

# 街角センサネットワークのデータ収集・管理システムにおける蓄積能力の平均化機構について

田中 孝浩<sup>†</sup> 中尾 太郎<sup>††</sup> 塚本 昌彦<sup>†</sup> 西尾章治郎<sup>†</sup>

<sup>†</sup> 大阪大学大学院情報科学研究科 〒 565-0871 大阪府吹田市山田丘 1-5

<sup>††</sup> 大阪大学大学院工学研究科 〒 565-0871 大阪府吹田市山田丘 1-5

E-mail: †{tanaka,tuka,nishio}@ist.osaka-u.ac.jp, ††taro@ise.eng.osaka-u.ac.jp

あらまし 本研究では、街角のいたる所に広範囲にわたってセンサを配置し、人の動きをセンシングする街角センサネットワークから発生するデータを、限られたストレージでなるべく長期的に収集・管理することを目的とする。そのために、発生する膨大なデータを分散ストレージで蓄積し、その一部を圧縮して長期保存するシステムを設計した。提案システムは、分散ストレージ間で自律的にデータ保持能力を均一化し、安定したデータへのアクセシビリティを提供する。本研究では、シミュレーションによるシステムの性能評価を行い、提案システムが時間帯や曜日による人の流れの変化といった外乱に適切に対応し、そのデータ保持能力を確保しつづけることを確認した。

キーワード センサネットワーク、分散ストレージ、動的負荷分散、ネットワーク帯域消費量

## A Balancing Mechanism of Data Storing Capacity in Data Collection and Management System for Ubiquitous Sensor Network

Takahiro TANAKA<sup>†</sup>, Taro NAKAO<sup>††</sup>, Masahiko TSUKAMOTO<sup>†</sup>, and Shojiro NISHIO<sup>†</sup>

<sup>†</sup> Grad. Sch. of Info. Science and Tech., Osaka Univ. Yamadaoka 1-5, Suita, Osaka, 565-0871 Japan

<sup>††</sup> Grad. Sch. of Eng., Osaka Univ. Yamadaoka 1-5, Suita, Osaka, 565-0871 Japan

E-mail: †{tanaka,tuka,nishio}@ist.osaka-u.ac.jp, ††taro@ise.eng.osaka-u.ac.jp

**Abstract** A ubiquitous sensor network collects the information of people's movement using the sensor devices which are located widely in a city. It is important for such a network to collect and manage data as long as possible with limited storage. We have designed a distributed storage system, which is capable of handling large amount of data from the sensor network. It has the function to balance the capability among multiple distributed storages to provide the stable data accessibility. In this paper, we describe the design of the system. Our simulation results show that the system is capable of adapting the disturbance of the environment, such as an occasional change of people's movement due to the spatio-temporal deviation, and as a result of sustaining data accessibility.

**Key words** sensor network, distributed storage, dynamic load balance, network bandwidth consumption

### 1. はじめに

近年、多くの小型センサを用いてネットワークを形成し、広範囲にわたるデータ収集を行うセンサネットワークが、今後のユビキタスコンピューティングの基盤として注目を集めている。

センサネットワークは、データ収集の対象となる空間領域の広さによって、システムの設計やその用途が異なる。例えば、部屋や建物の内部などの温度、湿度、照度や人の動きを検知して自動的に照明やテレビの電源を調節し、室内の気温や照度を自動制御するホームセンサネットワークでは、センサは屋内に固定配置され、結線される [3] [8]。

一方、ガスセンサを用いた大気汚染物質濃度の検出や、農作物や家畜に被害を与える自然状況の監視といった、広範囲からデータを収集するセンサネットワーク [10] [14] では、広範囲にわたるセンシングを低コストで実現するために、使い捨てのセンサを無線通信でアドホックにつなぐアプローチをとることが多い。このセンサネットワークを長期間にわたって動作させるため、センサの消費電力を抑え、データ通信量を削減する研究が盛んに行われている [1] [4]。その他にも、センサとサーバとの通信における片方向リンクを考慮した経路制御 [9] や、階層的にデータを集約し、少ないオーバーヘッドでクエリを配信し、データを取得するもの [12]、センサの位置情報の管理を効率化

するもの[2]など、無線センサネットワークの技術課題を解決するためにさまざまな研究が行われている。

本研究では、無線ではなく、有線で結ばれたセンサを都市部の電信柱、店舗、交差点など街角のいたる所に配置して、広範囲のデータ、特に人の流れのセンシングを行う街角センサネットワークを対象とする。有線によるセンサネットワークの構築には、センサノード以外にも配線などに多額の費用が必要となるが、無線を用いたセンサネットワークに生じるような技術課題は回避できる。都市のインフラストラクチャとして期待される街角センサネットワークでは、無線を用いることによるコスト削減よりも、センサネットワークを長期的に安定して動作させることが重要である場合が多い。その結果構築される大規模な街角センサネットワークは、局所的な人の動きの詳細な調査や、広域的な人の流れの傾向の導出を可能にする。

