

ゴシップ型マルチキャストにおける メンバーシップビューの構造化が信頼性に及ぼす影響に関する研究

On the Effects of Structuring Partial Membership Views
on the Reliability of Gossip-Based Multicast

五十嵐 大輔[†]
Daisuke Igarashi

土屋 達弘[†]
Tatsuhiko Tsuchiya

菊野 亨[†]
Tohru Kikuno

1. はじめに

ゴシップ型マルチキャストは実装が容易であり、信頼性が高いという特徴から、注目を集めている。ゴシップでは、メッセージを受け取ったプロセスは、他のプロセスを幾つかランダムに選択し、それらにメッセージを転送する。このような通信が繰り返されることによって、情報がシステム全体に伝搬される。この手法は分散データベースなど、様々な状況で用いられている [2]。

従来のゴシッププロトコルにおいては、プロセスはマルチキャストを行うグループの全メンバーの中からランダムに送信先のプロセスを選択する。そのため、各プロセスがグループ全体のメンバーシップ情報を保持しなければならない、スケーラビリティに問題があった。

この問題に対処するため、各プロセスがグループの部分的なメンバーシップ情報を保持して動作するプロトコルが開発されてきた [1, 3]。これらのプロトコルでは、メンバーシップ情報が均等に共有されるように、各プロセスの知り得る他プロセスがランダムに選択される仕組みをも提供している。つまり、これらは、ランダムなトポロジーを有するオーバーレイネットワークを用いたゴシッププロトコルと言える。実験的評価により、これらのプロトコルが高い信頼性を有することが分かっている。しかし、トポロジーがランダムでない場合のゴシップの信頼性については、殆ど研究がなされていない。そこで、本研究では、適切なグラフ構造を持つオーバーレイネットワークを利用することによって、負荷分散等の望ましい性質を保持しながら、同時に高信頼なマルチキャストを実現できることを示す。

しかしながら、正確なトポロジーのグラフを利用することは、プロセスのマルチキャストグループへの加入及びグループからの脱退が頻繁に起こる動的な分散環境において困難である。そこで、この問題を解決するため、与えられたグラフ構造を近似する自律分散的なメンバーシッププロトコルを提案する。この近似手法を用いることにより、満足できる信頼性が達成されうることを、実験結果により示す。

2. メンバーシップビューとグラフ表現

本研究で考えるゴシッププロトコルを図 1 に示す。各プロセス i はメンバーシップビュー $view_i$ を保持する。但し、この $view_i$ にはプロセス i 自身は含まれないものとする。マルチキャストメッセージを受信したプロセス i は $view_i$ の中からランダムに f 個のプロセスにメッセージを転送する。 f は定数であり、*fanout* と呼ばれる。

[†] 大阪大学大学院情報科学研究科情報システム工学専攻

m のマルチキャストを開始：

m を $view_i$ からランダムに選択された
 f 個のプロセスへ送信；

i が j からマルチキャスト m を受け取ったとき：

if (m を初めて受信)

m を $view_i \setminus \{j\}$ からランダムに選択された
 f 個のプロセスに送信；

図 1: 部分的メンバーシップ情報を用いたゴシップ

プロセスのビューから得られるオーバーレイネットワークのトポロジーは有向グラフで表現できる。各頂点は各プロセスに相当し、 $j \in view_i$ であるときに限り i から j への辺が存在する。以下に、システムが有すべき性質、及びそれを実現するためにこのグラフが満たすべき条件を示す。

スケーラビリティ：頂点の出字数は頂点数の増加に対して、緩やかに増加すること。

トポロジーグラフにおけるノード i の出次数は i のビューサイズと等しい。スケーラビリティを実現するためには、システムが大きくなっても各プロセスのビューサイズが小さく保たれる必要がある。

負荷分散：正則であること。

負荷分散を実現するため、各プロセスについて、1) ビューの中に含まれるプロセス数が等しく、2) そのプロセスを含むビューの数が等しい、という意味でバランスのとれたビューを持つことが望ましい。

伝搬速度：直径が小さいこと。

直径が小さければ、マルチキャストメッセージがグループ全体に伝搬する時間が短縮される。

本研究では、トポロジーグラフとして Kautz グラフ及び de Bruijn グラフ、Harary グラフの 3 種類のグラフを考える。これらのグラフは上記の特性をバランス良く有している。まず、正則もしくはほぼ正則なグラフである。次数及び直径についても優れており、例えばノード数 110、直径 2、次数 10 の Kautz グラフ、ノード数 100、直径 2、次数 10 の de Bruijn グラフが構成できる。

3. グラフを近似する自律分散的手法

以下に与えられたトポロジーを自律分散的に近似するアルゴリズムを示す。ここで、 E は与えられたトポロ

ジークラフであり、そのノード数を n とする。また、 E 中の各ノードには、0 から $n-1$ までの番号が重複なく付与されているものとし、各プロセスはこの E における各ノードの接続関係を把握しているものとする。

[加入 (subscription / join)]

