

大規模な街角センサネットワークにおける データ収集・管理システム的设计

田中孝浩^{†1} 中尾太郎^{†2} 小川剛史^{†3}
塚本昌彦^{†4} 西尾章治郎^{†1}

本研究では、街角のいたるところ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・管理し、データへの安定したアクセシビリティを提供するシステム的设计を目的とする。そのために、提案システムでは、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アプローチをとる。また、分散ストレージ間で自律的にデータ蓄積能力を均一化する機構により、システムに対するデータアクセス性能を向上させる。本研究では、シミュレーションによる性能評価を行い、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、そのデータアクセス性能を確保しつづけることを確認した。

A Design of a Data Collection and Management System for Large Scale Ubiquitous Sensor Networks

TAKAHIRO TANAKA,^{†1} TARO NAKAO,^{†2} TAKEFUMI OGAWA,^{†3}
MASAHIKO TSUKAMOTO^{†4} and SHOJIRO NISHIO^{†1}

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, and it is need for such a network to provide the stable data accessibility. 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. 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.

1. はじめに

近年、電子機器の小型化・低価格化にともない、多くの小型センサを用いてネットワークを形成し、広範囲からのデータ収集を行うセンサネットワークが、今後のコピ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、システム的设计やその用途が異なる。たとえば、部屋や建物の内部などの温度、湿度、照度、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たり、室内の気温や照度を自動制御したり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^{4),10)}。

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目的で、広範囲からデータを収集するセンサネットワーク^{12),21)}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、センサの位置情報の効率的な管理³⁾や、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

†1 大阪大学大学院情報科学研究科
Graduate School of Information Science and Technology, Osaka University

†2 株式会社 NTT データ
NTT Data Corporation

†3 大阪大学サイバーメディアセンター
Cybermedia Center, Osaka University

†4 神戸大学工学部電気電子工学科
Department of Electrical and Electronics Engineering,
Kobe University

経路制御¹¹⁾、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することによるデータの取得¹⁶⁾など、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さかんに行われている^{1),5)}。また、センサネットワークから得られるデータを効率的に利用するためのデータ分散手法¹⁴⁾も提案されている。

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ところに配置して、広範囲からのデータ収集を行う街角センサネットワークを対象とし、特に人の流れのセンシングを行うことを想定する。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により、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が可能となる。

街角センサネットワークで人の流れをセンシングする場合には、データ取得間隔が短いほど、より詳細な人の動きが調査できるが、取得するデータ量が膨大になり、データを長期間蓄積することが困難となってしまふ。一方、データ取得間隔を長くすると、データ量を少なく抑えることができるが、詳細な人の動きに関する調査などが困難となる。

本研究では、街角センサネットワークにおいて、取得された精細なデータを長期間蓄積し、アプリケーションから地理的単位での効率的なアクセスを実現するデータ収集・管理システムを提案する。本システムでは、人がセンサの前を通過したときのみデータを取得することで、詳細なデータを必要最小限取得するものとし、分散配置したストレージに取得したデータを蓄積する。十分なディスクスペースが確保できなくなると、古いデータを圧縮して他のストレージに退避させることで新しいデータの蓄積を継続する。

センサネットワークから発生するデータはセンサの前を通過する人の量に依存しているため、センシングする場所や時間帯によって発生頻度が大きく異なる。そのため、データ発生頻度の高いセンサを多く担当するストレージでは短い周期でデータを退避させなければならず、アプリケーションが必要とするデータにアクセスできなくなる問題が生じる。本システムでは、場所や時間帯によって変動するデータ発生頻度の偏り

を考慮して、各センサから得られるデータがストレージに滞留する時間を自律的に調整する機能を用いて、データへのアクセス性能を高める。

本システムを実環境で運用するために、ストレージ間の容量差やデータ発生パターンの偏りの時間的な変化といった条件下でも、提案システムが安定して性能を発揮できることを確認するため、シミュレーションによる性能評価を行う。

