

P2P による分散管理型地理情報システム

伊藤 雅亮 ・ 大沢 裕

A Distributed Geographic Information System on P2P Environment

Masaaki ITO, Yutaka OHSAWA

Abstract Sensor network is one of next generation ubiquitous networks and it is organized as an ad-hoc network. In the ad-hoc network, DHT (distributed hash table) is used for load balancing, and CAN (content-addressable network) is also used for concrete implementation. However, CAN is a kind of hash, then it does not preserve spatial locality. This paper proposes a new load balancing method suitable sensor network, which keeps spatial locality.

Keywords: センサーネットワーク (Sensor Network)、P2P、DHT、範囲検索 (Range Search)

1. はじめに

センサーネットワーク (I.F.Akyildiz et al., 2002) は次世代のユビキタスネットワークの一つである。これは、様々な場所に設置したセンサーをネットワークで繋げることにより、ユーザが遠隔地や広範囲の情報を自由に取得、利用できる環境である。大量のセンサーから継続的に情報を収集することで高精度・高信頼性な情報が得られる。また、それを利用することで、観測対象の予測や検知を実現できる。例えば、森林や海にセンサーを配置することで、森林火災や洪水等の自然災害、動植物の生息状況を知ることができる。

このような予測や検知を実現するには、大量・高密度にセンサーを配置する必要がある。そして、そのためには小型・低コストの自律動作型センサーが必要になる。しかし、それにより個々のセンサーは必然的に低電力、小容量となってしまう、

非常に脆弱となってしまう。その一方で、短距離無線通信、位置情報管理・検出機能が必要となるため、それらを兼ね合わせたセンサーが求められる。このような特徴から、従来のサーバ・クライアント型の環境ではなく、無線アドホックネットワーク環境で実現されることが多い。

しかし、ネットワークトポロジーが頻繁に変更される場合も多い。更に、個々を識別するグローバルな ID が無い、位置情報の絡んだ検索といった従来では実現困難な点が存在する。そのため、センサーネットワークの実現には新しい環境が必要であり、その例として LL-NET (金子ほか, 2005) がある。そこで、本稿ではこのようなセンサーネットワークを実現できる新しい環境を提案する。本稿で提案する環境では DHT (Distributed Hash Table) に着目した。

DHT は、ハッシュテーブルを複数のピアで管理することで、従来の P2P 環境の問題点を大幅に解消した Pure-P2P に分類される P2P 環境である。さらに、共有するデータのハッシュ値を論理空間上にマッピングし、その論理空間を複数のピアで分

伊藤：〒338-8570 埼玉県さいたま市桜区

下大久保 255 埼玉大学大学院理工学研究科,

TEL (FAX) : 048-858-9181

email: s07mm303@mail.saitama-u.ac.jp

散管理することにより，サーバを用いずに位置決定を可能にした．また，これは高速な完全一致検索も実現した．従って，DHT は十分な負荷分散，スケーラビリティ，頑健性を提供できるといえる．しかし，DHT は位置情報を考慮しないため，位置と関連した情報の扱いと範囲検索が頻繁に発生するセンサーネットワークに単純に適用できない．

本稿では，センサーネットワークを対象とした位置情報管理可能な新しい構造を提案する．そのために DHT の CAN の構造をベースに，GBD 木における領域式の手法を取り入れた．

2. 提案手法

本稿で対象にするセンサーネットワークには，実際にデータを収集するセンサーノードと，それを管理しデータを処理するシンクノードが存在する．センサーネットワークでは多数のセンサーノードが高密度に配置されている．従って，これらを複数のグループに分け，各グループにつき 1 つのシンクノードが管理する．ユーザはシンクノードにアクセスすることで欲しい情報を取得する．

これを踏まえた上で本手法の条件は，CAN のように負荷分散でき，なおかつ範囲検索を効率良く行えることである．そこで，CAN をベースに領域式を融合する．具体的には，CAN の論理空間を物理空間に置き換え，データをゾーンと MBR で管理する．

本章では CAN と領域式について簡単に述べた後，本手法のポイントと範囲検索について述べる．

2.1. CAN と領域式

本手法のベースである CAN (S. Ratnasamy et al., 2001) は DHT の構築法である．CAN は d 次元トラスの d 次元デカルト座標空間 U を用いて構築される．この空間 U は領域(ゾーン)に動的に分割され，各ゾーンを 1 つのピアが管理する．また，データは(キー, 値)のペアで扱われ，キーを空間上の点 P にマッピングすることにより，このデータが管理されるピアを決定する．図 1 は 5 個のピアで分割される二次元空間の例である．

CAN では空間局所性に対処できないため，本手法では CAN に GBD 木 (大沢・坂内, 1991) の領域式と MBR を適用する．MBR とはあるデータを内包する最もタイトな長方形であり，領域式とは「分割座標軸を巡回的に替えながら面積二等分割を繰り返して得られる長方形」の位置と大きさを表すのに用いる 0, 1 のビット列である．

