

データ集約システムにおける制御手法の考察

渡部智樹

日本電信電話株式会社 NTT サイバーソリューション研究所

視聴者参加型のインタラクティブ番組は、近年開始されているデジタル放送だけでなく、インターネットにおいてもますます盛んになってきている。このような番組では、ユーザデータの迅速かつ確実な集約が必要不可欠である。これまで筆者らは回線交換網におけるデータ集約システムを開発しその有効性を確認してきたが、今回インターネット向けにシステムを拡張した。本稿では、本システムの動作概要を述べ、インターネット向けの適用方法を提案し、シミュレーション実験の結果から考察する。

A Study of Control Method in Data Collection System

Tomoki WATANABE

NTT Cyber Solutions Laboratories, NTT Corporation

Recently, many internet users can enjoy interactive events. It is necessary to collect their data quickly and certainly in the event. We developed a data collection system that can collect efficiently using public telephone network, and checked the validity. This article describes an application to the Internet of this system, and considerations of the experiment.

1 はじめに

1.1 回線交換網向けのデータ集約

2000年末のBSデジタル放送, 2002年の110度CSデジタル放送に続いて, 本年2003年末には地上デジタル放送が三大広域圏で開始される。このように, 放送のデジタル化が急速に進む中で, 視聴者とのインタラクティブ性を盛り込んだ視聴者参加型番組が多く制作されている。このような番組では, 視聴者の意見や回答などのユーザデータを即座に収集し, 番組へ反映させるた

めの集約システムが必要となる。

現在のデジタル放送受信チューナにはモデムが搭載されており, これらのユーザデータはモデム通信により放送局へ送信される。このモデム通信では, 回線交換網を使用するため, 放送局に集中する呼の輻輳が懸念されていた。例えばクイズ番組で, ユーザが“回答ボタン”を押下した時点でデータを送信する場合, 他のユーザもほぼ同じタイミングで送信すると考えられるため呼が集中してしまうことになる。

1.2 インターネットにおけるインタラクション

このような番組へのインタラクションはインターネットでも行われており、近年の常時接続型インターネットの普及により、クイズ番組だけでなく、TVショッピングや生番組のアンケートなど、広く使われるようになってきた。

広帯域のネットワークであれば、これらのインタラクションに対するネットワーク負荷はそれほど大きくないと思われるが、今後画像や映像などデータ量が大きくなったりユーザ数が大多数になったりすると、サーバ側で能力の限界を超えてしまう可能性がある。

1.3 インターネットにおけるアクセス制御

デジタル放送における輻輳の問題に対して、筆者らは回線交換網を使ったデータ集約システム「TeleCollection」を開発してきた^{[1][2]}。このシステムは、チューナの発呼動作を放送局側から制御し、時間的に分散させることにより、輻輳を回避し効率良くユーザデータを集約する(図1参照)。受付サーバでの集約状況に応じたフィードバック制御を行うため、チューナ数を予測できない状況や視聴者数の変動が大き

い状況でもリアルタイムにフィードバックを行い集約効率を維持することができる。インターネットの場合においても、端末からのアクセス数を正確に予測することは困難であり、TeleCollectionによる制御は有効であると思われる。

そこで、回線交換網を使ったTeleCollectionのアルゴリズムをインターネット向けに適用し、デジタル放送の場合と同様のアクセス制御を実現することとした。

2 データ集約システムの概要

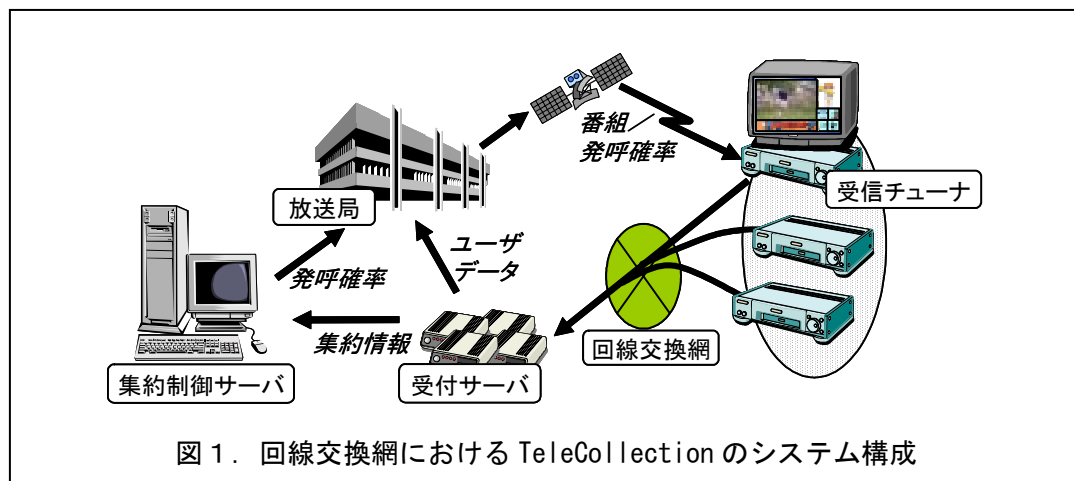
回線交換網におけるデータ集約システム「TeleCollection」の構成を図1に示す。また、ここで使用する言葉を以下の通り定義する。