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章で本論文をまとめる。

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。

センサが生成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、商店街のような局所的な地域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、センサが生成する時間的・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させるためには、センサから発生するデータを長期的に保存しなければならないが、たとえばセンサを数mから十数mの間隔で数十km²から数百km²という都市レベルの広さにわたって配置すると、生じるデータの量は1週間でテラバイトオーダーとなるなど、街角センサネットワークから発生するデータの量は膨大なものとなり、すべてのデータをサイズを縮小せずに長期間にわたって保持することは現実的に不可能である。これは、ストレージを多段化し、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非可逆圧縮によって数%程度にまで大幅に圧縮したうえで長期的に蓄積する二次ス

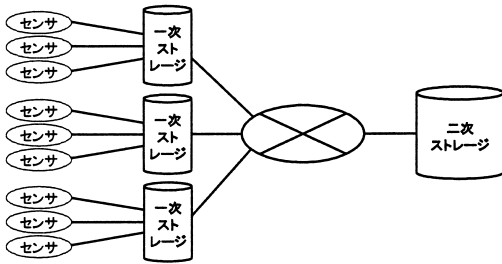


図 1 DSUSN の概略
Fig.1 Outline of DSUSN.

トレージを設けることで解決する。

2.2 DSUSN

本研究で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、広大な街角センサネットワークのセンシングエリアにおける効率的なセンサやデータの管理の実現などを目的として、以下の特徴を持つ。

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うえ、センサや一次ストレージのメンテナンスといった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、発生するデータの量も変動する。DSUSN では、一次ストレージにデータが滞留する時間 (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は 2.4 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と各一次ストレージはネットワークでつながっている。DSUSN の概略を図 1 に示す。

2.3 データへのアクセス

DSUSN では、システムに対して発生するデータアクセスを処理するために受付サーバを用意する。受付

サーバによって、DSUSN を利用するアプリケーションはシステムの内部構成を特に意識することなく必要なデータを取得できる。アプリケーション側は、たとえば駅を中心とした半径 1 km 以内の地域における昨日 1 日のデータといったように、地理的、時間的な条件をもとに DSUSN に対してクエリを発行する。受付サーバは一次ストレージのセンサ担当を示すインデックスを保持しており、与えられた条件に基づいて、適切な一次ストレージ、または二次ストレージにアクセスし、データを取得する。

センサネットワークからのデータの取得に関しては、アプリケーションからのクエリに応じてデータを提供するセンサデータエージェントの設計^{17),20)} や、センサネットワークを利用するマルチエージェントシステムのためのプラットフォームの実装²²⁾ に関する研究が行われているほか、ユビキタスなインターネット環境におけるネームサービスを中心としたインターネットアプリケーションプラットフォーム構築⁷⁾ についての研究が行われている。こうした研究の成果は、DSUSN における受付サーバの構築に利用できると考えられる。

2.4 DSUSN のデータ滞留時間均一化機構

データ滞留時間は、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。ただし、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ため、センサ担当を決定する際は一次ストレージの容量差も考慮する。具体的には、一次ストレージの使用率が閾値 th_{max} を超えた場合は、ストレージに空きを作るために以下のアルゴリズムを適用する。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、単位時間単位ストレージ容量あ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あ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 $t_{oldest-i}$, S_i の容量を $storage_i$ とし、以下の式によって求める。

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

- (3) S_s, S_d 間でセンサ担当を変更することによって、 S_s, S_d 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 の単位時間単位容量あ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) に進み、そうでなければ (6) に進む。

- (4) S_s が担当しているセンサのうち、 S_d に地理的に最も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る。
- (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 の使用率が閾値 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大幅に縮小したうえ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。

以降、このデータ滞留時間均一化手法を DSR (dynamic sensor relocation) 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 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って、データ発生頻度分布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越してデータ収集・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。そこで、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討を行う。