大規模な街角センサネットワークで考慮しなければならない問題に、広範囲に高密度で設置された多数のセンサから次々に発生する膨大なデータ量がある。筆者らはこれまでに、データストレージを多段化し、街角センサネットワークのアプリケーション利用を可能にするシステムを提案した[13]。提案システムは、通行人数の変動によってデータの発生頻度に偏りが生じて、ストレージがデータを保持する期間を自律的に調整することで、アプリケーションの設計を容易にしている。

本稿では、ストレージ間の容量差やデータの発生パターンの偏りといった、街角センサネットワークのためのデータ収集・管理システムを実環境で運用する上で考慮しなければならない条件下でも、提案システムが安定して性能を発揮できるようにする機構について述べる。提案機構の性能はシミュレーションによって確認する。

以降、2章では街角センサネットワークの特徴と提案システムの概要について述べ、3章で提案システムの機能と明らかにすべき課題について述べる。4章ではシステムの性能を評価するために行ったシミュレーションについて説明し、5章にその結果を示す。6章でシミュレーションの結果を考察し、最後に7章で本稿をまとめる。

## 2. 街角センサネットワークのためのデータ収集・管理システム

### 2.1 街角センサネットワークの特徴

本研究では、街角センサネットワークから発生するデータの利用形態として、以下の2つを想定している。

センサが発生する高精細なデータへのアクセス 刻々と変化する最新の人の流れの情報を知りたい場合や、商店街における人々の店から店への詳細な移動経路を調べる場合など、センサが発生する時間的・空間的に最も精細なデータへのアクセス  
長期にわたって蓄積されたデータへのアクセス 人の流れや分布状況の傾向を出店計画や都市計画の立案に利用するなど、長期にわたって蓄積した広域的なデータを解析する際のデータアクセス

この2つのデータ利用形態を両立するためには、センサから

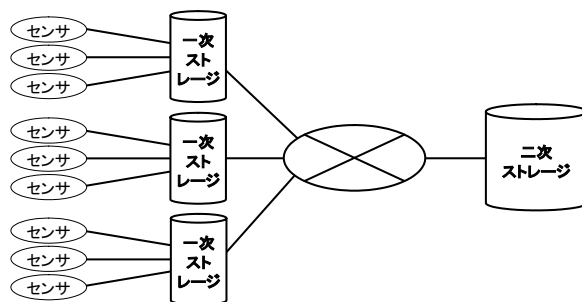


図1 DSUSNの概略

発生するデータを長期的に保存すればよいが、例えばセンサを数mから数十mの間隔で数十km<sup>2</sup>から数百km<sup>2</sup>という都市レベルの広さにわたって配置すると、生じるデータ量はテラバイト/週のオーダーとなるなど、大規模街角センサネットワークから発生するデータ量は膨大なものとなり、現実的に全てのデータを発生したまま長期間にわたって保持することはできない。また、街角センサネットワークのセンシングエリアの広さはセンサやデータの管理も困難にする。街角センサネットワークのデータを収集・管理するシステムには、これらの事項を考慮した設計が求められる。

### 2.2 DSUSN

2.1節の要求事項にもとづき、筆者らは街角センサネットワークのためのデータ蓄積システム(DSUSN: Data Storing system for Urban Sensor Network)を設計した。DSUSNは以下の特徴をもつ。

ストレージの多段化 街角センサネットワークから発生する膨大なデータへの直接のアクセスと、長期的なデータへのアクセスの両者を可能にするため、発生したセンサデータをそのまま蓄積する一次ストレージと、一次ストレージが保持できなくなったデータをカリングなどの手法によって数%程度にまで大幅に圧縮した上で蓄積する二次ストレージを設ける。この不可逆圧縮の過程でデータの大部分は失われる。

一次ストレージの分散配置 センサとセンサから発生するデータの管理を容易にするため、センシングエリア内に複数の一次ストレージを分散して配置し、近隣のセンサとそこから発生するデータを管理する。端末を分散配置して特定地域のデータを効率的に管理するため、データの可用性が向上する。センサのメンテナンスなどのシステム管理コストも削減できる。

データ滞留時間の均一化 人の流れをセンシングする街角センサネットワークでは、人通りの多さがセンサのデータ発生頻度に密接に関係する。人通りの多さが場所や時間帯によって変動すれば、発生するデータ量も変動する。DSUSNでは、一次ストレージにおけるデータ滞留時間を均一化し、アプリケーションの設計を容易にする。データ滞留時間を均一化するメカニズムの詳細は次節で述べる。

DSUSNでは、各センサは地理的に最も近い一次ストレージと有線で接続されており、発生したデータを一次ストレージへ送信する。データを長期的に蓄積する二次ストレージは、ネッ

トワークで一次ストレージとつながっている．DSUSN の概略を図 1 に示す．

### 2.3 DSUSN のデータ滞留時間均一化機構

一次ストレージにデータが滞留する期間は，データ発生頻度の変化に応じて一次ストレージ間でセンサ担当を動的に変更することで均一化する．具体的には，一次ストレージに空きを作る際に以下のアルゴリズム [13] を適用する．アルゴリズムは，ストレージの使用率が閾値  $th_{max}$  を超えた場合に適用する．以下に，一次ストレージ  $S_s$  の使用率が  $th_{max}$  を超えた場合を例としてアルゴリズムの動作を示す．