・集約対象端末 (集約対象数)

クイズなどの番組に参加し、回答など放送局に送信すべきユーザデータを保持し、未だ送信していない端末(数)

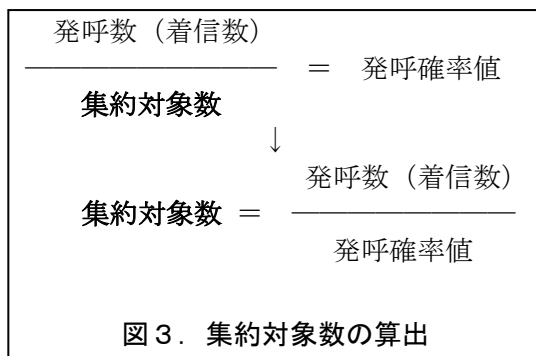
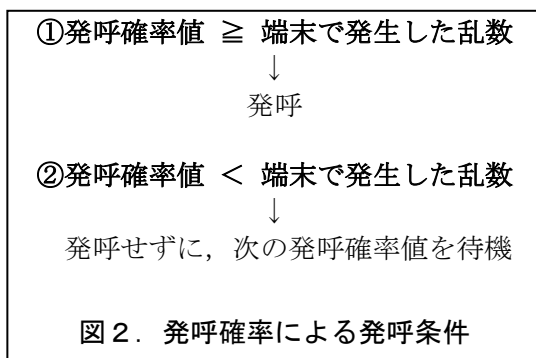
・窓口数

端末からの接続に応答するために放送局側(受付サーバ)に用意されたモ



デムの数. ある瞬間にこれ以上の発呼があると呼損となる.

TeleCollection では, チューナの発呼制御情報として発呼確率値を放送する. これを受信した集約対象端末は各々で乱数を発生させ, この乱数が発呼確率値以下であれば発呼し (図 2 の①), そうでなければ次の発呼確率値の受信を待機する (図 2 の②). 放送局側では, この発呼確率値



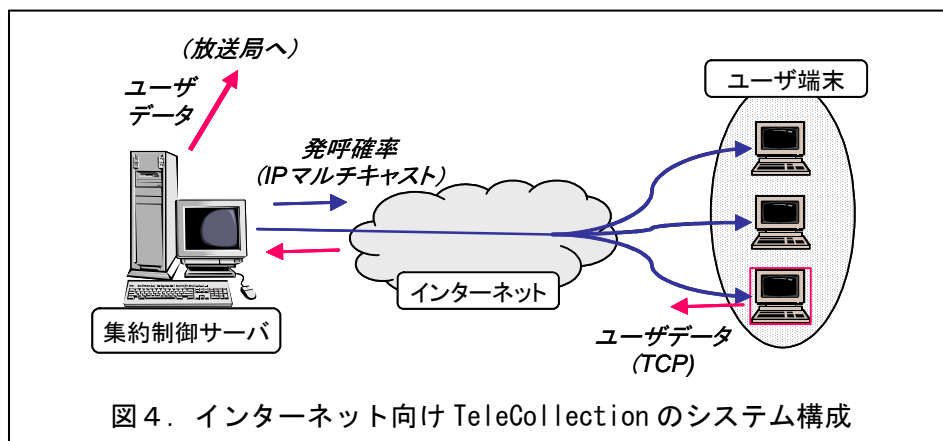
により着信した呼数 (着信数) を集計し, この時の集約対象数を推定する. つまり, 集約対象数に対して発呼確率の相当分が発呼し, センタに着信したと見なせる. よって, 発呼数が着信数と同値であるとする, その時の発呼確率から集約対象数を推定できる (図 3). そして, この発呼確率で発呼しなかった残りの集約対象数と窓口数から次の発呼確率値を設定し放送する. このようなフィードバックによる集約を集約対象がなくなるまで繰り返すことで, 全ての集約対象端末のユーザデータを効率的に集約することができる.

