

偏りのある配信木を用いた 低遅延 P2P 情報通知方式の解析

高原誠[†] 鈴木健二[†] 田上敦士^{††} 阿野茂浩^{††}

近年、プッシュ型サービスを用いた災害情報などの小さな情報の配信が増加している。これらの情報は、通知に時間がかかると、本来の意味を失う可能性がある。例えば災害情報速報においては、災害を検知してから通知完了まで長時間経過すると、その対策が間に合わなくなる。このため、これらの情報通知では配信開始から全ての端末への情報通知完了時間を最小にすることが望まれる。本稿では、シミュレーションを用いて、異なる方法で形成される配信木モデルについて、全ユーザに情報を配信するために要する総情報配信時間を比較検討した。さらに、その配信木モデルの中の偏りのある配信木に着目し、分散などのシミュレーションにパラメタを変更し、それがどのように総情報配信時間に影響を与えるかを解析した。

Analysis of Efficient Information Distribution Method with Skewed Distribution Tree over P2P Network

Makoto Takahara[†] Kenji Suzuki[†] Atsushi Tagami^{††}
and Shigehiro Ano^{††}

Recently, plenty of the short messages such as disaster information are distributed over Internet with correspond to the increasing use of Web services. These messages are time sensitive and need to be distributed within certain period. A sign and a warning of natural hazard should be distributed as fast as possible and quick distribution of stock information is very important to the enterprise management. Therefore, realization of the efficient push service over the Internet is expected. We have already reported that P2P communication is promising for the short message distribution to the city size of area with comparing to the server and client distribution. This paper analyzed the total information distribution time to all recipients with different network topologies, balanced and skewed tree, by simulation. Moreover, we have analyzed the contributions of skewed tree in details with several analytical parameters.

1. はじめに

近年、地震や津波などの災害情報通知において、迅速に情報を通知するシステムが注目されている。また教育機関では、コンピュータ上で講義やテストを行うことが増え、同時に多数の端末に対して資料やテスト問題を通知し、処理を行う必要もでてきている。これらの情報は、被害が予測される地域やテストを行う教育機関内など、地理的に限定された狭い範囲に通知することが多く、また、複数の受信者に対して可能な限り迅速に通知する必要がある。そのため、通知者が必要時に情報を通知できるプッシュ型サービスが望まれており、それを実現する同報配信技術が重要である。

一般に同報配信技術には、クライアント・サーバ方式 (C/S 方式)、IP マルチキャスト方式ならびに P2P 方式がある。C/S 方式では、配信サーバが全ユーザに対してユニキャストで情報通知するため、サーバの処理負荷が大きい。IP マルチキャスト方式は、中継するルータがパケットを複製しユーザに通知する方式で、ネットワーク効率は良いが各 ISP のポリシーの相違などから複数の ISP を利用するサービスに対応することが難しい。P2P 方式は、配信サーバが一部のユーザにのみ直接データを送信し、データを受信したユーザがデータ中継の役割を担い、他のユーザにパケットリレー式に通知する方式である。この方式のユーザ間のデータ配信には、ユニキャストを用いるため、ISP をまたがる環境下での同報配信が可能となり、柔軟なネットワークを組むことができる。さらに、C/S 方式でのサーバの役割をユーザが少しずつ分担するため、サーバの処理負荷が少なくなる。

従来 P2P 方式で同報配信を行う研究は、ストリーミング放送などの大容量情報に関することが多い[1][2][3][4]。これらの研究では、地理的にも大きな範囲に適用し、CPU およびネットワーク負荷が一部のノードに集中することを防ぐために P2P を用いている。これに対して本稿が想定する同報配信は、災害情報やテスト問題などストリーミング型コンテンツと比較して小容量な情報を迅速に通知完了することが求められる。さらに、通知される範囲は地理的に広範囲ではなく、局所的である。このため、既存研究とは異なる評価軸で P2P ネットワークの最適化を行う必要がある。