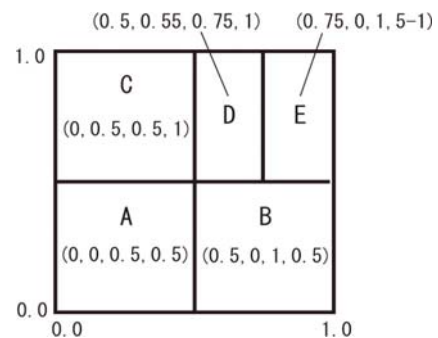


図 1: CAN の空間 U

2.2. ゾーンと MBR

CAN はハッシュを用いて負荷 (データ量) を一様にしたが，本手法では空間局所性を維持するためにハッシュを用いない．そのため，全空間で見ると非常に偏りのある分布となるが，領域式に従うゾーンの分割を行うことで，ノード単位で見れば十分負荷分散された状態にできる．つまり，これに MBR を導入すれば，範囲検索効率の良い負荷分散環境を提供できる．

本手法では CAN のデータをセンサーノード，ピアをシンクノードとして扱う．つまり，シンクノードは自身の領域として MBR とゾーンを持ち，その領域内のセンサーノードを管理する．従って，本手法はセンサーネットワークだけでなく一般の GIS にも適用可能である．さらに，CAN の構造をベースにしているため，CAN で用いる各種操作，拡張の多くをそのまま適用できる．

図 2 は，5 つのシンクノードで管理されるセンサーネットワークの例である．本手法と CAN の相違点は，ゾーンのサイズが疎らになる点と複数のゾーン (白色矩形)，MBR (灰色矩形) を管理する状況 (濃灰色部分) が発生する点である．

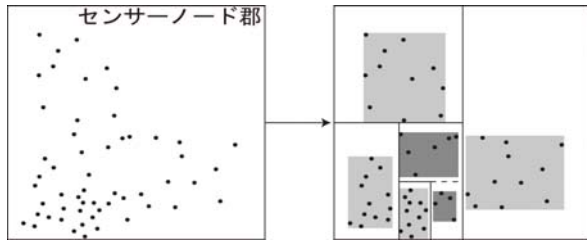


図 2：シンクノード 5 個のセンサーネットワーク

2. 3. シンクノードとセンサーノード

ある段階でシンクノード N_{new} を新たに設置する状況が発生した時、あるシンクノード N_{split} の領域を分割し、一方を N_{new} に割り当てる。この際の領域分割は、まずGBD木のように N_{split} のゾーンを領域二等分割する。そして、センサーノード数の多い方が $\lfloor 2 \times \text{総センサーノード数} / 3 \rfloor$ 個以下になるまで領域を二等分割する。その後、複数の断片ゾーンの内、最後に分割された断片ゾーンを N_{new} の、残りを N_{split} のゾーンとし、各々に適切なMBRを設定する。もし、 N_{split} のゾーンが複数ならば、ゾーンの分割をせずにセンサーノード数が同等になるよう 2 つのグループに分ける。

N_{new} の設置場所は、センサーノード数 K 個以上でゾーン数最多であるシンクノードのゾーン内である。 K は、自身と自身がリンクを張るシンクノードの管理するセンサーノード数の平均である。

また、シンクノード N_{del} が故障、破棄された場合はCANのゾーン結合法と同様の事を行う。つまり、 N_{del} の管理するセンサーノード群を、一意に決定する任意のシンクノードが代わりに管理する。

センサーノードの新規追加と故障は、MBR の拡大縮小量が最少であるシンクノードに挿入、削除することで行う。

3. 範囲検索

まず、ユーザがクエリを出したシンクノードから、検索範囲の中心点をゾーンに内包するシンクノードまでCANのルーティングを行う。そして検索に移り、各シンクノードが以下を行うことで実現する。

- ・ 検索結果と転送したメッセージ数をユーザに伝達
- ・ 転送元シンクノード以外で検索範囲と MBR がオーバーラップするリンクノードにメッセージを転送

転送したメッセージ数を結果報告に付与するのは検索の終了を知るためであり、「受信メッセージ数 - (転送メッセージ数 + 1) = 0」で検索終了となる。

4. 実験

実験により、本手法のルーティング性能をCANと比較し、範囲検索性能を調べた。使用したデータは以下であり、シンクノード数を 8000 個まで増やしながら各実験を行った。

使用したデータは、10 万個の点からなる一様分布データと正規分布データの 2 つである。正規分布データは、以下を 10 回繰り返すことで作成した。
Step1: 平均 μ が 0、分散 σ が 1 の正規分布データを作成

