

## プッシュ型データ配信における 遅延軽減のためのデータ選択方式

青野正宏<sup>†</sup> 上田尚純<sup>†</sup> 田窪昭夫<sup>†</sup>  
渡辺尚<sup>††</sup> 水野忠則<sup>††</sup>

データ放送において、放送するデータがサーバで更新されたとき、できる限り早く放送してクライアントが最新のデータを得られるようにすることが望まれる。更新データが複数発生した場合、サーバはいずれのデータから出力するかを選択しなければならない。サーバにおけるデータの更新から放送出力までの時間を遅延時間と定義し、クライアントによる情報の受信率を考慮して平均遅延時間を軽減する手法について提案する。さらに出力データ選択のための演算負荷を軽減する手法を提案し、評価する。

### A method of scheduling for minimum latency time in push-based communication

Masahiro Aono<sup>†</sup> Takasumi Ueda<sup>†</sup> Akio Takubo<sup>†</sup>  
Takashi Watanabe<sup>††</sup> and Tadanori Mizuno<sup>††</sup>

When broadcast data are updated at a server, it is desirable that the server broadcasts it as soon as possible and clients get fresh information in data broadcast system. When there are many pages to be updated at the same time, the server must chose a page to output from the candidacies. We define that delay time is the time when a page is output after it was updated at the server. Then, we propose a method to reduce average delay time using the access rate of clients to the pages. Then we propose some methods to reduce the runtime for the scheduling without increasing delay time almost.

#### 1. はじめに

インターネットの発展に伴い、情報収集の手段として、その利用は極めて重要な手段のひとつとなってきた。その利用形態もクライアントが情報を蓄積しているサーバのデータベースにアクセスして情報を引き出すプル型通信のみならず、クライアントが自動的に一定時間ごとに情報を採って収集する擬似プッシュ型通信の形態の利用も進んでいる。今後は、サーバから直接予め定められたクライアントに情報を配信するリアルプッシュ型通信の発展が見込まれる。また、不特定ユーザを対象とする無線データ放送もリアルプッシュ型通信のひとつである。リアルプッシュ型のデータ配信の利点はマル

チキャストやブロードキャスト通信の利用により、通信帯域とサーバの負荷が軽減できることと共に、サーバ側が把握している最新情報をクライアントが能動的にアクションを起こさなくても、主体的に配信することができる点である。サーバの情報が更新されると、サーバができる限り早くその情報をできるだけ早く配信することにより、クライアントは新鮮な情報を得ることができる。出力すべき情報が複数存在すれば、どの候補から出力すべきか判断しなければならない。本論文では、複数の出力候補が存在した場合、サーバから全クライアント平均として、どのように情報を選択して出力すべきかを判断するための、出力データ選択スケジュール方式を考察し、提案評価することを目的としている。また、簡易的に演算負荷を抑えられるスケジュール方式についても検討を行う。