3 インターネット向けの集約手法

回線交換網を使った TeleCollection のアルゴリズムをインターネット向けに適用するにあたり, 以下の変更を行った.

3.1 システム構成の変更

インターネット向けに変更した TeleCollection のシステム構成を図 4 に示す. 端末とサーバは常時接続によりインターネットで接続されているものとする. また, 発呼確率は IP マルチキャストで同報する.



3.2 制御手法の高度化

TeleCollection では、ユーザ端末の発信を制御する発呼確率値の決定が最も重要な処理であり、そのためには集約対象数を正確に予測する必要がある。

以前行った集約対象数の変化予測に関する実験結果から、直近の集約数に対して“平均値”を用いるよりも、“一次近似”による予測の方が、呼損（送信ロス）が少なく高精度に制御できることを導き出した^[3]。

インターネットでは集約対象数の変動がより一層大きいと思われるため、新たに以下の2つの手法を導入した。

・二次近似手法の適用

集約対象数の予測方法として、回線交換網では一次近似を用いていたが、より精度が高いと思われる“二次近似”を新たに適用する。

・減衰率の動的変更による制御

回線交換網の場合、輻輳の回避が必須課題であったため窓口数のある一定の割合（閾値）を超えると必然的に発呼確率値を減衰させていた。ここでは集約対象数の増減率に合わせて動的に変動させる手法を導入した。

3.3 発呼確率の配信周期の変更

発呼確率の配信周期とは発呼確率を配信する間隔、すなわち集約対象端末に制御を与える間隔である。この値が小さければ端末の緻密な制御が可能になる。

4 シミュレーション実験

4.1 実験の概要

3章で述べたインターネット向けのTeleCollectionにおいて、全ての集約対象端末から集約するための時間（集約時間）と、端末からの送信に失敗した数（送信失敗数）の観点から評価した。

実験環境として、TeleCollectionのクライアントモジュールを同時に400台動作させ、各モジュールが保持する1KByteのデータを100MbpsのLANを介してサーバに集約するシミュレーションシステムを構築した。また、サーバ側では処理負荷の増大を防ぐため、同時に接続可能なセッション数を窓口数として設定し、窓口数を超えるアクセスは送信失敗として再度送信させることとした。

4.2 予測手法の適用

3.2項に述べた制御手法を組み合わせ、表1の4通りを適用した。

表 1. 実験に適用した制御手法

	予測手法	減衰率
制御 A	一次近似	固定
制御 B	一次近似	動的変更
制御 C	二次近似	固定
制御 D	二次近似	動的変更

4.3 発呼確率の配信周期

2秒、5秒、10秒の3種類を適用した。

4.4 集約対象数の増減パターン

イベントに参加しデータ送信を待機している集約対象数について、以下の3通りのパターンを適用した。

・集約対象パターン A

一度に 400 台全てが集約対象端末となる。

・集約対象パターン B

400 台の集約対象端末が 3 回データ送信を行う。

・集約対象パターン C

最初は集約対象数 400 台で開始し、徐々に脱退していき集約対象端末が減少していく。

5 結果と考察

5.1 集約対象パターン A の結果

結果を図 5 に示す。

送信失敗が現れたのは一次近似だけであった。一方、二次近似の集約時間は一次近似の場合の約 2 倍となった。また、減衰率は固定された方が短時間で集約できる傾向にあった。二次近似では送信失敗とならないように制御が強く働き、減衰率の動的変更を適用するとその傾向はさらに強くなることが分かった。

配信周期については、10 秒の場合は 2 秒、5 秒のときと比べて約 2 倍の集約時間がかかっている。つまり、失敗せずに短時間で集約できるにもかかわらず、ネットワークやサーバのリソースを十分に活かせていない状態が続いていたと言える。