Step2: 各点の座標値が指定範囲内の整数値になるようにランダム量シフト

また範囲検索性能を調べるために各点データに対して矩形範囲を 1000 個ずつ以下のように作成した。

Step1: x_s, y_s, x_e を全空間の中からランダムに選択

Step2: 面積が全空間の 5% となるように y_e を選択

Step3: $(x_s, y_s), (x_e, y_e)$ を 1 つの検索範囲とする

4. 1. ルーティング性能

本手法とCANのルーティング性能を比較した。図 3 は、1000 回ルーティングを行い、ルーティング時に辿るシンクノード数の平均をプロットしたグラフである。これは、値が小さければ少ないノード数で目的ノードに辿れるため、ルーティングが効率的に行われていることを表す。

図より、全体的にルーティングにかかるノード数が約 6~8% 減っている。本手法のゾーンサイズはCANと同等以下であり、リンク数は本手法の方が少なくなる。従って、シンクノードのゾーン形状が不定になり、CANのルーティングで複数に渡

ったゾーンが1つに纏まった結果だと考えられる。

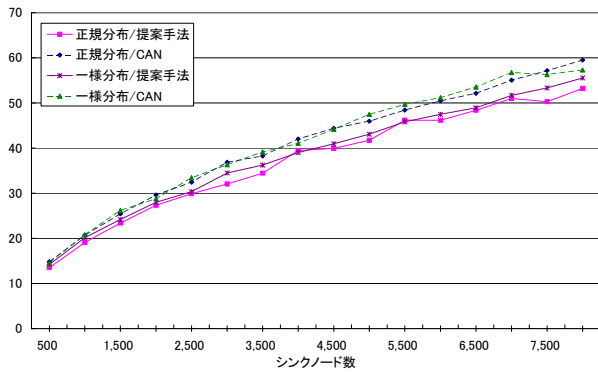


図 3：ルーティング性能

4. 2. 範囲検索性能

事前に生成した検索範囲を用いて範囲検索にかかるシンクノードのホップ数とメッセージ数を調べた (図 4)．の差が小さければ，範囲検索が効率的に実現できることを表す．ここで，ホップ数は辿らなければならないシンクノードの数であり，メッセージ数は結果を得るために調べたシンクノードの数である．

図より点データの粗密が大きければ，ホップ数，メッセージ数共に少ない数で範囲検索を実現できることが分かる．また，メッセージ数とホップ数の差が小さいことが分かる．

前者は，偏りのあるデータ分布に対して一様に配置された検索範囲を適用したためだと考えられる．つまり，正規分布では各ピアの持つゾーンサイズが異なっており，点の密度が薄い部分のみの検索も実行しているため平均したときに各数が減少していると考えられる．つまり，もし密度の濃い部分を重点的に調べるよう矩形を生成した場合，結果が大きく変わる可能性がある．

後者は，正規分布で設定される MBR が一様分布よりタイトになる (MBR の平均サイズが 10%程度減少) ために，探索ノードの選択が効率化されたからだと考えられる．これは，正規分布では粗密により管理 MBR 数が増え，一様分布より小さいサイズの MBR を複数管理する形になるためである．

以上より，均一的なデータよりも偏ったデータを対象にした範囲検索の方が，効率的に実現でき

ることが分かった．

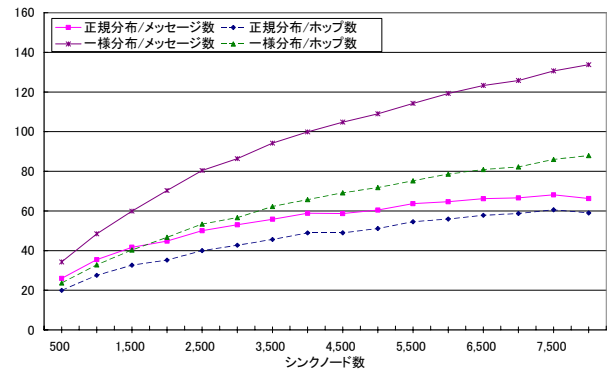


図 4：範囲検索性能

5. まとめ

本稿ではセンサーネットワークのための新しい環境を提案した．この環境は負荷分散と範囲検索を両立できる．そのために，CAN の構造に GBD 木の葉ノードの分割を導入した．また，実験により負荷を CAN と同程度に分散できること，CAN よりも効率的にルーティングできること，性能は粗密のあるデータの方が向上することを示した．

参考文献

- 大沢裕・坂内正夫 (1991) 種類の補助情報により検索と管理性能の向上を図った多次元データ構造の提案, 「電子情報通信学会論文誌」, 第 J74 巻, pp. 467-475.
- 金子雄・春本要・福本真哉・下條真司・西尾章治郎 (2005) ユビキタス環境における端末の位置情報に基づく P2P ネットワーク, 「情報処理学会論文誌」, 第 46 巻, pp. 1-15.
- I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, (2002) Wireless sensor networks: a survey, Computer Networks, Vol. 38, pp. 393-422.
- Sylvia Ratnasamy, Paul Francis, Mark Handley, Richard Karp, Scott Shenker, (2001) A scalable content-addressable network, Proceedings of the ACM SIGCOMM'01, pp. 161-172.