(1) 近隣の一次ストレージの集合  $S_{neighbors}$  から， $S_s$  より古いデータを保持している一次ストレージの集合  $S_{candidates}$  を選ぶ． $S_{candidates} = \phi$  ならば (6) に進む．

(2)  $S_s$  と  $S_{candidates}$  の中から，単位時間当たりのデータ発生数  $gen$  が最も少ない一次ストレージ  $S_d$  を選ぶ． $S_s = S_d$  ならば (6) に進む．なお，一次ストレージ  $S_i$  の単位時間当たりのデータ発生数  $gen_i$  は， $S_i$  が保持しているデータ数を  $num_i$ ，現在時刻を  $t$ ， $S_i$  が保持している最も古いデータの発生時刻を  $t_{oldest_i}$  として，以下の式によって求める．

$$gen_i = \frac{num_i}{t - t_{oldest_i}}$$

(3)  $S_s$ ， $S_d$  間でセンサの担当を変更することによって，各一次ストレージの  $gen$  の格差が縮まるかどうかを予測する． $gen_s$  を  $S_s$  の， $gen_d$  を  $S_d$  の単位時間当たりのデータ発生数として，センサ担当変更後の格差  $diff$

$$diff = |(gen_s - g) - (gen_d + g)|$$

ただし  $g = gen_s \times \frac{th_{max} - th_{min}}{th_{max}}$

が  $diff < gen_s - gen_d$  ならば格差が縮まると判断し，(4) に進む，そうでなければ (6) に進む．

(4) 格差が縮まると予測される場合は， $S_s$  が担当しているセンサのうち， $S_d$  に地理的に近いセンサの管理担当を  $S_d$  に変更する．このとき，センサが今までに発生したデータも同時に  $S_d$  に移動し，ストレージに空きを作っていく．

(5) (4) の操作を繰り返し， $S_s$  の使用率が  $th_{min}$  未満となればアルゴリズムの実行を終了する．

(6)  $S_s$  の使用率が  $th_{min}$  未満となるまで，古いデータから順に，保持しているデータのサイズを縮小した上で二次ストレージに移動し，アルゴリズムの実行を終了する．

このアルゴリズムによって，一次ストレージにおけるデータ滞留時間を十分に均一化できること，また，データ発生頻度分布がめまぐるしく変化する環境下においても，その均一化能力が有効であることをシミュレーションによって確認している [13]．

### 2.4 一次ストレージ間の性能格差の考慮

地理的に分散して配置され，ローカルに管理される DSUSN の一次ストレージでは，性能増強やメンテナンスが各一次ストレージの管理者によって独立に行われ，その結果，一次スト

レージ間に性能格差が生じることが考えられる．そこで [13] のデータ滞留時間の均一化アルゴリズムに一部変更を加える．2.3 節の (2) を以下のように変更する．

(2)  $S_s$  と  $S_{candidates}$  の中から，単位時間単位ストレージ容量当たりのデータ発生数  $gen$  が最も少ない一次ストレージ  $S_d$  を選ぶ． $S_s = S_d$  ならば (6) に進む．なお，一次ストレージ  $S_i$  の単位時間単位容量当たりのデータ発生数  $gen_i$  は， $S_i$  の容量を  $storage_i$  として，以下の式によって求める．

$$gen_i = \frac{num_i}{(t - t_{oldest_i}) \times storage_i}$$

以降，このデータ滞留時間均一化手法を DSR(dynamic sensor re-allocation) 法と呼ぶ．

## 3. DSUSN の機能と検証すべき課題

DSUSN では，以下の事項が明らかにすべき課題となっている．

### 3.1 ネットワーク帯域消費量

DSUSN は，DSR 法で一次ストレージのデータ滞留時間を均一化する際に，今までに発生したデータ系列を一次ストレージ間で転送し，ネットワークの帯域を消費する．

ネットワークの帯域を消費する妥当性を明らかにするために，センサ担当のみを変更し，その際のデータ系列の移動は行わない手法との比較検証が必要である．この比較を通して，ネットワーク帯域消費量とデータ滞留時間の均一化能力との間の関係を明らかにする．

### 3.2 データ発生パターン

実際の街では，駅や学校などの地理的条件や，昼と夜，あるいは平日と休日といった時間帯などによるデータ発生頻度分布の変化に何らかの傾向が見られることが多い．こうした場合，データの発生パターンの傾向を見こしてデータ収集・管理システムをあらかじめチューニングしておくことで，帯域を消費せずにデータ滞留時間をある程度均一化する手法の有効性が浮上する．環境の変化に対してシステムを自律的に最適化する DSUSN と，分散配置する一次ストレージの容量をあらかじめ設定してデータ保持能力を高めようとする手法との比較検証が必要である．