筆者らは先に、小容量情報の低遅延同報配信において、C/S 方式と P2P 方式の比較検討を行い、P2P 方式の方が優位である状況があることを見出した[5]。また、P2P 方式でも、配信木モデルの違いによって、想定される通知時間は異なることを示した[6]。本稿では、配信木モデルの違いを、シミュレーションで比較し、全ユーザに情報を通知するために要する時間 (総情報通知時間) の解析を行ったので以下に報告する。

[†] 電気通信大学大学院
Graduate School of Computer Science, The University of Electro-Communications

^{††} (株)KDDI 研究所
KDDI R&D Laboratories Inc.

2. 想定する配信木モデル

P2P を用いた配信木には、様々な形態が存在する。本章では、配信木モデルについて3種類の形式を挙げ、それらの特徴や構築アルゴリズムならびに構築に必要な最適なパラメタについて述べる。

2.1 配信木を用いた P2P 情報通知方式

筆者らの提案するP2P情報通知方式では、木構造のP2Pネットワークを用いて災害などの情報を通知する。この木構造では、オリジナルデータを持つノード(配信ノード)が根ノード、情報受信者が節点ノードもしくは葉ノードとなり、根ノードから葉ノードへの方向に木構造に沿って実データを通知する。木構造を用いることにより、データの流れる経路を把握し、それぞれのノードまでの通知にかかる時間を概算でき、通知に時間がかかるようならば、木構造を変えてデータ通知経路を変更することができる。このため、木構造をとらないメッシュ型のP2Pネットワークと比較し、通知時間を短縮しやすく、通知時間の揺らぎを抑制しやすい特徴がある。

2.2 配信木モデル

全ての情報受信者に情報が通知されるまでの時間は、データの流れる経路である配信木モデルに影響される。このため、本稿では、平衡配信木、偏りのある配信木、個々のリンク遅延を考慮した配信木の3つの配信木モデルの比較を行い、それぞれの特徴を解析する。以下に、それぞれの配信木の特徴と構築アルゴリズムの概略について述べる。ここで、 ΔT は次のノードに通知可能となるまでのプロセス処理時間、 R はノード間のリンク遅延とする。

ノードが新規参加するときの、各配信木の構築アルゴリズムを用いた親ノードの選定手順を図1に示す。配信ノードは、各配信木の親ノード候補優先方針に基づき、優先度の高いノードから親ノード候補に加える。ただし、総情報通知時間は配信木の高さに大きく影響されるため、高さが想定以上に高くないように、親ノード候補に必要条件(表1)を設ける。この必要条件は、親ノード候補が最優先の親ノード候補よりリンクの遅延の平均値 R_{ave} より大きくならないようにそれぞれ設定している。木を構成するノードの中からの親ノード候補検索は、配信木を構成するノードの中に必要条件を満たすノードがなくなったとき、もしくは優先順にノードを親ノード候補に加えてX個以上になったときに終了する。ここで、親候補ノード数がX個より大きくなるのは、優先度が同じノードが複数存在するためであり、このとき、親ノード候補がX個になるように、最後に親ノード候補に加えた優先度の同じノード群の中から、削除するノードをランダムで選択する。これにより、作成された親ノード候補は、新規参加ノードに通知され、新規参加ノードと各親ノード候補の間のリンク遅延を計測し、リンク遅延が最小となるノードを親ノードとする。

