのまま取り入れ、懸賞モデルを実装することもできる。

サンプルの選定

アンケートを実施する対象者は、不特定多数の場合もあれば、社内の人間に限定したアン

ケートも存在する。このような、サンプルの選定を実現するため、作成者がアンケートを登録する時点で、対象者の設定を行う。具体的には、回答可能なユーザの IP アドレスの限定である。アドレスを限定することにより、想定した対象者のみにアンケートを配布することができる。

3-6 分野定義

アンケートには、スポーツ、哲学など様々な分野がある。従って、分野に基づく検索を行う際に、システムの分野の定義が緻密なものでなければ、回答者は意図した分野のアンケートを取得することができない。本モデルは、現時点では既存のポータルサイトの分野定義を採用している。また、ユーザから新たな分野定義の要求があった場合、柔軟に対処できるシステム構築が必要となる。これについては、ファイル情報を管理するサーバに分野項目を追加する。サーバに情報を追加することで、ユーザが所持するアプリケーションが瞬時に更新を反映することができる。

3-7 ユーザビリティ

提案するモデルは汎用的な集計システムであるため、作成者はアンケートのデザインを考慮する必要がない。従って、デザインによるアンケートの質の差は無くなり、純粋なデータとして扱うことができる。しかし文献[4]が述べている通り、システムのユーザインタフェースの設計が悪く、ユーザビリティが低いものであれば、回答者はアンケートどころかアプリケーションですら利用しない可能性がある。ユーザの趣味に応じたスキンセレクトの実装、さらにはアプリケーションの使いやすさを考慮した設計が求められる。

4 プロトタイプの実装

3で述べたモデルを参考に、プロトタイプを作成した。本プロトタイプは、数多くのピア（アンケート作成者と回答者）と、アンケートファイルを所有するピア存在を管理するサーバで構成されている（図1参照）。作成者側のピアからアンケートを取得し、回答後に直接作成者に送信する形態を取っている。以下にプロトタイプのプロセスを示す。

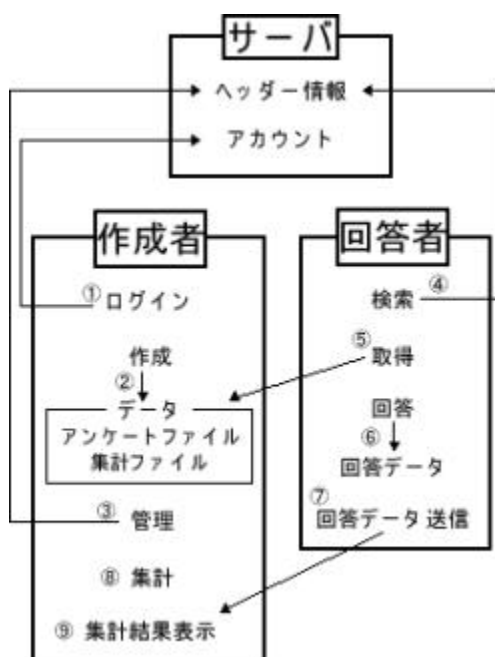


図1 本プロトタイプ構成図とプロセス

ログイン（作成者）

作成者はアプリケーションの起動時にログイン手続きを行う。ログイン後、作成済みのアンケートファイルのヘッダ情報はサーバに登録される。つまり、データベースの情報には、現在ログインしている作成者の有するファイル情報のみが入ることになる。従って、回答者は確実に存在するファイルを取得す

ることができる。

アンケート作成（作成者）

アンケートの作成は、専用のアンケート作成フォーム（図2参照）を用いて行う。アンケートのタイトル、分野を入力し、複数の質問文と対応する選択肢を入力する。入力後、作成ボタンを選択することでアンケートファイルが作成される。また、アンケートの集計結果を保存する集計ファイルも同時に作成される。

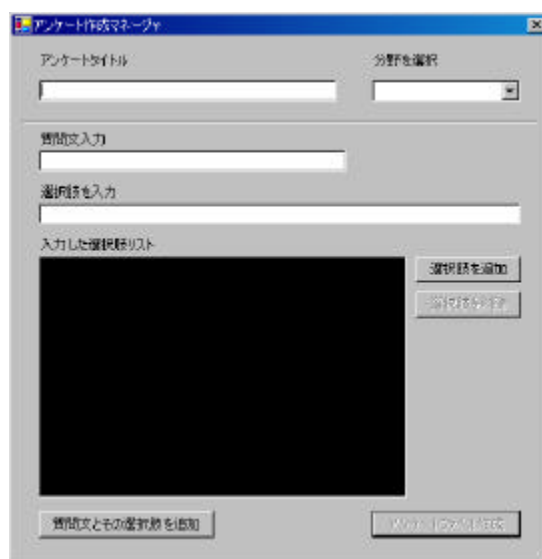


図2 アンケート作成フォーム

アンケート管理（作成者）

アンケートは、作成してすぐに第3者へ公開されるわけではない。アンケートの公開や非公開はアンケート管理メニューから行う。これは、公開する前にアンケートの修正などを行えるようにするためである。また、公開中のアンケートを非公開にすることや、アンケートの削除も管理メニューから行う。