## 4. シミュレーション

シミュレーションでは，分散配置した一次ストレージにセンサから発生したデータが保持され，データに安定してアクセスができるかどうかを，データ滞留時間の均一化能力によって評価する．まず，データの発生に地理的条件や時間帯による傾向がある環境下で，データ発生傾向に基づいてシステムパラメータをチューニングしておいたデータ保持システムと，環境の変化に自律的かつ動的に対応する DSUSN との比較を行う．また，データ滞留時間の均一化能力とネットワーク帯域消費量との関係を明らかにするために，DSUSN の帯域消費量を計測する．その際，DSR 法の帯域消費量をより少なくするための一つのアプローチとして，DSR 法においてデータ系列の移動を行わないものを用意し，DSR 法と比較する．

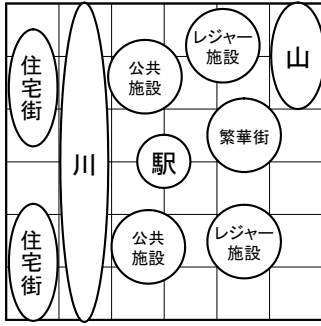


図2 シミュレーション領域のイメージ

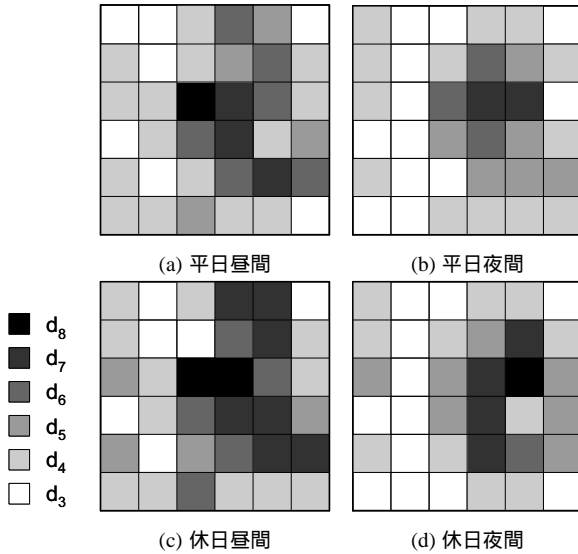


図3 データ発生パターン

#### 4.1 環境設定

数  $m$  から数十  $m$  間隔にセンサを配置した数  $km$  四方のフィールドを想定し、正方形の仮想領域を 36 の小領域に分割し、センサを格子状に 90,000 個配置したシミュレーション領域を用意した。仮想領域には一次ストレージを 36 個格子状に配置した。実際の街では、センサや一次ストレージの配置には地理的な偏りが生じるだろうが、DSUSN ではセンサの配置状況ではなく、一次ストレージにおける単位時間当たりのデータ発生数が結果に影響を与えるので、本シミュレーションでは配置パターンを考慮する必要はない。

各小領域には 10 段階のデータ発生頻度  $d_1, \dots, d_{10}$  を設定する。データ発生頻度  $d_n$  は、初項  $d_1 = 10$ 、公差  $\frac{d_c - d_1}{9}$  の等差数列とした。以降、 $d_c (= d_{10})$  を基準データ発生数と呼ぶ。

データ発生頻度の分布は平日昼間、平日夜間、休日昼間、休日夜間の 4 パターンを作成した。シミュレーション領域には、図 2 に示すように、駅やレジャー施設、住宅街などの人工建造物や、山、川といった自然環境を想定し、データの発生頻度を設定した。作成した 4 つのデータの発生パターンを図 3 に示す。図のように各小領域には  $d_3, \dots, d_8$  の発生頻度が設定されている。駅は常時、レジャー施設は休日の昼間、繁華街は夜間にデータ発生頻度が高くなるようにし、住宅街や公共施設は全体的にデータ発生頻度を低めに設定した。山と川には常にほとん

ど人がいないものとした。この発生パターンから、5%の確率で 2 段階、10%の確率で 1 段階、発生頻度レベルをランダムに変化させる。

シミュレーションでは、PC 程度の一次ストレージ端末がカメラ画像を扱うことを想定して、1 つのデータの大きさは数十 KB、一次ストレージの容量は数十 GB 程度とした。すなわち、一次ストレージの容量は 250,000 から 2,000,000 の範囲でランダムに定めた。これは、40KB のデータに対して、一次ストレージの容量が 10GB から 80GB あることに相当する。シミュレーションは、以上の環境下で 12 時間ごとに昼夜を繰り返す平日 2 日、休日 2 日、平日 3 日の 1 週間とした。

#### 4.2 比較手法

3.1 節の課題に対して DSR 法の性能の特徴を明らかにするために、センサ担当のみを変更し、その際そのセンサからこれまでに発生したデータを移動しない手法として DSRwithoutDS (DSR without data sequence) 法を用意した。sensor, num をそれぞれ  $S_s$  が現在担当しているセンサ数と保持しているデータ数、ratio をデータ移動量が移動元の一次ストレージ容量に占める割合 ( $th_{max} - th_{min}$ ) としたとき、withoutDS 法は、 $S_s$  の使用率が  $th_{max}$  を超えた場合、 $sensor \times ratio$  個のセンサの担当変更を行った後、センサ担当変更の有無に関わらず、 $S_s$  の使用率が  $th_{min}$  未満となるまでストレージ内のデータを古い順に二次ストレージに移動する。