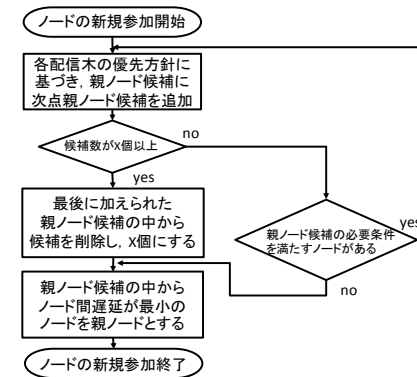


図1 親ノードの選定手順
 Fig. 1 The parent node decision flow

表1 親ノード候補の必要条件

Table 1 Prerequisites for parent node nominee

配信木の種類	親ノード候補の必要条件
平衡配信木	ノードの深さが 最優先親ノード候補の深さ+1以内
偏りのある配信木	ノードの次配信時刻が 最優先親ノード候補の次配信時刻 (概算値)+平均リンク遅延 R_{link} 以内
個々のリンク遅延を 考慮した配信木	ノードの次配信時刻が 最優先親ノード候補の次配信時刻 (計測値)+平均リンク遅延 R_{link} 以内

(1) 平衡配信木 (Balanced Tree : BT)

平衡配信木は、根ノードから全ての葉ノードまでの距離(ホップ数)をなるべく等しくした木である。さらに、ここでは配信ノード、全ての中継ノードの持つ最大子ノード数を等しくしたものを扱う。ここで、 n を総受信ノード数、 m を1ノードあたりの子ノードの数の最適値、 h を配信木の高さ(葉ノードから根ノードまでの枝の数)とする。図2に、木の高さ $h=2$ 、子ノード数の最適値 $m=3$ とした配信木モデル図を示す。図中の円内の番号はノードの深さを表す。全ユーザに対して通知し終わるまでの時間を最小化するために、最後に通知されるノードまでにかかる時間を総情報通知時間 D とする。

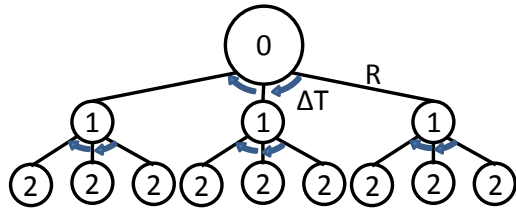


図 2 平衡配信木モデル

Fig. 2 The balanced tree model

最後に通知されるノードは、 $(m-1) \times h$ 個の処理時間 ΔT と h 個のリンク遅延 R を要するため、総情報通知時間 D は以下の式で表される。

$$D = ((m-1) \times \Delta T + R) \times h$$

また、木の高さ h は、高さは自然数であることから、

$$h = \lceil \log_m \{(m-1)(n+1) + 1\} \rceil - 1$$

となる。ただし、 $\lceil x \rceil$ とは x 以上の最小の整数を表す。この数式より、総情報通知時間 D を最小化する子ノード数の最適値 m を導き出すことができる。この子ノード数の最適値の算出を新規参加ノードがある場合に行い、最適値 m を更新する。この結果、算出後の最適値が前の最低値よりも小さくなる場合には、単純にノードを参加させるのではなく、木の再構築が必要となる。

親ノード候補は、算出した最適値より少ない子ノードしか持たず、根ノードからの距離（ホップ数）が小さいノードを優先する。

(2) 偏りのある配信木 (Skewed Tree : ST)

偏りのある配信木は、平衡配信木のように全ての中継ノードが同じ子ノード数を持つのではなく、時間的に早く情報通知されたノードが、より多くの子ノードを担当するように偏りをつけた木である。そうすることにより、平衡配信木で起こりうる未受信ノードが存在するにも関わらず、配信データを保持し通知可能であるノードが配信しないという状況が少なくなる。一方で、平衡配信木よりも負荷が分散せず、一部のノードに負荷が集中する特徴がある。

親ノード候補の優先順を決定するために、リンクの遅延の平均値 R_{ave} 、 ΔT 値をあらかじめ算出しておき、各ノードに情報が通知される時間を概算する。この概算時間を TimeLevel (以後、TL) とする。 R_{ave} を 40、 ΔT を 1 とした配信木モデル図を図 3 に示す。図中では、各ノードに到達する概算時間 TL をノード円内に示し、配信木は TL=82 までに通知できるノードで構成している。