5.2 集約対象パターン B の結果

結果を図 6 に示す。

一次近似では、パターン A に比べて全体的に送信失敗が多く発生した。このパターンでは、各端末の送信が 3 回になり、全体の送信回数が増えたため、一次近似では集約対象数の変化に対応できず送信失敗が発生してしまったと考えられる。

しかし、二次近似ではほとんど発生しておらず、ほぼ正確に対応できたと言える。このことは、配信周期 2 秒の一次近似・動的に比べて、二次近似・動的では集約時間が多少長くても失敗数が全くないことから分かる。

5.3 集約対象 C パターンの結果

結果を図 7 に示す。

このパターンでの集約時間はほぼ同じであった。また、配信周期が 2 秒、5 秒の一次近似では送信失敗が多数発生しているが、5 秒、10 秒の二次近似では全く発生していない。

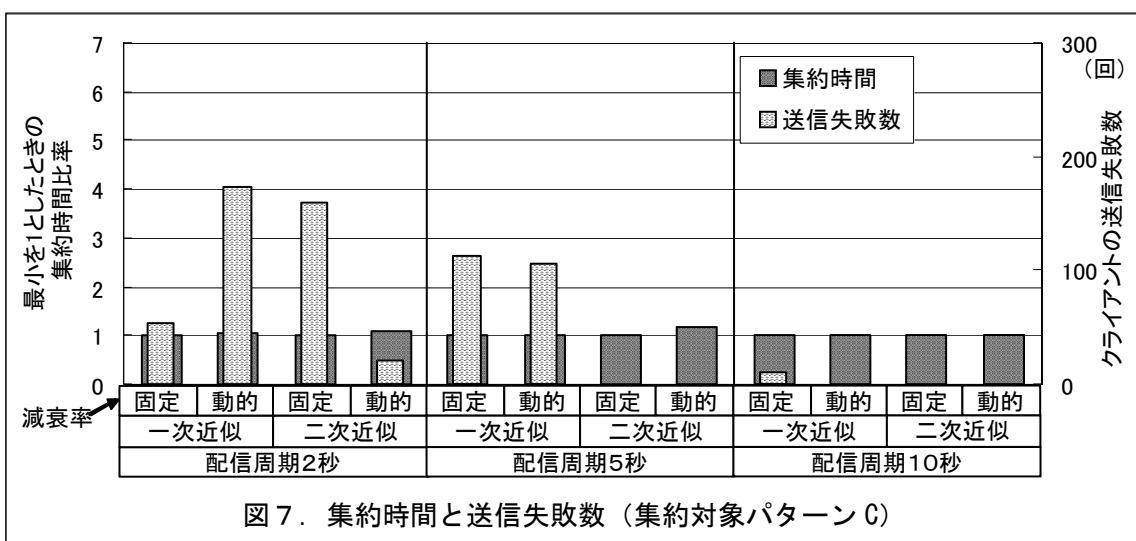
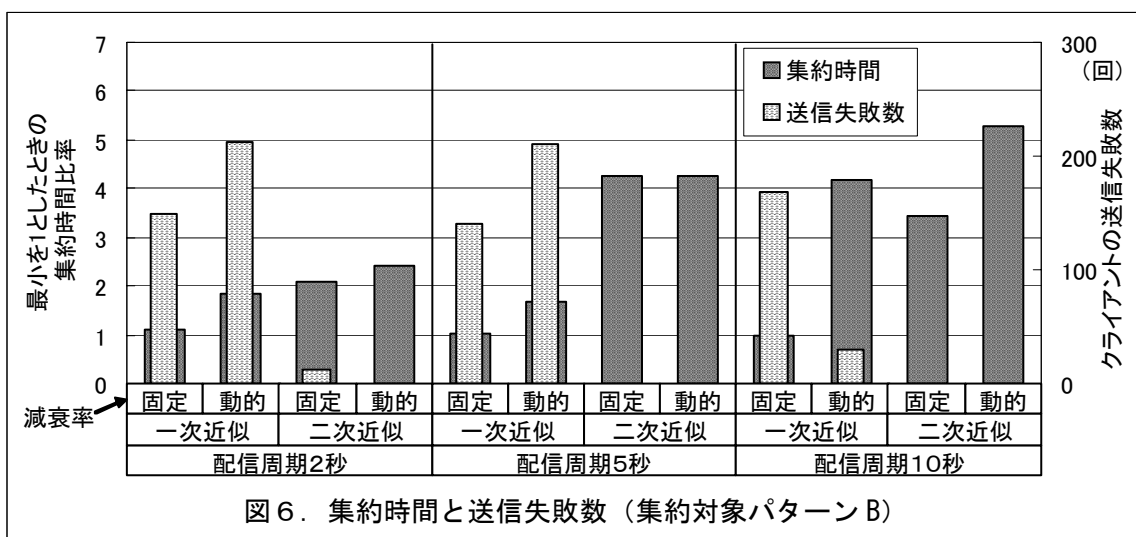
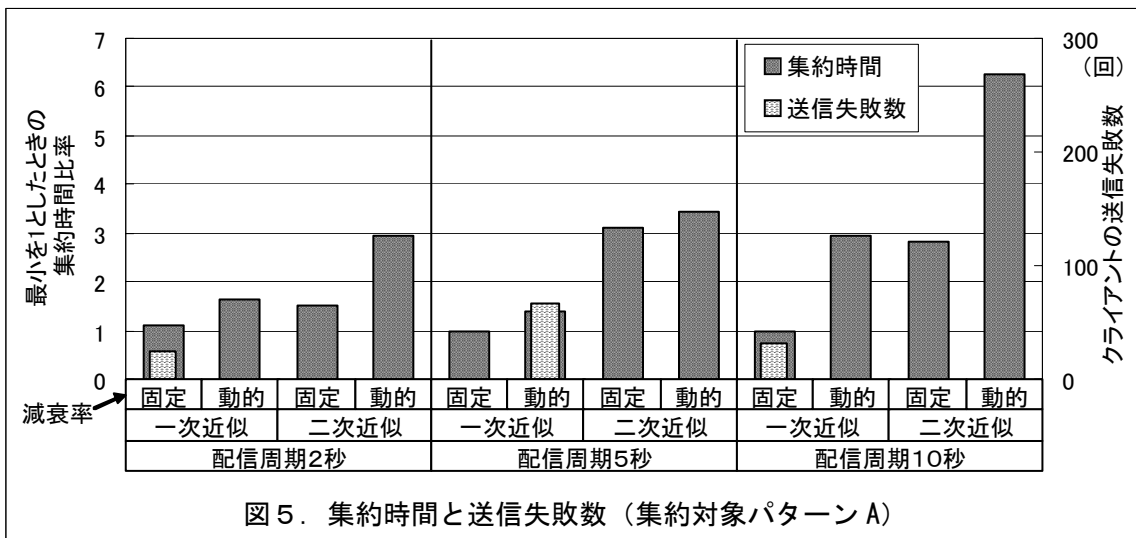
配信周期が短くなると一回の発呼確率での着信数は少なくなり、さらに集約対象数が減少していくこのパターンでは、この着信数はさらに小さくなる。このように着信数が小さくなると、1 台の端末が送信するか否かの判定の重みが大きくなり、集約対象数の予測精度に大きく影響する。各端末で発生させる乱数の精度にも起因するが、着信数が小さい時の対処が必要である。