アンケート検索（回答者）

回答者は公開されたアンケートを、アンケートのタイトル、分野などの条件に基づいて検索する。検索を実行すると、プログラムはサーバに問い合わせ、条件を満たす情報を表示する。

アンケートデータ取得（回答者）

回答者は公開された作成者のアンケートファイルを、アンケートのタイトル、分野で検索する。検索を実行すると、プログラムはヘッド情報を登録したサーバに問い合わせ、アンケートのタイトルなどを取得する。回答者はリストから答えたいアンケートファイルを選択することで、作成者のマシンからアンケートファイルを直接ダウンロードする。

アンケート回答（回答者）

実際にアンケートに答えるメニューが別個存在し、回答はその専用フォームで行う。

回答済データ送信（回答者）

回答者は、回答済アンケートリストから任意のファイルを選択することで、ファイルを作成者側のマシンへ直接送信する。作成者がネットワーク上に存在しない場合（プログラムを起動していない場合）は、作成者がネットワーク上に表れしだい送信される。

アンケート集計（作成者）

作成者のマシンが回答済みアンケートファイルを受信すると、自動的にそのアンケートファイルの集計が行われる。集計ファイルを読み出し、送信されてきたファイルの情報を抽出したのち、そのアンケートの回答結果データ集計ファイルにデータを追加する。

アンケート集計結果表示（作成者）

作成者は、リストから集計したいアンケートを選択するだけで、リアルタイムの集計結果を見ることが可能である。リストからアンケートを選択すると、そのアンケートのグラフが表示される（図3参照）。左側にはグラフの凡例、右側にはグラフを表示する。

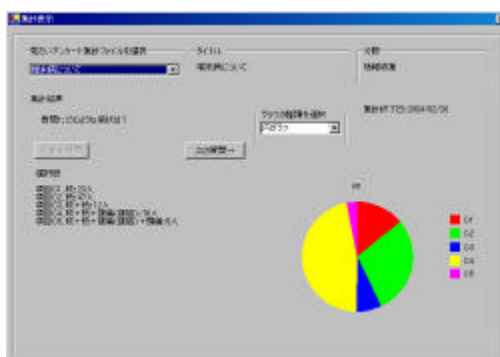


図3 集計結果表示フォーム

5 プロトタイプ評価

動作テストとして大学のコンピュータの性能の良し悪しを問う選択式アンケートを作成し、数人に答えてもらうことで、今回作成したシステムの評価を行った。先に述べたプロセスを経て、作成者は集計結果表示画面において最終的な集計結果を知ることができた。負荷分散の性能評価は、負荷分散であるが故、既存のクライアント・サーバモデルとの相对比较が不可能であるため、実施しなかった。

6 おわりに

本稿の集計システムは、既存の集計専門サイトの、サーバへの負荷の問題の解決を主たる目的とした。

今回のモデルでは、ハイブリッド型 P2P 技術を利用し、サーバにファイルの作成者情報などを保存している。従って、膨大なファイルが P2P ネットワークを流れる場合（膨大なヘッダ情報が存在する場合）、クライアント・サーバモデルと比較すれば小さなものであるが、サーバに負荷がかかってしまう。だが、サーバを介さない通信モデルであれば、不正なアンケートが流通した場合に、作成者の摘発や、流通したデータへの対応が困難になる。

今後の課題として、サーバを介して負荷を極力抑える方法を検討するか、サーバを介さずに不正なアンケートをユーザから通知、もしくはシステムが探知するといったような、セキュアなシステムの構築を検討する予定である。

謝辞

2003 年度企業研修を受け入れていただき、数々の技術的指導を賜った独立行政法人 理化学研究所ゲノム 科学総合研究センタータンパク質構造・機能研究グループ 応用化プロテオミクス研究チーム、有限会社 ネットコンパス、株式会社 ソフテム、株式会社 インテリジェントウェイブの方々にこの場を借りて感謝致します。

参考文献・URI

- [1] 富士総合研究所：“e-リサ+(プラス)システム”
, <http://eresearcher.biz/gamen.html>
- [2] 坂田岳史：“P2P イノベーションのすべて”，日本実業出版社，第 6 章
- [3] 小野束：“電子透かしとコンテンツ保護”，オーム社，2001
- [4] Alertbox：“オンライン・アンケートは簡潔に”
<http://www.usability.gr.jp/alertbox/20040202.html>