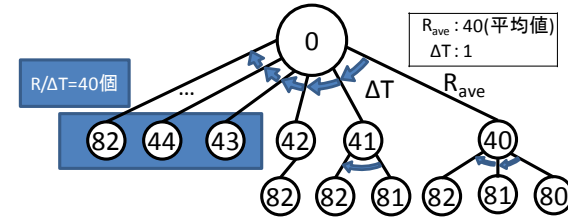


図 3 偏りのある配信木モデル

Fig. 3 The skewed tree model

R_{ave} を用いて最適トポロジを計算しているため、分散 $\sigma^2=0$ のときは、配信ノードや中継ノードは一斉に配信完了となる。この概算時間 TL を用いて、新規参加ノードは最も早く通知可能なノードを優先して親ノード候補として選ぶ。

ここで、総ノード数 N (総受信ノード数 n + 配信ノード) と配信木の最大 TL (MaxTL) の関係について述べる。TL 時間までに通知できる総ノード数 $N(TL)$ は、 $N(TL - \Delta T) + \Delta T$ ごとの N の増加量で表わすことができる。この ΔT ごとの N の増加量とは、TL が ΔT 増えたときに新たに通知可能となるノードによって増加するノードであるため、 $R_{ave} / \Delta T$ 時間前に配信木に参加したノード数と一致する。このため、総ノード数は下記の式で表わせる。

$$N(TL) = \begin{cases} 1, & 0 \leq TL < R_{ave} / \Delta T \\ N(TL - \Delta T) + N(TL - R_{ave} / \Delta T), & R_{ave} / \Delta T \leq TL \end{cases}$$

MaxTL は TL を増やしていき、初めて総ノード数を越えた TL の値となる。次いで、この式を下記のように変形させる。

$$\begin{aligned} N(TL) &= N(TL - \Delta T) + N(TL - R_{ave} / \Delta T) \\ N(TL - \Delta T) &= N(TL - 2\Delta T) + N(TL - \Delta T - R_{ave} / \Delta T) \\ &\vdots \\ +) N(R_{ave} / \Delta T) &= N(R_{ave} / \Delta T - \Delta T) + N(0) \end{aligned}$$

$$N(TL) = N(R_{ave} / \Delta T - \Delta T) + \sum_{k=0}^{TL - R_{ave} / \Delta T} N(k)$$

$N(R_{ave} / \Delta T - \Delta T) = 1$ であるため、下記の式となる。

$$N(TL) = \begin{cases} 1, & 0 \leq TL < R_{ave} / \Delta T \\ 1 + \sum_{k=0}^{TL - R_{ave} / \Delta T} N(k), & R_{ave} / \Delta T \leq TL \end{cases}$$

この式は、配信ノード(1) + 子ノードを根とする部分木ノード数の和を表わしている。この式により、各子ノードに割り当てるノード数を計算することができる。

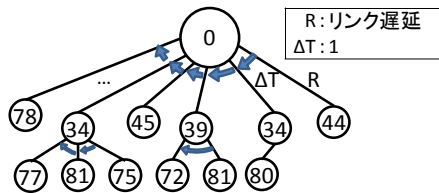


図 4 個々のリンク遅延を考慮した配信木モデル

Fig. 4 The skewed tree model based on link delay

(3) 個々のリンク遅延を考慮した配信木 (Skewed Tree based on Link Delay : STLD)

個々のリンク遅延を考慮した配信木は、個々のリンク遅延を計測し、最適な親ノードを選択するよう構成した木である。ΔT と実際に計測したリンク遅延 R を用いて、各ノードへの情報通知時間を正確に計算する。この値を用いて、新規参加ノードは、最も早く通知可能なノードを優先して親ノード候補として選ぶ。配信木モデル図を図 4 に示す。図中では、R を各リンク遅延、ΔT を 1 とし、各ノードにデータが到達する時間をノード円内に示す。平衡配信木と偏りのある配信木よりも正確なリンク遅延を用いているため、最適な木構造を取りやすいが、最適な親ノード候補を選出するときは、最も早く通知可能なノードを選ぶ必要があるため、全ノードの通知される時間を管理する必要がありコストが高い。