5.4 考察のまとめ

二次近似による予測は、一次近似よりも送信失敗数が少なく有効である。

減衰率は失敗数の増減に関わらず集約時間が長くなり、固定の場合よりも優位な傾向は見られなかった。減衰率の決定方法を改善する必要がある。

配信周期については、短く設定する方がきめ細かい制御が可能になるが、1 回の発呼確率での着信数が少なくなり、予測の精度が低下する。着信数が少ないときの予測手法を改善する必要がある。



6 今後の課題

集約対象数をさらに増やし、また様々な増減パターンにおいて再度実験を行い、考察を深める。また、制御手法としては、発呼確率の配信周期を動的に変更する方法や、着信数が少ないときの予測手法について検討を行う。

7 おわりに

本稿では、インターネットにおけるデータ集約手法として、筆者らが回線交換網向けに開発してきたデータ集約システム TeleCollection をインターネット向けに適用したシステムを提案し、シミュレーション実験から得られた結果について考察した。

[参考文献]

- [1] 酒井 他, 放送と通信の結合サービスにおけるデータ集約技術, 信学会 96 秋季大会 B-7, 1996
- [2] 酒井 他, TeleCollection 大規模データ集約システム, 情処学会第 55 回全国大会, 2V-4, 1997
- [3] 渡部 他, インタラクティブ番組におけるデータ集約制御手法, 情処 DPS 研究報告 No.107, 2001