以下、2章で本論文に用いる用語や前提条件の定義を行い、3章で遅延時間軽減のためのスケジュール方式に

<sup>†</sup>三菱電機

Mitsubishi Electric Corporation

<sup>††</sup>静岡大学

Shizuoka University

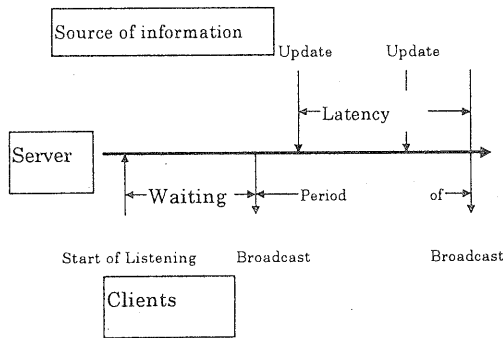


図 1. 待ち時間と遅延時間の概念  
Fig.1 Waiting time and Latency time

ついて提案を行い，4章でその評価を実施する．さらに，5章で先行研究との関連について述べ，6章で全体のまとめとする．

## 2. 用語と前提条件の定義

本論文では検討にあたり，以下の用語と前提条件の定義を行うとする．放送と受信の形態は次の状況を想定する．多数のクライアントが，ひとつのサーバから放送されている情報を受信しているものとする．クライアントは放送される情報のなかから，自己が必要とする情報のみを選んで利用する．クライアントは任意の時点で受信を開始する．クライアントは目的データを受信すると，以後，そのコンテンツの更新を把握するため受信を続け，ある時点で受信を中止する．サーバは，前回にある情報を出力後，その情報の更新データが最初にサーバに入力されてから，最新の更新データが出力されるまでの時間を遅延時間とする．クライアントは自分が興味ある情報を受信しているものとする．出力されるページは時々更新される．このとき，サーバは更新された情報をできるだけ早く出力すれば，クライアントは早く情報を得られることになる．ある情報に関する更新データが出力されないうちに，その情報の新しい更新データがサーバに入力されれば，サーバは古い更新データを捨てる．しかし，評価としては，古い更新データをクライアントに知らせることができなかったので，評価としては上述の定義とする．図 1 に待ち時間と遅延時間の概念を示す．なお，放送の帯域は一定とする．その他の論文全体に共通する用語と前提は以下のとおりである．

- (a) アクセス：クライアントが放送される情報を受信して内容の判断まで行うこととする．
- (b) ページ：サーバの出力，クライアントのアクセスがなされる情報の単位であり，全て同じ大きさとみな

す．放送されるページの新規生成と消滅はこの論文の範囲内では考えない．

- (c) コンテンツ：ページの内容である．コンテンツは時々更新される．コンテンツ更新やアクセスの単位はページとする．ページのコンテンツの更新があった場合もページの大きさは変わらないものとする．
- (d) タイムスロット：1 ページをサーバが出力する時間である．以下の図 2～9 における遅延時間の単位はタイムスロットである．
- (e) 遅延時間：前回にあるページを出力後，そのページの更新データが最初にサーバに入力されてから，最新の更新データが出力されるまでの時間である．クライアントは自分が興味あるページにアクセスしているものとする．出力されるページは時々更新される．このとき，サーバは更新されたページをできるだけ早く出力すれば，クライアントは早く情報を得られることになる．あるページに関する更新コンテンツが出力されないうちに，そのページの新しい更新コンテンツがサーバに入力されれば，サーバは古い更新コンテンツを捨てる．しかし，評価としては古い更新データをクライアントに知らせることができなかったので，評価としては上述の定義とする．図 1 に待ち時間と遅延時間の概念を示す．
- (f)  $i$ ：各々のページの番号を示す．
- (g)  $m$ ：出力される全ファイルの総ページ数を示す． $i=1 \sim m$ である．
- (h)  $p_i$ ：定常要求率．クライアント総数に対し，ページ  $i$  に継続してアクセスしようとしているクライアントの比率を示す． $0 \leq p_i \leq 1$ である．
- (i)  $T_i$ ：ある時点において，ページ  $i$  の前回出力後に最初のコンテンツ更新が発生してからその時点までの経過時間（遅延時間）を示す．
- (j)  $q$ ：1 タイムスロットあたりコンテンツが更新される全ページの平均確率を示す．
- (k)  $q_i$ ：ページ  $i$  のコンテンツが 1 タイムスロットあたりに更新される確率を示す．以下，コンテンツの更新率は単に更新率と略する．

## 3. 遅延時間軽減のためのページ選択方式

コンテンツが更新されたページが放送されず累積されていた場合や同時に複数のページのコンテンツが更新された場合，サーバから出力すべきページが複数存在することがある．そのとき，どのページを選択出力すれば，全体として最も遅延が小さくなるかを検討するのが本章の目的である．そのため，ページの選択方式を検討する．

### (1)基本方式 (TP 法)

まず、いくつかの基本的なページ選択方式を挙げる。第1は、コンテンツの更新時刻が古いページ順に出力すること、つまり遅延時間が最大のページを選択する方式である(先着順方式)。第2は定常要求率が最大のページを選択する方式である(要求率順方式)。第3にそれらを折衷する方法である。折衷の方法は、次のように考える。遅延時間と定常要求率の両者の要素を取り入れて選択基準値を算出し、その値が最大の値を持つページを選ぶものとする。折衷係数を  $k$  ( $0 \leq k \leq 1$ ) とするとページ  $i$  の選択基準値  $SD_i$  は、 $T_i$ ,  $P_i$ ,  $k$  の3つの要素からなる関数として表わせる。

$$SD_i = f(T_i, P_i, k) \dots \dots \dots (1)$$

簡易な計算で遅延時間と定常要求率の両者の要素を取り込もうとすると、加算は両者のデータの性質が全く異なるため不適当であり、積算で選択基準値を選ぶ方式の方が適している。 $\alpha$ の値により、 $SD_i$ において相対的に  $P_i$ の要素を無視して  $T_i$ の要素だけで決まる場合(先着順方式)から、その逆の場合(要求率順方式)までを簡易な形で表現する関数として、 $T_i$ ,  $P_i$ をそれぞれ折衷係数  $k$ ,  $1-k$  でべき乗変換する方法を考察する。式(1)を、次のように改める。