3. 配信木シミュレーションを用いた総情報通知時間の比較

本章では、前章で述べた P2P 情報配信木モデルについてシミュレーションを用いて総情報通知時間 D を比較する。このシミュレーションでは、リンク遅延 R を正規分布によって与え、ノードの参加のみについて扱う。

3.1 シミュレーション設定

本シミュレーションは地震発生時に、震源近くのすべての受信者に災害情報を通知することを想定してパラメタを設定する。地震発生時に通知したい受信者は、地理的に局所にいるため、一般的なメッセージ通知に対して、リンク遅延のばらつきが小さい。最大受信ノード数は地域に住む家庭数であるため 10,000 ノードとし、ノード間のリンク遅延は平均 $R_{ave}=40ms$ 、分散 σ^2 の正規分布に基づくものとする。実ネットワークを用いて、処理遅延の計測[5]の計測に基づき、処理遅延 $\Delta T=0.4ms$ に設定する。さらに、シミュレーションでの総情報通知時間 D の計測は、200 ノードが新規に参加するごとに行う。シミュレーションの結果は、本試行を 50 回繰り返す、それぞれの平均値についての結果を示した。

また、平衡配信木ではノードの参加毎に最適な子の数を計算し、その値が現在の値が減少したときの再構成処理を行っている。再構成方法は、まず、配信ノードが現在

の配信木からノード間のリンク遅延の小さなノードを選び、子ノードとする。次いで、新たな配信木に加わったノードがさらに遅延の小さなノードを選び子ノードとする。この工程を繰り返すことで新たな配信木を作成した。

3.2 シミュレーションの結果

(1) 配信木モデルの比較

パラメタ(親ノード候補数 $X=5$, 分散 $\sigma^2=\{0,5,10,15\}$) が等しいとき、平衡配信木 (BT)、偏りのある配信木 (ST)、個々のリンク遅延を考慮した配信木 (STLD) のそれぞれのモデルにおける総情報通知時間 D の違いを調査する。シミュレーション結果を図 5 に示す。すべての場合において、個々のリンク遅延を考慮した配信木 (STLD) の総情報通知時間 D が最も短い。また、ノード数が増加するに従って総情報通知時間 D は増加するが、平衡配信木 (BT) の場合は受信ノード数 6,000 ノードで総情報通知時間 D が短くなった。これは、受信ノード数の増加により、算出した最適な子ノード数 m が変化して再構築が発生し、リンク遅延のより小さなリンクを用いて配信木が構築されたためである。平衡配信木 (BT) と偏りのある配信木 (ST) の比較では、受信ノード数 6,000 ノード以下の範囲において、偏りのある配信木 (ST) の総情報通知時間 D は短く、分散が大きくなるに従いその差がなくなる傾向にある。偏りのある配信木 (BT) と個々のリンク遅延を考慮した配信木 (ST) の比較では、分散 $\sigma^2=0$ のとき両方の木とも等価となり、総情報通知時間 D も同じ値をとった。そして、分散が高くなるに従い、総情報通知時間 D の差は大きくなった。