- 1) プロセス i が加入要請を任意のプロセスに送ることによりグループに加入する。
- 2) プロセス i から加入要請を受け取ったプロセスは、0 から $n-1$ まで一様に分布するランダムな整数 $i.rnd$ を生成し、 i に通知する。また、そのノードは i のアドレスと共に番号 $i.rnd$ をゴシップを用いて既存のグループにマルチキャストする。
- 3) プロセス j が i のアドレスと割り当てられたランダムな数 $i.rnd$ を受け取ったとき、 j は以下のような処理を行う。
 - (a) $(j.rnd, i.rnd) \in E$ ならば、 j は新たな加入プロセス i を $view_j$ に加える。
 - (b) $(i.rnd, j.rnd) \in E$ ならば、 j は $view_i$ に j を保持させておくために、 i に対して j のアドレスを送信する。
- 4) 加入プロセス i は $view_i$ に受け取った全てのアドレスを追加する。

[脱退 (unsubscribe / leave)]

- 1) プロセスは脱退メッセージをマルチキャストした後、グループから脱退する。
- 2) プロセスがビューに保持しているプロセスから脱退メッセージを受け取ったら、単に脱退プロセスをビューから削除する。

4. 実験的評価

この節では、プロセス数 500 のマルチキャストグループを対象として、シミュレーションにより得られた実験結果を示す。全プロセスが同確率でマルチキャストを開始するという仮定の下で、すべての生存プロセスがマルチキャストメッセージを受け取る平均確率を計算した。fanout f は 13 とし、通信路故障による各メッセージの損失確率を 0.05 と仮定した。また、プロセスの故障は全プロセスの半数までとした。結果はプロセス故障確率の各々について 20 個の故障パターンの平均をとった。

評価したトポロジークラフは以下の 3 種類である。

1)Kautz グラフ (直径 2, 次数 22), 2)de Bruijn グラフ (自己ループなし, 直径 2, 次数 23), 3)Harary グラフ (次数 23)。ここで、ノード数 500 の Kautz グラフや de Bruijn グラフは存在しないため、500 以上の最小ノード数を持つこれらのグラフから余分なノードを除去したものを利用した。

上記 3 種類のグラフに加え、4)Harary グラフ (次数 13), 5) ランダムにビューを生成したグラフ、の 2 種類のグラフを比較した。次数 13 の Harary グラフの利用は、決定的フラッディング [4] と等価である。一方、各ノードが同じ次数のランダムにビューを生成したグラフの利用は LPBCAST プロトコル [1] に相当する。

得られた結果を表 1 に示す。故障プロセスの割合に関わらず、次数 13 の Harary グラフを除く 4 つのグラフの信頼性は、ほぼ同程度であることが分かった。

表 1: マルチキャストの信頼性 (単位は%)

	プロセス故障割合					
	0 %	10 %	20 %	30 %	40 %	50 %
1)Kautz	100	100	100	99.99	99.98	99.95
2)de Bruijn	100	100	100	99.99	99.99	99.98
3)Harary	100	100	100	100	100	99.99
4)Harary	100	100	100	99.99	99.81	97.74
5)Random	99.83	99.87	99.88	99.81	99.93	99.76

表 2: 近似手法を用いた場合のマルチキャストの信頼性

	プロセス故障割合					
	0 %	10 %	20 %	30 %	40 %	50 %
1)Kautz	99.12	98.82	98.62	99.32	99.28	99.76
2)de Bruijn	99.45	99.80	99.73	99.65	99.47	98.50
3)Harary	100	100	100	100	100	99.93
4)Harary	99.97	99.45	99.83	96.10	87.73	27.82
5)Random	99.83	99.87	99.88	99.81	99.93	99.76

次に、グラフを近似する自律分散的手法についてのシミュレーション結果を表 2 に示す。この結果から、表 1 に示したものよりも信頼性が低下しているものの、次数 23 の Harary グラフのマルチキャスト成功率はほぼ 100 % と高い数値を示している。

5. まとめ

システムが有すべき負荷分散等の望ましい性質を、上記 3 つのグラフが持ち、そのようなトポロジークラフにおいて、ゴシップ型マルチキャストの信頼性が高く維持できることをシミュレーションにより示した。プロセス故障の割合に関わらず、ランダムに生成した同サイズのビューから得られるオーバーレイと比べて、上記の 3 つのグラフは同程度の高い信頼性を示した。更に、与えられたトポロジークラフを自律分散的に近似する手法により得られたグラフを用いた場合でも、元のグラフに匹敵するほどの高信頼性が達成できることがわかった。

参考文献

- [1] P. T. Eugster, R. Guerraoui, S. B. Handurukande, A.-M. Kermarrec, and P. Kouznetsov. Lightweight probabilistic broadcast. In *Proceedings of the 2001 International Conference on Dependable Systems and Networks (DSN '01)*, pages 443–452, Gothenberg, Sweden, July 2001.
- [2] P. T. Eugster, R. Guerraoui, and A.-M. Kermarrec. Epidemic information dissemination in distributed systems. *IEEE Computer*, 37(5):60–67, May 2004.
- [3] A. J. Ganesh, A.-M. Kermarrec, and L. Massoulié. Peer-to-peer membership management for gossip-based protocols. *IEEE Transactions on Computers*, 52(2):139–149, February 2003.
- [4] M.-J. Lin, K. Marzullo, and S. Masini. Gossip versus deterministically constrained flooding on small networks. In *Proceedings of 14th International Conference on Distributed Computing, (DISC 2000)*, volume 1914 of *Lecture Notes in Computer Science*, pages 253–267, Toledo, Spain, October 2000.