$$SD_i = T_i^k \cdot P_i^{1-k} \dots \dots \dots (2)$$

$0 \leq k \leq 1$  とする。ここで  $k=1$  とおくと、 $P_i$ の要素は無視して、 $T_i$ の要素だけで決定する先着順方式となり、 $k=0$  とおくと、 $T_i$ の要素は無視して、 $P_i$ の要素だけで決める要求率順方式となるので、折衷方式は先着順方式と要求率順方式を包含することとなる。ゆえに、以下は先着順方式と要求率順方式を省略し、それらは折衷方式の両方の極限を示すものとして捉える。ところで、選択基準値は絶対的な値に意味はなく、比較のための相対値に過ぎない。 $T_i$ と  $P_i$ の両方の要素をべき乗計算するのは計算負荷上得策でなく、いずれかをそのままとした値と他方をべき乗変換した値との積を選択基準値とすれば計算は楽になる。 $T_i$ は計算する度に値は必ず変わるが、 $P_i$ はリアルタイムに測定できないため一定と仮定するか、または変わるとしてもその頻度は遅延時間と同等か少ないと想定して良い。計算量削減の観点から  $T_i$ より  $P_i$ をべき乗変換の方が望ましいと判断できる。 $k>0$ のとき、

$$\alpha = (1-k)/k \dots \dots \dots (3)$$

とおき、置換評価係数  $CSD_i$ を定義する。

$$CSD_i = SD_i^{1/k} = T_i \cdot P_i^\alpha \dots \dots \dots (4)$$

この折衷方式において、全てのページの  $T_i$ と  $P_i^\alpha$ との積を計算し、その最大値を求め、該当するページを出力データとして選択する方式をTP法と呼ぶこととする。

### (2)選択評価対象打ち切り方式 (RxW 法)

TP法の計算において、 $P_i$ が変わらないという前提をおき、 $P_i^\alpha$ の値は変化しないとしても、 $T_i$ については毎回値が変わる。ページ数が多い場合、TP法は計算負荷が大きい。

このため、TP法を簡略化した計算方式を考察する。Aksoyらは、オンデマンド型通信における出力ページ簡易選択法として、RxW法を提案し、評価を行っている<sup>2)</sup>。 $R_i$ をページ  $i$ の要求クライアント数、 $W_i$ を最初にクライアントが要求開始してからの時間を意味するものとする。  $R_i$ が多い順から一定数ページと  $W_i$ が大きい順から一定数ページを選び、それぞれについて  $R_i$ と  $W_i$ との積を計算し、そのなかで最大の値を持つページを選択する方式がRxW法である。我々はこの手法を更新における遅延軽減のための出力ページ選択法に応用してみる。 $R_i$ の代りに  $T_i$ を用い、 $W_i$ の代りに  $P_i^\alpha$ を用いて  $T_i \times P_i^\alpha$ が最大のページを探すものとする。

### (3)グループ化方式 (GP 法)

もうひとつのTP法の計算簡略化手法として定常要求率が近似したページを集めてグループを作り、グループ単位で比較計算を行う方式を提案する。各グループのページの合計出力配分率なるべく等しくなるようグループ分けを行う。その方法は次のとおりである。各ページ番号  $i$  は定常要求率の大きさにより降順に振られているものとする。各ページの出力配分率を  $s_i$  とすると、 $P_i$ と  $\alpha$ により、以下の式で各ページの出力配分率を定める。

$$s_i = \frac{P_i^\alpha}{\sum_{j=1}^n P_j^\alpha} \dots \dots \dots (5)$$

$numgrp$  はグループの数とする。 $gid_{i,numgrp}$  は各グループの先頭ファイルを指すものである。 $i, j, sum$  は計算のための一時的な使用値である。

- step1:  $gid_j=1, j=2, sum=0, i=1$  とする。
- step2:  $j$  が  $numgrp$  を越えなければ step3~6 を実行し step2に戻る。
- step3:  $sum$  に  $s_i$  を加算する。
- step4:  $sum$  が  $(j-1)/numgrp$  より小であれば step5を省略する。
- step5:  $gid_j$  を  $i+1$  とする。 $j$  をカウントアップする。
- step6:  $i$  をカウントアップする。

オフライン処理で以上のグループ分けを行った後、オンライン時に、次の方法で出力するページを選ぶ。まず、グループ別に更新ページの先着順に待ち行列を作り、各々のグループにおける先頭順位のページの遅延時間