3.2 ネットワーク帯域消費量

DSUSN では、DSR 法でセンサ担当を自律的に変

更することによって一次ストレージのデータ滞留時間を均一化するが、センサから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する際、センサから発生したデータを有線で接続されていない一次ストレージへネットワーク経由で格納する際に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討を行う。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。

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 1 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。

4.1 環境設定

10m 間隔にセンサを配置した 1.5 km 四方のフィールドを想定して、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 22,500 個配置したシミュレーション領域を用意した。DSUSN では、システム導入コストや 1 つの一次ストレージあたりに発生するデータの量を考慮した場合、特定地域のデータの効率的な管理を目的として分散配置される一次ストレージは、都市部の最小行政単位に 1 つ程度の割合で配置されることになる予想される。本シミュレーションでは、シミュレーション領域の面積を考慮し、仮想領域に 16 個の一次ストレージを格子状に配置した。

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であろうが、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あたりのデータ発生数が結果に影響を与えるため、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}

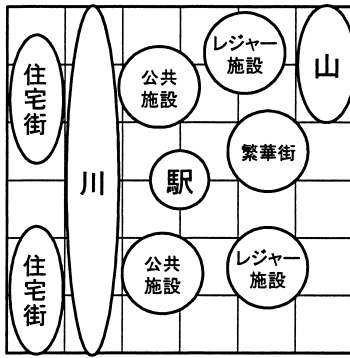


図2 シミュレーション領域のイメージ
Fig. 2 Image of simulation area.

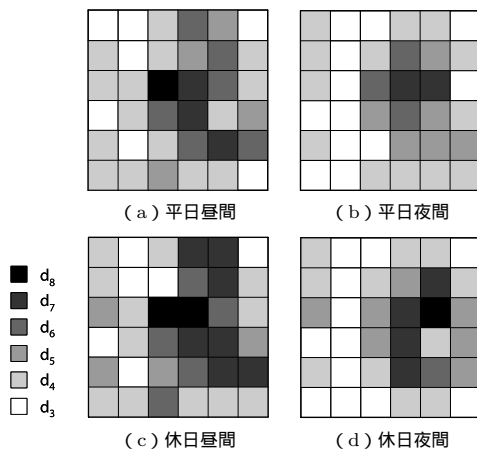


図3 データ発生パターン
Fig. 3 Data generation patterns.

を設定する．データ発生頻度 d_n は、小領域から単位時間あたりに発生するデータ数の平均値である．本シミュレーションでは、 d_{10} を基準データ発生数 d_c と定義し、 d_n を初項 $d_1 = 10$ 、公差 $\frac{d_c - d_1}{9}$ の等差数列とした．シミュレーションは、 d_c の値を 2500、3125、3750 として行った．これは、最もデータ発生頻度の高い小領域 ($d_n = d_c$) に 500 人から 750 人程度の人間が存在し、それぞれが 1 単位時間あたり約 5 回センシングされるという想定に基づいている．なお、本シミュレーションにおいて 1 単位時間は 1 分としている．

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設定し、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．山と川にはつねにほとんど人がいないものとした．実際の街では、同じ時間帯であっても、日によってデータ発生頻度分布に多少の違いがあるため、作成した発生パターンから 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．

本研究では、街角センサネットワークの利用例として画像データの収集を想定し、一次ストレージとして、コストなどの観点から一般的な PC 程度の計算機を想定した．シミュレーションでは、1 つのデータのサイズを、カメラ画像から人物追跡を行う研究^{2),19)} で一般に用いられている画像を参考に数十 KB とした．また、1 つの一次ストレージの容量はデータ 200 万個分から 1,000 万個分の範囲でランダムに定めた．これは、50 KB のデータに対して、一次ストレージの容量が 100 GB から 500 GB 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 週間として行った．

