

インターネット上のフィードバック情報分散集約方式における
集約周期制御機能の考察

3 G - 8

村田 博史, 小泉 泰則, 長岡 満夫
日本電信電話(株) ソフトウェア研究所

1. はじめに

オープンなネットワーク上の情報提供において、利用者からのフィードバックを得るためのリターンチャンネルの提供、得た情報の集約結果のフィードバックが、主に放送と通信との融合の観点から注目されている。一方、放送においては不特定多数の視聴者が、短時間に特定のサービス、特定のチャンネルに集中するため、サービスの円滑な進行のためには、通信網の輻輳やサーバの処理能力が問題になる。

本稿では、利用者からのフィードバック情報によって構成されるサービスにおける、情報の集約処理における課題およびその対処の方式について述べる。

2. 応答収集サービスの課題

インターネット上で、利用者からの応答を求めるサービスが増えているが、利用者からの応答の集中に際して、以下のような問題が生じることが多い。

- ・[アクセシビリティ] 応答の集中により、ネットワークやサーバマシンへの負荷が大きくなり、輻輳やサーバダウンといった障害が発生する。
- ・[不正投票] 特定の項目に対して集中的・恣意的な投票がなされ、その結果、本来の意見の全体傾向とは異なる結果を得ることになる。

負荷集中の回避策として現在、プロキシサーバによるキャッシュや、ミラーサーバ等によるトラフィックの分散が行われている。これらの方法により、利用者からのアクセス先マシンおよびアクセス経路を分散することで、ネットワーク全体の負荷は軽減される。ところが各サーバに蓄積された一次集計を収集して全体集計を得る際、また全体集計を各サーバにフィードバックする際の負荷については、解決策が得られていない。即ち、

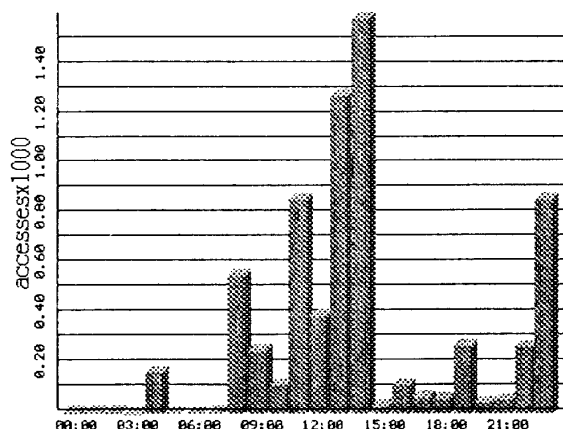


図 1. 投票集中の例

応答情報の集中および結果配送についての負荷軽減が必要である。

ある投票サービスの Web ページにおける、実際の投票集中の傾向を図 1 に例示する。毎日 16:00 受付開始、翌日 15:00 受付終了としたサービスの、ある一日の時間別投票数のグラフである。

この例では、投票受付終了の直前 (14 時台) に投票が殺到している。これは、各利用者が各自の意向を最終結果に反映させようとし、項目間の投票の競争状態が生じたためである。これにより、受付終了の直前に投票数のピークが訪れ、また最終結果は実際の利用者全体の意向に沿わないものになる。

3. 応答情報の収集および結果配送の負荷軽減

集中的に受け付けている投票を収集する際、生ずる負荷を軽減する諸方式について述べる。

負荷を軽減する観点には、以下の方法が挙げられる。

- ・[空間分散] ネットワーク上の複数箇所に利用者を振り分ける方法、
- ・[時間分散] 輻輳を予測し、集計や応答を意図的に遅延させる方法。

空間分散とは、物理的に複数のアクセス先サーバに、利用者を割り振る方法である。割り振る観点としては、ネットワーク上の距離、サーバマシンやネットワークの負荷などがあり、プロキシサ

A Concept of Controlling Interval to Collect and Distribute Feedback-Type Information on the Internet.

Hiroshi MURATA, Yasunori KOIZUMI and Mitsuo NAGAOKA

Software Laboratories, Nippon Telegraph and Telephone Corporation

3-9-11 Midori-cho Musashino Tokyo 180-8585 Japan.