を  $T_i$  とする。各先頭ページの中で  $T_i \times p_i^\alpha$  が最大のページを選択する。この方法は定常要求率が近似するページをまとめ、グループ内では先着順方式を採用することにより、負荷の軽減を図るものである。以下、GP 法と呼ぶ。

#### (4) 基準時刻設定方式 (ST 法)

別の演算負荷軽減方法として、各ページに基準時刻を設定しそのもっとも早い時刻を選択する方式を提案する。これは、最初にそのページの更新率と定常要求率から各ページの望ましい放送周期を求める。放送周期を求める方法はべき乗指数法<sup>14)</sup>を用いる。その手法は、以下のとおりである。各ページについて、パラメータ  $\alpha$  と更新率をべき乗変換するパラメータ  $\beta$  の値を仮に想定し、以下の式で周期  $d_i$  を求める。

$$d_i = \left[ \frac{\sum_{j=1}^m p_j^\alpha \cdot q_j^\beta}{p_i^\alpha \cdot q_i^\beta} + 0.5 \right] \dots \dots \dots (6)$$

周期を仮に設定するとページ  $i$  の期待遅延時間 LATENCY <sub>$i$</sub>  は、各タイムスロットにおける遅延時間に新規コンテンツ更新の発生確率を乗じた値の総和に周期内におけるコンテンツ更新確率で除すと得られる。

$$\text{LATENCY}_i = \frac{\sum_{k=0}^{d_i-1} ((d_i - k) \cdot q_i \cdot (1 - q_i)^k)}{1 - (1 - q_i)^{d_i-1}} \dots \dots \dots (7)$$

$$d_{i+1} = \frac{1 - (1 - q_i)^{d_i+1}}{q_i}$$

$$= \frac{1 - (1 - q_i)^{d_i+1}}{1 - (1 - q_i)^{d_i-1}}$$

これから、全体の遅延時間

$$\text{LATENCY} = \sum_{i=1}^n \text{LATENCY}_i \cdot p_i \dots \dots \dots (8)$$

を最小にする  $\alpha$ ,  $\beta$  を  $\alpha$ ,  $\beta$  の値をいろいろ変えて試算し、探索法により求める。近似的に遅延を最小とする  $\alpha$ ,  $\beta$  が定まれば、 $d_i$  が定まる。ここまでは、事前のオフラインの計算、又は想定更新率が変わったときのみに行う計算である。

次に当該ページを放送出力する基準時刻  $T_s$  を計算する。基準時刻を  $T_s$  とすると新たにページの更新が発生したとき、以下の式で求める。

$$T_s = T_s + d_i \cdot \frac{d_i + 1 - (1 - q_i)^{d_i+1}}{q_i} - 1 + T_s \dots \dots \dots (9)$$

$$= T_s + d_i \cdot \frac{1 - (1 - q_i)^{d_i+1}}{q_i} + T_s$$

$T_s$  は基準時刻計算のときの時刻、すなわち当該ページのコンテンツが更新されたときの時刻である。更新があった時刻に固定スケジュール方式における平均遅延時間を加えて 1 を減ずる。なお、式(7)の場合は、当該ページの放送出力間に更新が発生しなかった場合を遅延時間の計算から除くため、当該区間の更新率で除しているが、式(9)の場合は更新が発生した時点で計算を行うので更

新率で除する必要がないのが異なる。1 を減ずる理由は、最初に基準時刻を計算する可能性のある時刻は出力後 1 タイムスロット後になってからのためである。その値に調整値  $T_s$  を加える。 $T_s$  の初期値はゼロとし、当該ページが実際に出力された時刻 ( $T_{out}$ ) と基準時刻との差異を調整用に保存するものである。 $(T_s = T_s - T_{out})$  当該ページが基準時刻より早く出力されれば次の基準時刻設定では遅くなるよう、遅く出力されれば早くなるよう調整するのに用いられる。この方式は平均して 1 タイムスロットに 1 回以下の基準時刻設定のみで済む。基準時刻が最も小さいページを放送出力ページとして選択する。これを ST 法と呼ぶこととする。ただし、以下の提案方式の評価で、個々のページの更新率は事前に把握できないものとして、実際には更新率に差異があっても、式(6)において  $\beta=0$  とし、個別の更新率の要素は基準時刻設定上不明とした場合の効果について試算するものとする。