また、DSR 法におけるネットワーク帯域消費量の理論的な最小値を得るための OPT 法を用意した。OPT 法は、ある与えられたデータ発生頻度パターンに対し、各一次ストレージが自分に有線で接続されたセンサを理論的に可能な限り多く担当した場合を示す。OPT 法におけるデータを格納する際の一次ストレージ間のデータ転送量が、DSR 法のアプローチにおけるデータ転送量の理論上の下限値となる。

さらに、3.2 節の課題に対して DSR 法の特徴を明らかにするために、一次ストレージの容量をあらかじめ設定しておき、データの移動を全く行わない手法として FIX 法を用意した。FIX 法では、センサの担当変更をせず、 $S_s$  の使用率が  $th_{min}$  未満となるまで  $num \times ratio$  個のデータを二次ストレージに移動する。ある領域を担当する一次ストレージの容量  $S$  は、その小領域における平日昼間、平日夜間、休日昼間、休日夜間のデータ発生頻度をそれぞれ  $d_{wd}, d_{wn}, d_{hd}, d_{hn}$  として、以下の式によって算出する。

$$S = \frac{\{(d_{wd} + d_{wn}) \times 5\} + \{(d_{hd} + d_{hn}) \times 2\}}{7} \times const$$

const の値は、一次ストレージの総容量が DSR 法の場合と等しくなるように設定した。図 3 のデータ発生パターンから得られた、FIX 法における一次ストレージ容量の分布を図 4 に示す。

#### 4.3 実験

以上のような環境設定の下で、各一次ストレージにおける保持データ数、保持している最も古いデータの発生時刻、単位時間単位ストレージ容量当たりのデータ発生数の 3 項目を 50 単位時間ごとに計測した。1 単位時間は 1 分に相当する。また、ネットワークを介して移動されたデータの総数も計測した。な

0.51	0.42	0.49	0.72	0.67	0.42
0.56	0.42	0.54	0.77	0.83	0.56
0.6	0.49	0.96	1.0	0.93	0.53
0.42	0.49	0.77	0.93	0.67	0.65
0.58	0.42	0.53	0.81	0.86	0.79
0.49	0.49	0.63	0.56	0.56	0.49

図4 FIX 法における一次ストレージ容量比

お、ネットワーク上をデータが移動するのは以下の場合である。

(1) センサ担当変更によって、センサが有線で接続されていない、地理的に離れた一次ストレージの管理下にあるセンサから発生したデータを、そのセンサを管理している一次ストレージに送信する際

(2) センサ担当を変更する際に、ストレージ間でデータを移動する際

二次ストレージへデータを移動する際にもネットワーク上をデータが流れるが、その量はいずれの手法も同じなので、今回のシミュレーションでは省略している。

## 5. シミュレーション結果

### 5.1 データ滞留時間

$ratio$ (データ移動量が移動元の一次ストレージに占める割合( $th_{max} - th_{min}$ ))を0.02, 0.05, 0.1としたそれぞれの場合において、50単位時間ごとに測定した各一次ストレージにおけるデータ滞留時間の標準偏差を算出し、その平均をとったものを図5, 6, 7に示す。横軸は $d_c$ である。 $d_c$ が大きいほど、データの発生に時間的・空間的な偏りが大きいことを示す。標準偏差の平均値は、小さいほどデータ滞留時間の平均からの誤差が小さく、一次ストレージにいつアクセスを行っても滞留時間の格差が小さく抑えられていることを示す。

データ発生頻度分布の傾向を考慮してシステムをあらかじめチューニングし、データの移動を行わないFIX法よりも、環境の変化に応じて自律的にデータ滞留時間を均一化するDSR法が良い結果を示している。ただし、withoutDS法の場合は一次ストレージ間でデータ保持能力を適切に均一化できていない。

また、DSR法、withoutDS法では、 $d_c$ の値が大きいほどデータ滞留時間を均一化できていることがわかる。さらに、DSR法においては $ratio$ の値が小さいほど均一化能力が高いことが示されている。これは、均一化のためのセンサの担当変更が頻繁に行われるためだと考えられる。

### 5.2 一次ストレージの活用効率

一次ストレージ全体で保持している総データ量の平均が、一次ストレージの総ストレージ容量に占める割合を一次ストレージの活用効率として、 $ratio$ を0.02, 0.05, 0.1としたそれぞれ

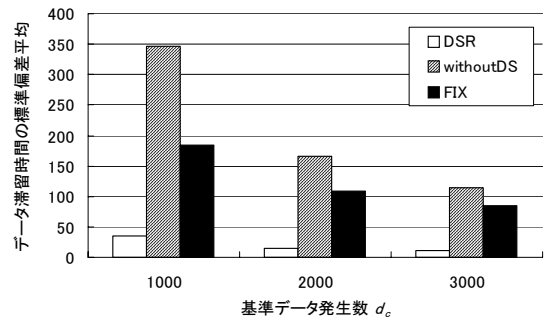


図5 データ滞留時間の標準偏差平均 ( $ratio = 0.02$ )

