

# アドホックネットワークにおけるデータ更新を考慮した複製配置について

原 隆浩

3M-01

大阪大学大学院工学研究科情報システム工学専攻

## 1 はじめに

近年、移動体にルータの機能をもたせて、移動体のみの一時的なネットワークを形成するアドホックネットワークに関する注目が高まっている [1, 2]. アドホックネットワークでは、移動体の移動によってネットワークが分断された場合に、分断された部分ネットワーク内のデータに対してアクセスできないため、データの利用率が低下してしまう。この問題を解決する手法として、データの複製を作成し、オリジナルデータをもつ移動体以外に配置することが有効である。

そこで、筆者らはこれまでに、各移動体の物理領域に制限があり、データの更新が起こらないアドホックネットワークを想定して、データの複製を配置する手法を提案した [3]. 更新が起こらないデータの例としては、発掘調査のデータなどのように追加のみが起こるものや、気象データのように最新の情報であることが厳密には要求されないようなものがある。

一方、センサーから定期的に得られるデータなど、更新が頻繁に起こるデータも考えられる。そこで本稿では、データの更新が定期的に起こるアドホックネットワークを想定して、データの利用率向上のために複製を配置する手法を提案する。

## 2 複製配置方式

本稿では、次のようなシステム環境を想定する。

- サイズの等しい  $n$  種類のデータ ( $D_1, D_2, \dots, D_n$ ) が存在し、各々が特定の移動体にオリジナルデータとして保持されている。
- 各移動体は、自身をもつオリジナルデータ以外に、データ  $C$  (定数) 個分の物理領域をもち、複製を作成する。
- 各データ  $D_i$  ( $i = 1, \dots, n$ ) の更新が周期  $T_i$  で発生する。オリジナルが更新された後、古い複製は無効になる。

自身をもつオリジナルデータと複製に対してアクセス要求が発生したとき、即座にアクセスは成功する。自身が、アクセス対象のデータやその複製をもっていない場合は、ネットワーク内にアクセス要求をブロードキャストし、相互接続された移動体のいずれかがそのデータや複製をもっているとき、アクセスは成功する。それ以外は失敗である。

ここで、最適な複製配置の決定は計算量も大きく非常に困難であることから、本稿では、次の方針に基づくヒューリスティックな方式を提案する。

- 複製の再配置を、一定の周期 (再配置周期) ごとに行う。
- 各移動体の各データへのアクセス頻度と、次のデータ更新までの時間、および、再配置周期の時点でのネットワークの接続状態に基づいて複製の配置を行う。

### 2.1 E-SAF (Extended Static Access Frequency) 方式

E-SAF 方式では、再配置周期ごとに、各移動体が物理領域の許す限り、自身をもつオリジナルデータ以外で、次の値が高いデータから順にその複製を配置する。

$$p_i \cdot \tau_i = p_i \cdot (T_i - t_i) \quad (1)$$

On Replica Allocation Considering Data Update in Ad Hoc Networks  
Takahiro HARA  
Department of Information Systems Engineering, Graduate School of Engineering, Osaka University.

ここで、 $p_i$  はデータ  $D_i$  へのアクセス頻度、 $\tau_i$  は次に  $D_i$  が更新されるまでの時間、 $t_i$  は前回  $D_i$  が更新されてから経過した時間を表している。本稿では、上式を PT 値と呼ぶ。PT 値は、そのデータが次に更新されるまでに発生するアクセス要求の平均回数を表している。PT 値の大きいデータの複製を配置することで、再配置周期の時点でのアクセス成功率を最も高くできる。データの更新後、保持している複製は無効となるが、その後、更新後のデータへのアクセスが成功したとき、新たにその複製を配置する (他の 2 方式でも同様)。

E-SAF 方式は、複製配置の際に他の移動体と情報交換を行う必要がないため、複製配置のオーバーヘッドが小さい。一方、各移動体が自身のアクセス頻度のみに基づいて複製を配置するため、同様のアクセス特性をもつ複数の移動体では同じデータの複製を作成してしまう。相互接続されている移動体同士は、他の移動体が所持しているデータや複製にアクセスできるため、データや複製の重複を少なくして効果的に共有する方が、アクセス成功率を向上できる。したがって、この方式は、同様のアクセス特性をもつ移動体が多い場合に性能が低下してしまう。