## 4. 提案方式の評価

### (1) 各方式比較

各方式において、 $T_i$  の要素と  $p_i$  の要素の比重をどの程度にすれば良いか、すなわち  $\alpha$  の値をどう設定すれば、遅延時間を低く抑えることができるか、また、TP 法、RxW 法、GP 法、ST 法の各手法を採用するといかなる遅延値となるかを、シミュレーションにより検討する。シミュレーションに用いる  $p_i$  の分布は Acharya<sup>9)</sup>, Vaidya<sup>9)</sup>, Wong<sup>10)</sup> などの研究でも使用されている式(10)で示す Zipf 分布に従うものとする。

$$p_i = \frac{(1/i)^\theta}{\sum_{j=1}^n (1/j)^\theta} \dots \dots \dots (10)$$

この分布の式は、 $\theta$  をべき乗値として要求率と更新率各々に適用することにより分布の傾きを定量的に表すことができる。 $\theta$  の値が大きくなるほど、分布の傾きは大きくなる。本論文では  $\theta=1$  とし、 $\alpha$  の値を 0 から 2 まで変化させて遅延が最小となる  $\alpha$  を求める。各ページの 1 タイムスロットあたり平均更新率 ( $q$ ) が等しく、その値を  $2^{-1}$  から  $2^{-9}$  とした場合の各  $\alpha$  値による平均遅延時間を図 2 (TP 法)、図 3 (RxW 法)、図 4 (GP 法)、図 5 (ST 法) に示す。なお、ページ数は  $2^8=256$  ページ、RxW 法における打ち切り数は  $T_i$ ,  $p_i^\alpha$  をそれぞれ上位から 8 ページ、GP 法におけるグループ数は 5 を例として用いている。いずれの手法においても、 $q$  が極めて大きい場合は、 $\alpha=0.5$  が最も遅延時間が少ない結果となっている。その理由は容易に証明できる。ページの更新率が非常に大きい場合、すなわち各タイムスロットにおいて常時コンテンツが更新されている場合を想定すると、ページの

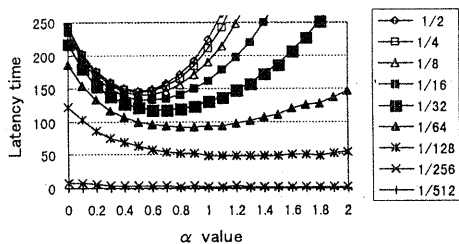


図2 TP法による遅延時間  
Fig.2 Latency time with TP method

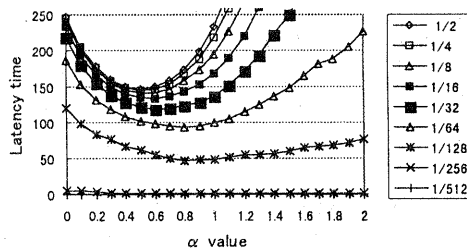


図6 最小遅延時間  
Fig. 6 Minimum Latency time

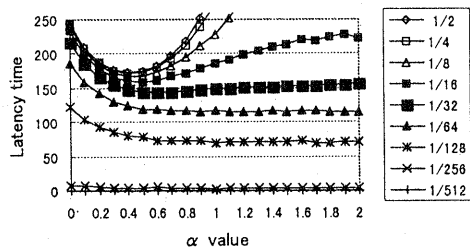


図3 RxW法による遅延時間  
Fig.3 Latency time with RxW method

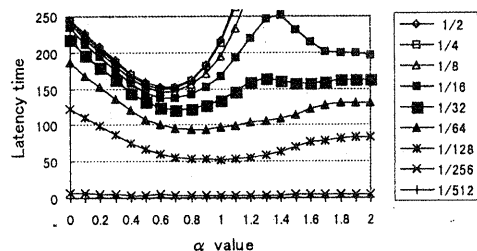


図4 GP法による遅延時間  
Fig.4 Latency time with GP method

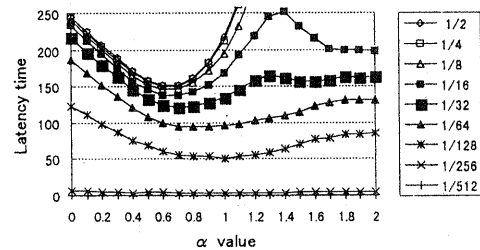


図5 ST法による遅延時間  
Fig.5 Latency time with ST method

表1 スケジューリング方式

Table 1 Scheduling Method

手法	方式	遅延時間	演算時間	備考
TP法	受信率のべき乗変換値×経過時間の最大値を選択.	もっとも小さい.	もっとも大きい.	基本的な方式
RxW法	TP法の簡易型. 評価対象を限定.	大きい.	大きい.	評価対象を増やせば遅延時間が減少し, 演算時間が増大する.
GP法	TP法の簡易型. 類似受信率をグループ化して比較.	比較的小さい.	小さい.	グループ数を増やせば遅延時間が減少し, 演算時間が増大する.
ST法	出力基準時刻を設定しその大小で選択.	小さい.	比較的小さい.	更新率が判明していることが望ましい.