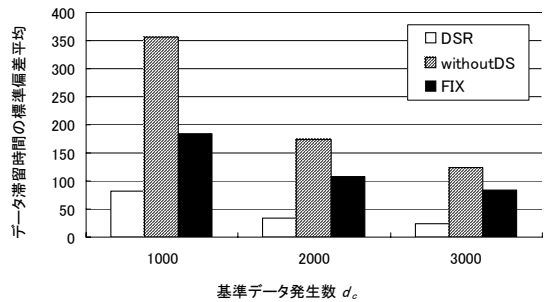


図6 データ滞留時間の標準偏差平均 ( $ratio = 0.05$ )

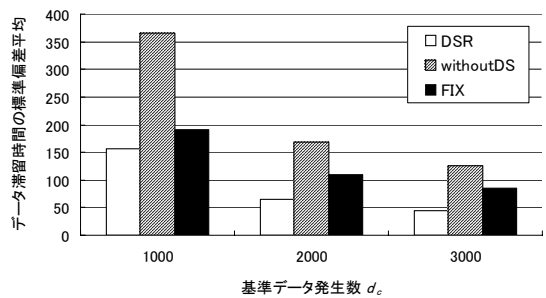


図7 データ滞留時間の標準偏差平均 ( $ratio = 0.1$ )

の場合について、それぞれ図8, 9, 10に示す。横軸は $d_c$ である。なお、本シミュレーションにおいて、一次ストレージ活用効率の最大値は $th_{max} = 0.9$ である。

$d_c = 2,000, 3,000$ の場合、各手法における一次ストレージ活用効率にほとんど差はみられないが、 $d_c = 1,000$ の場合、withoutDS法の活用効率が約1.2%から1.4%低下している。これは、一次ストレージ全体で保持できるデータ量が約18GB減少することに相当する。これは、5.1節で示されたように、withoutDS法の均一化能力が他の手法に比べて著しく低いためだと考えられる。

一方、DSR法、FIX法の活用効率が $ratio, d_c$ の値に関わらずほぼ等しいことから、データ発生頻度分布の変化に応じた動的なセンサ担当の変更と一次ストレージ活用効率の間には相関関係は存在しないと考えられる。

### 5.3 帯域消費量

$ratio$ を0.02, 0.05, 0.1としたそれぞれの場合における、一次ストレージ間での総データ移動数を図11, 12, 13に示す。横軸は $d_c$ である。なお、 $ratio = 0.02$ のとき、withoutDS法が消費する帯域は、 $d_c = 1,000$ で約38.7Mbps、 $d_c = 2,000$ で

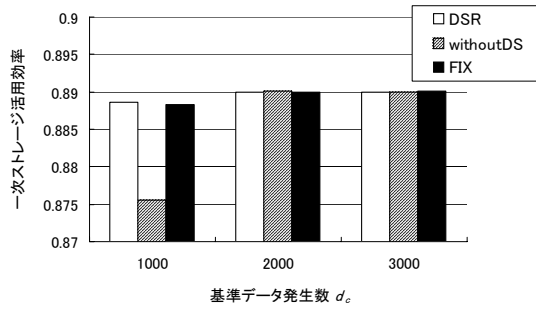


図 8 一次ストレージの活用効率 ( $ratio = 0.02$ )

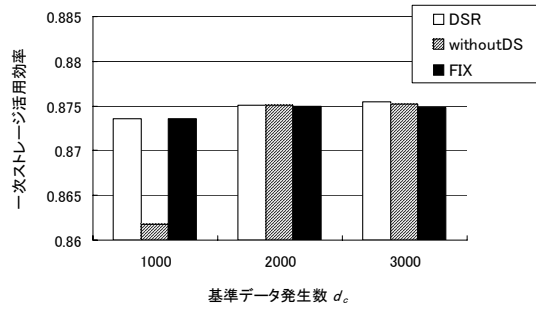


図 9 一次ストレージの活用効率 ( $ratio = 0.05$ )

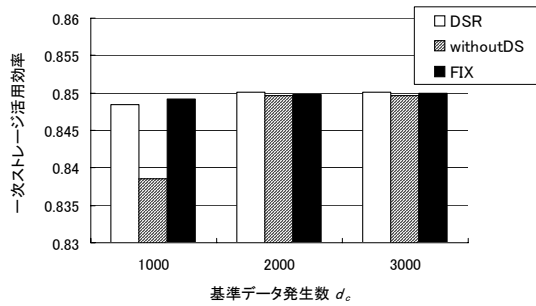


図 10 一次ストレージの活用効率 ( $ratio = 0.1$ )

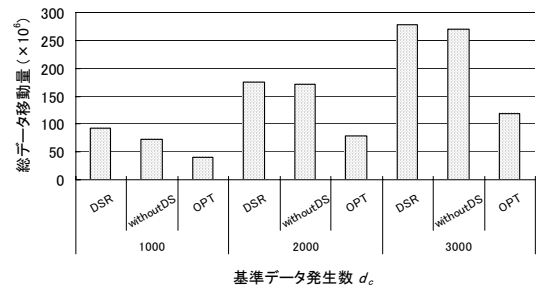


図 11 データ移動数 ( $ratio = 0.02$ )