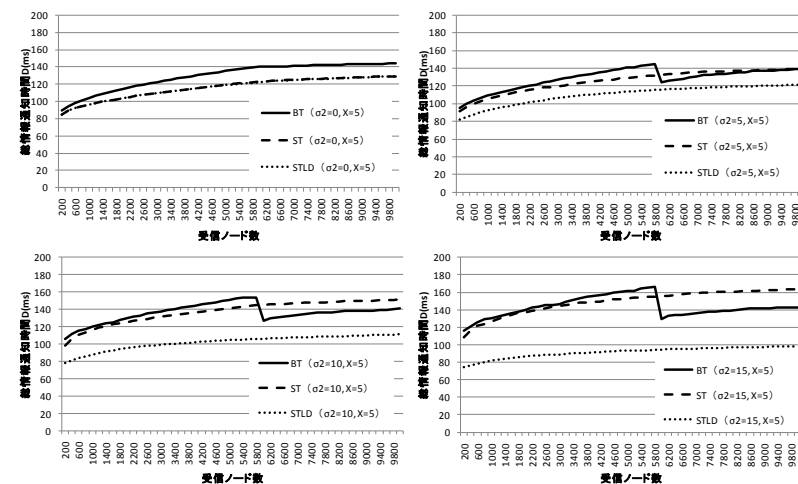


図 5 配信木モデルの比較

Fig. 5 Comparisons of BT, ST and STLD model

(2) 分散 σ^2 の比較

災害情報を通知する範囲が広がると、ノード間のリンク遅延にばらつきが大きくなる。そのため、各配信木においてリンク遅延の分散が総情報通知時間 D にどのような影響があるのかを調べる。分散 $\sigma^2=\{0,5,10,15\}$ 、親ノード候補数 $X=5$ とし、各配信木モデルのシミュレーション結果を図 6 に示す。平衡配信木 (BT) と偏りのある配信木 (ST) は分散が高くなるに従い、総情報通知時間 D が増える傾向になった。これは、平均リンク遅延 R_{ave} に基づき最適なトポロジを形成しているため、経路のリンク遅延が R_{ave} よりも小さい箇所を有効に使えないためである。

一方、個々のリンク遅延を考慮した配信木 (STLD) では、分散が高いほど、リンク遅延の小さなリンクを有効に使うことができるため、総情報通知時間 D は短くなった。このように、平衡配信木と偏りのある配信木は、配信木のトポロジは改良の余地があるため、近隣ノードの情報を用いて自律的に改善するような仕組みを組み込む必要がある。

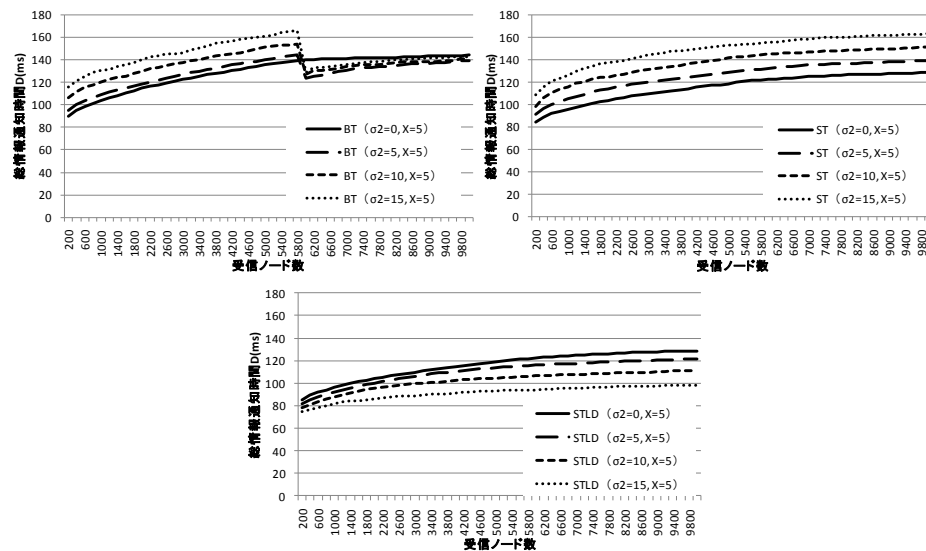


図 6 分散 σ^2 の比較

Fig. 6 Comparisons of variance σ^2 for BT, ST, STLD model

(3) 親ノード候補数 X の比較 (傾きのある配信木)