出力直後に更新が発生するので, ページの出力間隔と遅延時間はほぼ等しくなる. ページ  $i$  の送信間隔を  $d_i$  とする. また,  $s=1/d_i$  とおく. 以下の式が成立する.

$$\sum_{i=1}^m s_i = 1 \dots \dots \dots (11)$$

平均遅延時間  $LATENCY$  は次の式で得られる.

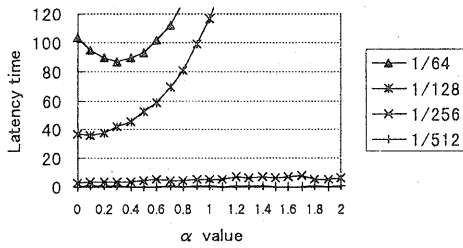


図7 TP法による遅延時間（更新率順方向）  
Fig.7 Latency time with TP method(The update probability is forward distribution.)

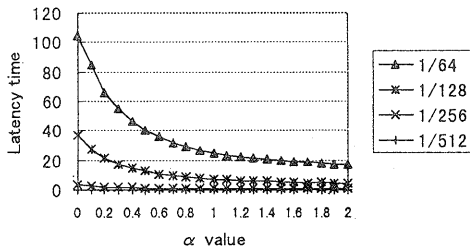


図8 TP法による遅延時間（更新率逆方向）  
Fig.8 Latency time with TP method (The update probability is reverse distribution.)

$$LATENCY = \sum_{j=1}^m \left( \frac{P_j}{\sum_{j=1}^m P_j} \right) \dots \dots \dots (12)$$

$$= \frac{\sum_{j=1}^m \frac{P_j}{s_j}}{\sum_{j=1}^m P_j} = \frac{1}{\sum_{j=1}^m P_j} \left( \frac{P_1}{s_1} + \frac{P_2}{s_2} + \sum_{j=3}^m \frac{P_j}{s_j} \right)$$

$$= \frac{1}{\sum_{j=1}^m P_j} \left( \frac{P_1}{s_1} + \frac{P_2}{1-s_1-\sum_{j=3}^m s_j} + \sum_{j=3}^m \frac{P_j}{s_j} \right)$$

LATENCYを  $s_1$  で微分し、LATENCYが最小値となる  $s_1$  を求める。

$$\frac{\partial LATENCY}{\partial s_1} = \frac{1}{2 \sum_{j=1}^m P_j} \left( \frac{P_1}{s_1^2} + \frac{P_2}{(-s_1 - \sum_{j=3}^m s_j)} \right) = 0 \dots \dots \dots (13)$$

$$s_1 = \frac{1 - \sum_{j=3}^m s_j}{1 + \sqrt{\frac{P_2}{P_1}}} \dots \dots \dots (14)$$

同様に LATENCYが最小になる  $s_2$  を求める。

$$s_2 = \frac{1 - \sum_{j=3}^m s_j}{1 + \sqrt{\frac{P_1}{P_2}}} \dots \dots \dots (15)$$

ここから  $s_1$  と  $s_2$  の最適の比を求めると

$$\frac{s_1}{s_2} = \sqrt{\frac{P_1}{P_2}} \dots \dots \dots (16)$$

が成立する。同様に  $s_2$ 以降についても成立するので、遅延時間を最小にするためには、各ページの出力配分を定常要求率の平方根に比例させるのが最適である。TP法、RxW法、GP法の場合、折衷法を基本としているが、 $\alpha=0.5$ とすると定常要求率の平方根に比例する結果となる。また、ST法の場合も同様に定常要求率の平方根に比例した配分を設定基準値とすることになる。平均更新率が低くなるにつれて、最適  $\alpha$  の値は大きくなる。GP法において  $\alpha$  の値に対して遅延時間が複雑な曲線を示しているが、これは各ページのグループ化において配分率合計が平等に近く分割ができるか否かによるためである。更新の頻度が極めて低い場合には、要求率順方式で十分であることは、容易に判断できる。しかし、更新率が大きく過密な場合は  $\alpha$  の値による平均遅延率の差異は大きい、更新率が小さくなると  $\alpha$  の値が 0.5 以上であれば、 $\alpha$  の値による差異は小さくなる。また更新率とページ数の積が 1 未満となると、出力すべきページが 1 タイムスロットに平均 1 つ以下となるので遅延時間は大幅に減少する。