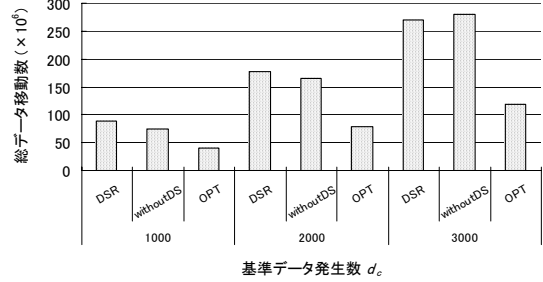


図 12 データ移動数 ( $ratio = 0.05$ )

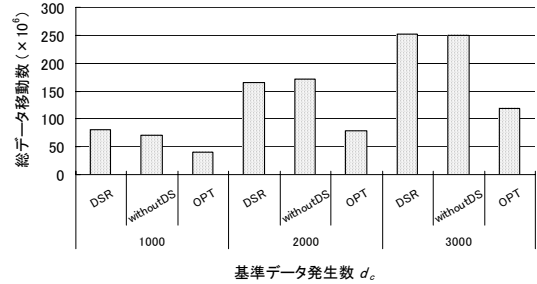


図 13 データ移動数 ( $ratio = 0.1$ )

約 91.0Mbps,  $d_c = 3,000$  で約 142.9Mbps に相当する。一方,  $ratio = 0.02$  のとき DSR 法が消費する帯域は,  $d_c = 1,000$  で約 49.5Mbps,  $d_c = 2,000$  で約 92.5Mbps,  $d_c = 3,000$  で約 146.6Mbps に相当する。

ほぼ全ての場合において, DSR 法に比べ withoutDS 法が帯域の消費を抑えているが, OPT 法と比較するとその削減量はわずかなものであり, 図 5, 6, 7 に示された withoutDS 法の低い均一化能力とその帯域消費削減量は見合っているとはいえない。また, withoutDS 法ではセンサ担当を変更する際にデータ系列を移動しないため, 特定地域のデータが複数の一次ストレージに分散して格納されることになり, DSUSN に対するデータアクセス性能の低下につながると考えられる。つまり, センサ担当を変更する際にデータ系列を移動しないことによるメリットはほとんどないといえる。

OPT 法は, DSR 法におけるネットワーク帯域消費量の理論的な下限値を示すものであるが, あらかじめデータ発生頻度分布がわかっていなければセンサ担当の最適配置を決定できないため, DSR 法の帯域消費量を下限値まで削減することは現実的に不可能である。DSR 法において帯域消費量をさらに削減する手法についての考察は 6 章で行う。

## 6. 考 察

大規模な街角センサネットワークが情報社会のインフラストラクチャとして有することになる公共性を考えれば, データ保持機構はそこから発生するデータをさまざまなアプリケーションで利用できるように設計するべきである。街角センサネットワークを汎用利用するために, そのデータ保持機構はセンサから発生する精細なデータをより多く, より安定して保持できる必要がある。

DSUSN は, 従来の SAN(storage area network) [6] [7] などの分散ストレージシステムと違い, ストレージ容量に比べて圧倒的に発生するデータ量が大きい環境を想定している。そうした環境では, 従来の分散ストレージシステムにおける I/O のバランシングといった負荷分散メカニズム [5] [11] に加えて, ストレージが溢れることを前提とした DSR 法のような新しい負荷分散戦略が必要となる。

DSUSN は, データの発生源であるセンサの担当を分散したストレージ間で自律的に変更することでストレージ容量を有効活用し, データの発生頻度を予想しにくい街角センサネットワークからのデータを収集・管理する。センサの担当を分散ストレージ間で動的に変更するため, 各ストレージは自分と有線

で接続されていないセンサを担当する機会が多くなり、センサからのデータをネットワークを介してストレージに格納する必要がでてくる。5章に示したように、DSR法とwithoutDS法の帯域消費量にあまり差が認められないことから、両手法とOPT法におけるネットワーク帯域消費量の格差は、主にこのデータ転送によるものと考えられる。

図11, 12, 13に示されたように、DSR法は帯域を消費してしまうため、DSUSNを実環境に適用するためには、これをさらに削減するための手法の検討が必要となる。各ストレージが、自分と有線で接続されているセンサをできるだけ多く担当できるように、例えば、駅など、データ発生頻度が高いと予想される地域にあらかじめ一次ストレージを密に配置したり、センサを複数の一次ストレージと有線で接続しておくといったようなシステムのチューニングが考えられる。

一方、DSUSNを利用するアプリケーションによっては、こうしたデータ滞留時間の均一化をそれほど必要としないものもある。また、DSUSNを適用する環境下で利用できるネットワーク帯域が限られている場合も考えられる。そこで、状況に応じてDSR法の均一化能力をチューニングし、ネットワーク帯域消費量とのバランスをとる機能をDSUSNに付加するアプローチが有効だと考えられる。