サーバやミラーサーバ、ルータ等により利用者毎に接続先を制御し、同一のサービスを提供する。

時間分散では、各項目の得票数や参照数、マシン負荷、投票集中などに基づき、集約や配送のタイミングを変更させる。時間軸での負荷の観点として、Web ページの参照状況や利用者の応答状況、項目毎の応答集中の状況などが挙げられる。それらにより、集約や配送の周期や時刻を可変にする。

4. 負荷分析

我々は、各利用者からのアクセスを分散させるフィードバック情報分散集約方式 WebGONG を研究している(文献[1], [2])。ここでは、短期に集中するトランザクションを、複数のサーバに振り分け分散させることにより、サーバやネットワークの負荷を軽減し、そのうえで特定の周期で各サーバ毎の集計結果を集約し、全体に反映させる。現在、図2のような空間分散のアーキテクチャに基づきシステムを実装し運用している。

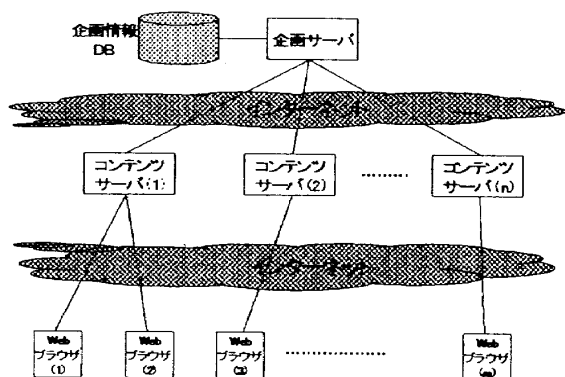


図2. WebGONG アーキテクチャ

ある投票サービスに WebGONG を適用した際の利用者の応答状況を、表1, 2に例示する。15日間の運用の平均値、中央値および最大値である。

平均アクセス	402 件/時	6.7 件/分
中央値	64 件/時	1.1 件/分
最多アクセス	7,058 件/時	117.6 件/分

表1. アクセスの集中

全項目の平均	483 件/日
中央値	68 件/日
最大得票項目	23,299 件/日

表2. 投票の偏り

なお表2において、得票が最大(23,299件)だった項目について、そのうち2万件以上が、僅か2つのホストマシンからのアクセスであった。

5. 考察

前述の2つの表から判明することとして、以下のことが挙げられる。

- アクセスの集中時には、通常の100倍以上にまで集中的にアクセスがある。
- 特定の項目に、わずかの利用者からのアクセスが集中し、その項目だけ得票数が突出する。
- 集中的なアクセスがあるのは、ある特定の時間帯に限られる。

これらにより生ずる負荷を回避するため、集中的なアクセスのある特定の時間帯においては、一旦、一次集約先に応答情報を貯えておき、集中のピークを回避した後に集約を行うのが望ましい。これにより、全体集約における負荷を平準化できる。図3にその概念を記す。ここでXは、集約におけるサーバの処理能力を示す。

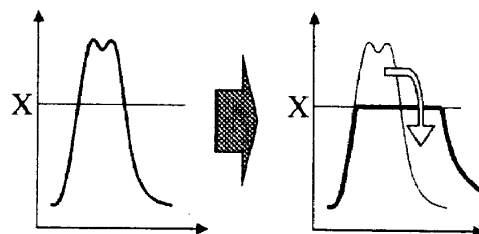


図3. 集約分散のイメージ

即ち、集約処理においては、実際の応答における集中のピーク自体を、時間軸の方向に和らげ、それ以外の時間帯に割り振ることで、各サーバの処理能力を有効に活用できる。

また同時に、集約を遅延させることにより、応答の反映が遅れ、項目間の投票の競争による偏りが生じにくくなることが見込まれる。

6. おわりに

本稿では、利用者からのフィードバック情報によって構成するサービスにおいて、集中アクセスを分散する方式について述べ、また時間分散の導入効果について考察を行った。今後、時間分散を実際のシステムに適用し、効果を検証する。

参考文献

- [1] 小泉, 吉田, 長岡 “インターネットにおけるフィードバック型情報集約方式の構想”, 情報処理学会マルチメディアと分散処理研究会 84-6, pp.31-36, 1997.
- [2] 小泉, 吉田, 長岡 “インターネット活用FB型情報集約技術の評価 -サービス適用とその性能評価-”, 情報処理学会マルチメディアと分散処理研究会 86-25, pp.141-146, 1998.