### 2.2 E-DAFN (Extended Dynamic Access Frequency and Neighborhood) 方式

E-DAFN 方式では、E-SAF 方式における複製の重複が多いという問題を軽減するために、隣接する移動体間で重複を解消する。そのアルゴリズムは、次のようになる。

1. 再配置周期がくると、各移動体は自身の識別子をブロードキャストする。すべての移動体のブロードキャストが終了すると、各移動体は受け取った識別子の情報から、自身と相互接続されている移動体を知ることができる。
2. 各移動体は、E-SAF 方式に基づいて、配置する複製を暫定的に決定する。
3. 識別子  $M_i$  の添字  $i$  が最も小さい移動体から、幅優先探索の順に全ての隣接する移動体間で、以下の手順を繰り返す。隣接する移動体とデータや配置する複製に重複があるとき、一方がオリジナルデータの場合は複製をもつ移動体、両方とも複製の場合はそのデータの PT 値が低い方の移動体において、配置すべき複製を変更する。変更の際には、両方の移動体において複製の配置対象となっていないデータのうち、その移動体からの PT 値が最も高いものの複製を配置する。

E-DAFN 方式では、隣接移動体間のデータの重複を解消することで、数多くの種類の複製を隣接移動体間で共有できる。ただし、幅優先探索の順でネットワークを 1 回走査することにより重複を解消するため、隣接移動体間での複製の重複が完全に解消されるわけではない。さらに、再配置周期ごとに隣接移動体間での情報交換と複製の再配置が行われるため、E-SAF 方式に比べてオーバーヘッドが大きい。

### 2.3 E-DCG (Extended Dynamic Connectivity based Grouping) 方式

E-DAFN 方式が隣接移動体間で複製を共有するのに対して、E-DCG 方式ではさらに広範囲の移動体間で複製の共有を行う。効率的に複製を共有するためには、複製を共有する移動体のグループの安定性が高く、トポロジの変化によって分断されにくいことが好ましい。そこで、複製再配置時のネットワークトポロジにおいて、2 連結成分 (biconnected component)

をグループとする。2連結成分とは、1つの頂点を除いてもグラフが連結されている（分断されない）極大な部分グラフのことである。DCG方式では、再配置周期ごとに、次のアルゴリズムに基づいて複製の再配置を行う。

1. E-DAFN方式と同様に、再配置周期がくると、各移動体は自身の識別子をブロードキャストする。これにより、自身と相互接続されている移動体を知る。
2. 識別子  $M_i$  の添字  $i$  が最も小さい移動体から、ネットワークの2連結成分を発見するアルゴリズムを実行する。その結果発見された2連結成分をグループとする。2つ以上の2連結成分に含まれる移動体は、最初に発見された2連結成分のグループのみに含まれるものとする。
3. グループとしての各データのPT値を、グループを構成する移動体からのPT値の合計として計算する。
4. グループとしてのPT値が高いデータから順に、その複製のグループ内での配置を決定する。ただし、グループ内の移動体のいずれかがオリジナルデータとして所持しているものについては配置を行わない。グループ内での配置先は、空き領域がある移動体のうちで、そのデータのPT値が最も高いものとする。
5. すべての種類のデータ（複製）の配置を決定した後、グループ内の移動体に空き領域がある場合は、空き領域がなくなるまで、グループとしてのPT値が高いデータから順にさらに複製の配置を決定する。グループ内での配置先は、そのデータ（複製）を所持しておらず、かつ、空き領域のある移動体のうちで、そのデータのPT値が最も高いものとする。

この方式では、分断されにくい移動体のグループにおいて、多くの種類のデータ（複製）を共有するため、アクセス成功率が向上するものと考えられる。ただし、再配置周期ごとにグループ内での情報交換や複製の再配置が行われるため、その他の2方式に比べてオーバーヘッドが大きい。