## 7. おわりに

本研究では、街角センサネットワークから発生するデータを適切に収集・管理するシステムDSUSNを設計した。DSUSNでは、ストレージを多段化して、街角センサネットワークから発生する高精細なデータへのアクセスと、長期的に蓄積されたデータへのアクセスを両立させる。DSUSNは、データ発生頻度の偏りに応じてDSR法によって動的にストレージ間でセンサの担当を変更し、ストレージにデータが滞留する時間を均一化できる。そのため、ストレージにどれくらい前までのデータが残っているのかを予想しやすく、街角センサネットワークのアプリケーションを設計しやすくなっている。

本研究ではシミュレーション実験を行い、DSUSNがDSR法によって環境の変化に自律的に対応し、ストレージにデータが滞留する時間を均一化できることを確認した。

今後の課題として、よくアクセスされるデータは一次ストレージに残しておくなど、二次ストレージに移動する際のデータの選別におけるチューニングを可能にし、DSUSNのデータアクセス性能をさらに向上する手法の検討や、センサから発生したデータを格納する際の一次ストレージ間でのデータ転送量を抑えてネットワーク帯域消費量を削減するため、データ発生頻度の分布の傾向を考慮して最適な一次ストレージの配置場所を決定する手法の検討などが挙げられる。

## 謝 辞

本研究は、文部科学省21世紀COEプログラム「ネットワーク共生環境を築く情報技術の創出」および文部科学省特定研究領域(C)「Grid技術を適応した新しい研究手法とデータ管理技術の研究」(プロジェクト番号:13224059)によっている。ここに

記して謝意を表す。また、研究を進める上で貴重なご意見を頂いたサイバーメディアセンターの秋山豊和助手、小川剛史助手に深く感謝する。

## 文 献

- [1] H. Gupta, S. R. Das, and Q. Gu, "Connected Sensor Cover: Self-Organization of Sensor Networks for Efficient Query Execution," Proc. 4th Annual ACM International Symposium on Mobile Ad Hoc Networking and Computing(MOBIHOC2003), Annapolis, Maryland, pp.189-200, June 2003.
- [2] 廣井幹也, 南正輝, 森川博之, 青山友紀, "高密度センサネットワークにおける統計的位置管理手法," 情報処理学会マルチメディア, 分散, 協調とモバイル(DICOMO2002)シンポジウム, pp.237-240, June 2002.
- [3] 楠本晶彦, 中澤仁, 戸辺義人, 徳田英幸, "A Location-Adaptive Virtual Networked Appliance," 日本ソフトウェア科学会, Systems for Programming and Applications(SPA2001), March 2001.
- [4] S. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, "The Design of an Acquisitional Query Processor For Sensor Networks," Proc. Special Interest Group on Management of Data(SIGMOD2003), San Diego, CA, pp.491-502, June 2003.
- [5] F. Matthijs, Y. Berbers, and P. Verbaeten, "A Flexible I/O Framework for Parallel and Distributed Systems," Proc. 4th International Workshop on Object-Orientation in Operating Systems(IWOOS1995), Lund, Sweden, pp.187-190, August 1995.
- [6] X. Molero, F. Silla, V. Santonja, and J. Duato, "Modeling and Simulation of Storage Area Networks," Proc. 8th International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems(MASCOTS2000), IEEE, San Francisco, pp.307-314, August 2000.
- [7] X. Molero, F. Silla, V. Santonja, and J. Duato, "On the Interconnection Topology for Storage Area Networks," Proc. 15th International Parallel and Distributed Processing Symposium(IPDPS2001), IEEE, San Francisco, pp.162, April 2001.
- [8] 森川博之, 南正輝, 青山友紀, "ユビキタスネットワークへの道," 情報処理学会誌, vol. 43, no. 6, pp.631-638, June 2002.
- [9] 永原崇範, 今井尚樹, 國頭吾郎, 森川博之, 青山友紀, 篠田庄司, "センサネットワークのための二段階ルート構築手法," 電子情報通信学会総合大会, March 2001.
- [10] T. Nakamoto, Y. Suzuki, and T. Moriizumi, "Study of VHF-band QCM gas sensor," Sensors and Actuators B, Chemical, vol. 84, Issues 2-3, pp.98-105, May 2002.
- [11] X. Qin, H. Jiang, Y. Zhu, and D. R. Swanson, "A Dynamic Load Balancing Scheme for I/O-Intensive Applications in Distributed Systems," Proc. 32nd International Conference on Parallel Processing(ICPP2003), Kaohsiung, Taiwan, pp.79-86, October 2003.
- [12] 関根理敏, 瀬崎薫, "センサネットワークにおける階層的データ集約法," 電子情報通信学会ソサイエティ大会, pp.252, September 2003.
- [13] 田中孝浩, 中尾太郎, 塚本昌彦, 西尾章治郎, "大規模なユビキタスセンサネットワークにおけるデータ収集・管理システムの設計," 情報技術レターズ, 情報科学技術フォーラム(FIT2003), pp.325-326, September 2003.
- [14] B. Warneke, M. Last, B. Liebowitz, and K. S. J. Pister, "Smart Dust: Communicating with a Cubic-Millimeter Computer," Computer Magazine, IEEE, Piscataway, NJ, pp.44-51, January 2001.