図6に各  $q$  の値に対して、各手法が最適の  $\alpha$  値をとった場合の遅延時間を示し、表1に各手法についての比較を示す。ST法はTP法とほぼ同等の遅延時間であり、遅延時間を抑えるには有効な手法である。比較の基本となる数値は平均して、1 タイムスロットに1回の計算で済み、TP法に比べて計算時間が小さくて済む。しかし、出力順位保守のための計算が必要であり、この計算は2分探索のため、 $\log_2 m$  回以上の大小判定が必要であるのでGP法よりは一般に計算負荷が大きい。GP法は各グループの代表のみ選択して比較すれば良いので計算負荷はもっとも小さい。しかし、ST法よりは遅延時間が大きい。単純に評価対象を削減したRxW法はあまり効果が上がらず、この手法の採用は得策でない。

(2)定常要求率と更新率分布による差異比較

TP法において、各ページの定常要求率と更新率の分布傾向が異なる場合における遅延時間の傾向を確認する。そのため、次の2つの端的な場合の例を示す。図7は、定常要求率と更新率が、共に  $\theta=1$  の Zipf 分布に従い、各々の率の順番が一致している場合のTP法による遅延時間を示す。図8は図7と同様に共に  $\theta=1$  の Zipf 分布に従うが、各々の率の順番が逆転している場合を示す。他の条件は図2の場合と同じとする。(ただし、各ページの更新率は1以下である。想定した更新率分布において更新率最大のページのページを1以下とするた

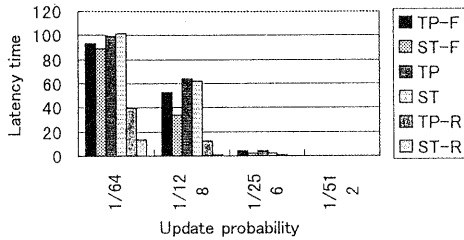


図9 更新率分布別による最小遅延時間の比較  
Fig.9 Minimum latency time at each distribution of update probability

めには、平均更新率は  $2^{-5}$  未満の値としなければならない。) 図7の場合は、最適 $\alpha$ の値は図2で示す更新率が等しい場合より、小さい方向に振れ、図8の場合は、大きい方向に振れている。

個々のページの更新率を把握できない場合を想定する。図2, 7, 8から判断すれば、 $\alpha=0.5$ とすると、各更新率に共通してある程度遅延を小さく抑えることは可能である。図2, 7, 8と同じ条件で各ページの定常要求率と更新率が分布している場合を考える。しかし、その分布が不明であると想定して、各ページの更新率は一定であると仮定し、 $\alpha=0.5$ のパラメータでもってTP法とST法各々を用いてスケジュールを決定した場合の遅延時間を図9に示す。TP-F, ST-Fは、更新率の分布が定常受信率と同じ分布をしており、分布の傾向が更新率と定常受信率とで同じである場合、TP, STは、更新率が各ページで同じである場合、TP-R, ST-Rは更新率の分布が定常受信率と同じ分布をしており、分布の傾向が更新率と定常受信率とで逆の場合(最大の更新率を持つページが最小の定常受信率を持つ。)を示す。更新率が各ページで同じである場合は、TP, ST両手法とも大差はないが、更新率が各ページにより差異がある場合は、TP法よりST法の方が良い結果を生じている。特に分布の傾向が更新率と定常受信率とで逆の場合に、その差が大きい。更新頻度がサーバの放送出力能力より小さければ、遅延時間は極めて小さくなるのでいずれの方式をとっても大差がないので、アルゴリズムを簡単にして最大の $p_i$ 値を選択する要求率順方式、あるいは先着順方式で十分と考えられる。更新率が正確に把握できるのであれば、更新率に応じて最適に近似した $\alpha$ 値を定めればさらに良い結果が得られると予測できる。