前節で述べたように、傾きのある配信木 (ST) において分散 σ^2 が大きい場合、リンク遅延の小さいリンクを利用できていない。これに対して、親ノード候補数 X を大きくとることにより、参加時に実測するノードを増加させ最適化が可能であるかどうかの確認を行った。

分散 $\sigma^2=\{5,10,15\}$ のとき、親ノード候補数 $X=\{1,5,10\}$ を変化させたときの総情報通知時間 D のシミュレーション結果を図 7 に示す。分散 $\sigma^2=0$ のときは、親ノード候補数による違いがほぼないため省略した。偏りのある配信木 (ST) では、親ノード候補数を増やすと、総情報通知時間 D は短くなる傾向にある。親ノード候補数 1 と 5 では大きく差があるが、5 と 10 では差が小さい。これは、親ノード候補数 5 で、十分な親候補が選ばれているためである。また、親ノード候補数の増加による総情報通知時間 D の収束は、分散が大きくなるに従い遅くなった。これは、分散が大きければ、最適に近い親ノードを選ぶことができる親ノード候補数が増加するためである。

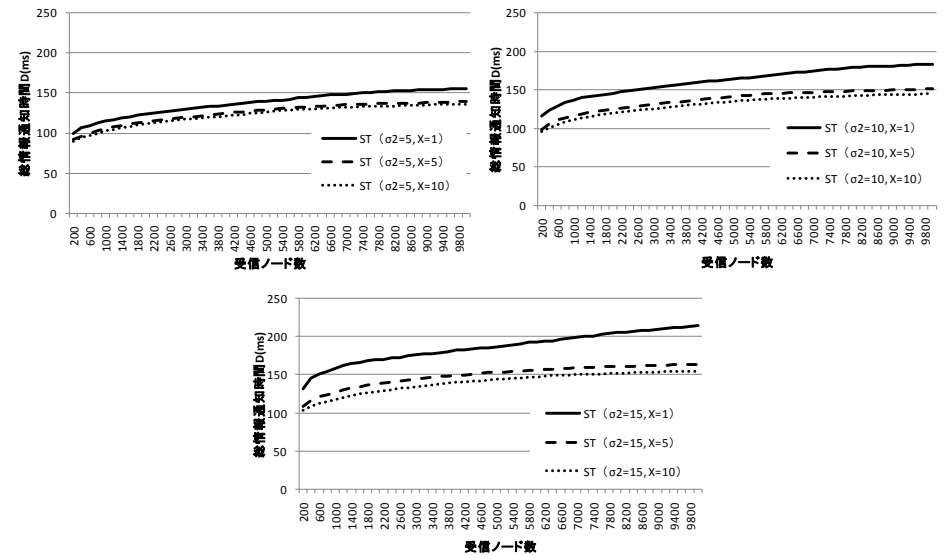


図 7 親ノード候補数比較

Fig. 7 Comparisons of parent node nominee for ST model

4. 考察

3章の評価より,最も総情報通知時間 D が短いのは,いかなる条件下においても個々のリンク遅延を考慮した配信木 (STLD) であった. さらに, STLD は新規ノード参加時に最も早く通知可能なノードを親候補として選択可能であるため,分散が高いほど総情報通知時間 D の短縮に有効であることが分かった. しかしながら, STLD を実現するためには,配信木の構造やそれぞれのリンク遅延を1カ所(通常は配信ノード)で保持し,最短で通信可能なノードを導出する必要があるため,配信木の維持管理に必要なコストが高くなる.一方,平衡配信木 (BT) と偏りのある配信木 (ST) は配信木のトポロジさえ分かれば,参加すべきノードが分かるため,配信木の維持管理に必要なコストは低い.さらに,リンク遅延の分散 σ^2 が小さい場合には,STLD と BT/ST の総配信遅延時間の差は小さい.そのため,本稿が前提とする受信者が局所的に配置されている環境では,BT や ST の利用も現実的であると言える.