### 3 シミュレーション評価

本章では、本稿で提案した複製配置方式の性能評価のために行ったシミュレーション評価の結果を示す。

50×50の二次元平面上に、40個の移動体 ( $M = \{M_1, \dots, M_{40}\}$ ) が存在する。各移動体は、すべての方向に等確率に、0から3の範囲でランダムに決定した速度で移動する。各移動体の無線通信範囲は、半径7の円とする。ネットワーク内に、40種類のデータ ( $D = \{D_1, \dots, D_{40}\}$ ) が存在し、 $D_i$  は  $M_i$  にオリジナルデータとして保持されている。各移動体は、最大で10個の複製を作成する。各移動体の  $D_i$  へのアクセス頻度は、 $p_i = 0.5(1 + 0.01i)$  とする。これは、すべての移動体が同一のアクセス特性をもち、各データへのアクセス頻度の差が小さい場合を表している。各方式では、再配置周期200（単位時間）で複製の再配置が行われる。データの更新周期を  $T$  とし、 $D_{i-1}$  の更新の  $T/40$  後に  $D_i$  の更新が発生するものとする。シミュレーション実験では、 $T$  の値を変化させて、各方式のデータアクセス成功率とトラフィックを比較する。ここでは、トラフィックを複製再配置のための通信のホップ数と定義する。

初期位置として各移動体をランダムに配置し、50,000単位時間を経過させたときのシミュレーション結果を図1および図2に示す。図では、横軸が更新周期  $T$ 、縦軸が各方式のアクセス成功率およびトラフィックをそれぞれ表している。図1の結果から、E-DCG方式、E-DAFN方式の順で高いアクセス成功率を示している。これは、E-SAF方式において複製の重複が非常に多いことから、E-DAFN方式とE-DCG方式によって成功率が向上するためである。更新周期が長くなるにつれて、各方式のアクセス確率は高くなり、その性能差も大きくなっていく。一方、図2の結果から、トラフィックもE-DCG方

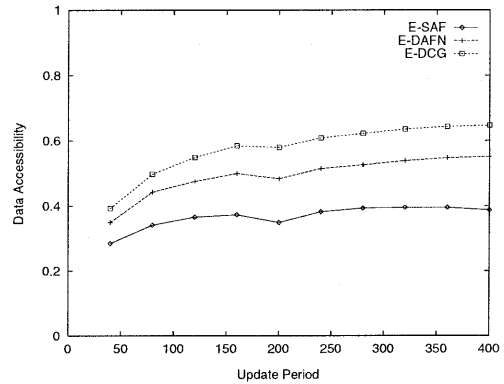


図1: 更新周期とアクセス成功率

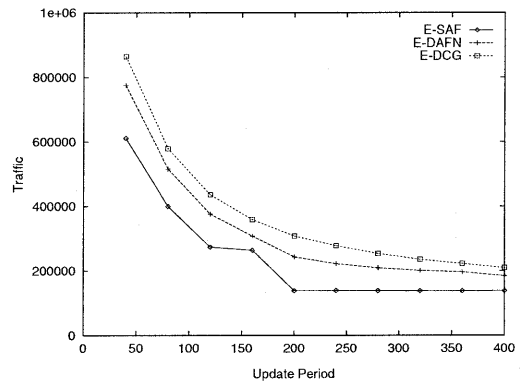


図2: 更新周期とトラフィック

式、E-DCG方式の順で大きくなっている。更新周期が長くなると、各方式のトラフィックは小さくなっている。

### 4 おわりに

本稿では、データの更新が定期的にかかるアドホックネットワークを想定して、データの利用率向上のために複製を配置する方式を提案した。性能評価のために行ったシミュレーション実験の結果から、E-DCG方式のアクセス成功率が最も高くなることを確認した。また、トラフィックに関しては、E-SAF方式が最も低くなることを確認した。

今後は、データの更新が不定期に発生する環境を想定し、有効な複製配置方式を検討する予定である。

謝辞 本研究は、日本学術振興会未来開拓学術研究推進事業における研究プロジェクト「マルチメディア・コンテンツの高次処理の研究」(Project No. JSPS-RFTF97P00501)の研究助成によるものである。ここに記して謝意を表す。

### 参考文献

- [1] D.J. Baker, J. Wieselthier, and A. Ephremides, "A distributed algorithm for scheduling the activation of links in a self-organizing, mobile, radio network," Proc. IEEE ICC'82, pp.2F6.1-2F6.5, 1982.
- [2] J. Broch, D.A. Maltz, D.B. Johnson, Y.C. Hu, and J. Jetcheva, "A performance comparison of multi-hop wireless ad hoc network routing protocols," Proc. Mobicom'98, pp.159-164, 1992.
- [3] 原 隆浩, 西尾 章治郎: "アドホックネットワークにおけるアクセス可能性向上のためのデータ複製配置方式," 情報処理学会マルチメディア, 分散, 協調とモバイルシンポジウム論文集, pp.7-12, 2000.