## 5. 関連研究

繰り返し出力されるデータをデータベースとみなす

考え方は DataCycle<sup>4)</sup>に見られるが、単一周期の基礎的なスケジュールについて述べているのみである。Imielinskiらは、モバイル端末に対する放送サービスとしてのスケジュール方式について検討している<sup>9)</sup>。しかし、その主眼点は、電力節約のためアクセス回数を少なくするスケジュール方式についてであり、データ更新については述べていない。Acharyaらは、Broadcast Disksという概念を打ち出し、目的とするデータにアクセスするまでの待ち時間を小さくする手法について論じている<sup>10)</sup>。初期要求率の大きい情報は小さい周期で、初期要求率の小さい情報は大きい周期で放送すること、各情報の周期は一定とする概念を打ち出している。しかし、待ち時間軽減を中心としており、遅延時間については触れていない。Acharyaら<sup>11)</sup>、Barbaraら<sup>12)</sup>はそれぞれコンテンツの更新についての研究を発表しているが、コンテンツの更新率と遅延時間の関係についてなされたものではない。我々は、データの定常要求率と更新率から平均遅延時間を小さくする手法について提案した。Aksoyらは、プル型通信において計算が容易な待ち時間を抑えるデータ選択法を述べている<sup>13)</sup>。我々はその手法をプッシュ型通信(データ放送)のデータ選択法に適用し、我々が提案した手法との比較検討を行った。我々は、スケジュールを固定とした場合に遅延時間を軽減する手法<sup>14)</sup>について検討を行った。本論文においては、それを拡張し可変スケジュールにおける遅延時間軽減について考察を行った。

## 6. おわりに

これまでのデータ放送スケジュールの研究には、目的データの受信の待ち時間を短くするという観点から発表されたものが多くある。本論文は、データが更新されてから出力されるまでの平均遅延時間を短く抑えるという観点から考察を行った。特に毎出力のタイミング毎に出力ページを選択する可変スケジュールが可能な場合に遅延時間を軽減する方法を検討した。その方法として先行順方式と要求率順方式を両極端とし、それらを折衷する方式として、ページの定常要求率をべき乗変換してそれまでの待ち時間との積をとる方式(TP法)を考察した。また、TP法の計算負荷を削減するため、評価対象を限定するRxW法とグループ化を行うGP法を定義し、評価を行った。また、それらとは全く別の観点から、コンテンツの更新が発生したとき、次の出力時刻を想定する基準時刻方式(ST法)を提案し、評価を行った。その結果ST法が他の方法に比べて遅延時間が同等以下で候補ページを選択する計算も比較的容易である点を示した。

今後の課題として、ページの大きさが可変長の場合への拡張、ページ受信分布や更新分布がより一般的な場合への拡張、受信障害との関係などが挙げられる。

#### 参考文献

- 1) Acharya S., Franklin M. and Zdonik S., "Disseminating Updates on Broadcast Disks", 22<sup>nd</sup> International Conference on Very Large Data Bases (VLDB96) Bombay India, 1996
- 2) Aksoy D. and Franklin M., "Scheduling for Large Scale On-Demand Data Broadcasting", Proc. IEEE INFOCOM Conf. San Francisco CA, March 1998
- 3) Barbara D. and T. Imielinski, "Sleepers and Workholics: Caching Strategies in Mobile Environments", Proc. ACM SIGMOD Conf. Minneapolis, MN, May, 1994
- 4) Bowen F., Gopal G., Herman G., Hickey T., Lee K.C., Mansfield W. H., Railz J. and A. Weinrib, "The Datacycle Architecture", Communications of ACM, Vol.35, No.12, pp71-81, December 1992
- 5) Franklin M. and Zdonik S., "Dissemination-Based Information Systems", IEEE Data Engineering Bulletin, Vol.19, No.3, September 1996
- 6) Imielinski T. and Viswanathan S., "Adaptive Wireless Information System", Proc. of SIGDBS(Special Interest Group in DataBase Systems) Conference, pp.19-41, Tokyo Japan, October 1994
- 7) Chi-Jiun and Tassiulas L., "Broadcast Scheduling for Information Distribution", IEEE INFOCOM Kobe Japan 1997
- 8) Vaidya N. H. and Hameed S., "Scheduling Data Broadcast in Asymmetric Communication", Proc. of Workshop on Satellite-based Information Services(WOSBIS), New York November 1996
- 9) Vaidya N. H. and Hameed S., "Improved Algorithms for Scheduling Data Broadcast", Texas A&M University Technical Report 96-029 December 1996
- 10) Wong J. H., "Broadcast Delivery", Proceedings of the IEEE, VOL.76, No.12, December 1988 pp1566-1577
- 11) 青野正宏, 田窪昭夫, 渡辺尚, 水野忠則, "データ放送におけるスケジュール決定法「二重循環法」の提案と評価", 情報処理学会論文誌 March 1999