3.2(1)節で述べた配信木モデルの比較より,BT と ST では,10ms 程度 ST の方が総配信遅延時間 D は短くなると言える.この時間は,人が感知して動くには小さな時間であり意味をなさないが,情報通知システムと同期して,地震時に家電を止める等の管理を行う場合は有効になりうる.また,株価など種類が多く,迅速に通知する必要がある情報を同じ通知プラットフォームで扱い,必要に応じて差分や結合などの処理を行う場合は, ΔT が大きくなるため,BT と ST の差が更に大きくなると考えられる.配信形式として,ノード間のリンクを保持する方法と必要時にリンクを繋ぐ方法がある.リンクを保持する場合,遅延等を計測するなど最適化することが可能であるが,その分,普段より処理コストが必要となる.一方,必要時にリンクを繋ぐ方法では,IP アドレス等の情報送信に必要な情報を配信元ノードで一元管理し,情報通知が必要となるとき,配信ノードが受信ノードに配信先リストを実データと共に通知し,その受信ノードに一部の配信を任せることが可能である.このとき,通知順と宛先リスト配分を調整することで,コストを抑えて BT と ST を実現することができる.ただし,中継をノードの離脱が生じたときに,配信しようとする,余計な遅延が発生するため,中継ノード,特に,責任の大きなノードの選定は,ノードの生存率や生存確認の頻度を工夫する必要がある.

5. おわりに

本稿では,地震災害警報などの情報通知を迅速におこなうため,低遅延 P2P 配信木の構築法について,配信木モデルを変えてシミュレーションを行い,総情報通知時間について評価をした.この結果,総情報通知時間に関して,個々のリンク遅延を考慮した配信木 (STLD) が最適であることを明らかにした.また,リンク遅延の分散が小さい場合は,個々のリンク遅延を考慮した配信木 (STLD) と偏りのある配信木 (ST) の差は小さいことを示した. STLD は全てのリンク遅延を配信ノードが管理する必要があり,管理コストが大きい.これに対して配信ノードが局所的に配置されており,リンク遅延の分散が十分小さい場合には,管理コストが小さい ST を用いることが可能であると言える.

今後は,ノードの離脱やネットワークの近接性を考慮した評価,および実ネットワークを用いた P2P 遅延計測シミュレーションを検討する予定である.

参考文献

- 1) Hosseini, M., Ahmed, D. T., Shirmohammadi, S., Georganas, N. D.: A Survey of Application-Layer Multicast Protocols, IEEE Communications Surveys & Tutorials 9(3),pp.58-74 (2007)
- 2) D. Pendarakis, S. Shi, D. Verma, and M. Waldvogel.: ALMI: An Application L Multicast Infrastructure, 3rd Usenix Symp. Int'l. Tech. and Sys., pp. 49-60 (2001)
- 3) Y. Chu, S. G. Rao, and H. Zhang: A Case for End System Multicast, IEEE JSAC, vol. 20, no. 8, pp. 1456-71 (2002)
- 4) M. Castro, P. Druschel, A.-M. Kermarrec, and A. Rowstron.: SCRIBE: A large-scale and decentralized application-L multicast infrastructure, IEEE Journal on Selected Areas in Communications Vol.20 No.8, pp.1489-1499 (2002)
- 5) 高原誠, 鈴木健二, 田上敦士, 阿野茂浩: 低遅延 P2P 情報配信プラットフォームの一考察, 情報処理学会第 71 回全国大会, 1E-6, pp.37-38 (2009)
- 6) 高原誠, 鈴木健二, 田上敦士, 阿野茂浩: P2P プラットフォームによる更新情報の低遅延配信方式の提案, マルチメディア・分散・協調とモバイル(DICOMO2009) シンポジウム, 5E-3, 047-1054(2009)