4.2 比較手法

3.1 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、データの移動をまったく行わない手法として FIX 法を用意した． $sensor, num$ をそれぞれ一次ストレージ 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、FIX 法では、センサ担当の変更を行わず、 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、 S 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、 $d_W = d_{wd} + d_{wn}$ 、 $d_H = d_{hd} + d_{hn}$ と定め、以下の式によって算出する．

$$S = \frac{(d_W \times 5) + (d_H \times 2)}{7} \times const$$

$const$ 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．

さらに、3.2 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系列を移動しない手法として DSRwithoutDS (DSR without data sequence) 法を用意した．withoutDS 法は、

0.54	0.52	0.80	0.60
0.60	0.77	1.0	0.72
0.52	0.67	0.92	0.78
0.55	0.61	0.71	0.67

図 4 FIX 法における一次ストレージ容量比

Fig. 4 Ratio of first storages' capacity of FIX method.

S_s の使用率が th_{max} を超えた場合, $sensor \times ratio$ 個のセンサ担当の変更を行った後, センサ担当変更の有無にかかわらず, S_s の使用率が th_{min} 未滿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に OPT 法を用意した. OPT 法では, あるデータ発生頻度パターンが与えられた際, 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するようにただちにセンサ担当を変更し, データ系列の移動は行わない. 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.

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単位時間単位ストレージ容量あたりのデータ発生数, 保持している最も古いデータの発生時刻の 2 項目を 50 単位時間ごとに計測した. 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. なお, ネットワーク上をデータが移動するのは以下のいずれかの場合である.

- センサ担当変更によって, 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場合
- センサ担当を変更する際に, 今までそのセンサから発生したデータ系列を一次ストレージ間で移動する場合

二次ストレージへデータを移動する場合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.

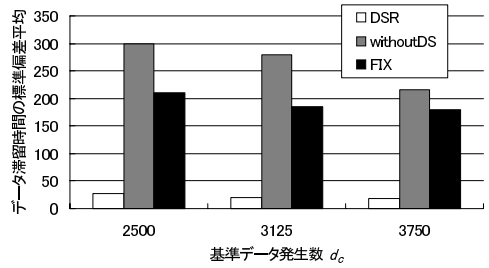


図 5 データ滞留時間の標準偏差平均 (ratio = 0.02)

Fig. 5 Average of standard deviation of data staying time.

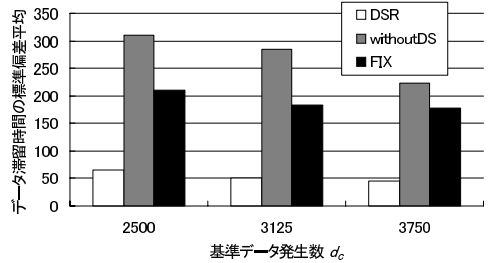


図 6 データ滞留時間の標準偏差平均 (ratio = 0.05)

Fig. 6 Average of standard deviation of data staying time.

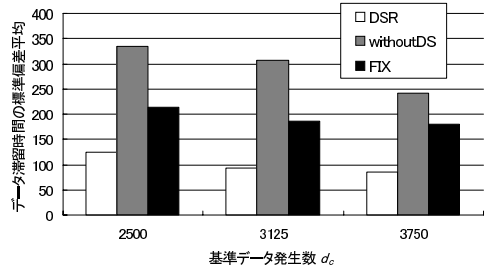


図 7 データ滞留時間の標準偏差平均 (ratio = 0.1)

Fig. 7 Average of standard deviation of data staying time.

5. シミュレーション結果

5.1 データ滞留時間

$ratio$ 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データ滞留時間 (現在時刻 - 最も古いデータの発生時刻) から標準偏差を算出し, その平均をとったものを図 5, 図 6, 図 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.

$d_c = 3,750$ で $ratio$ を 0.02, 0.05, 0.1 としたそれぞれの場合において, 各一次ストレージの単位時間単位容量あたりのデータ発生数から算出した標準偏差

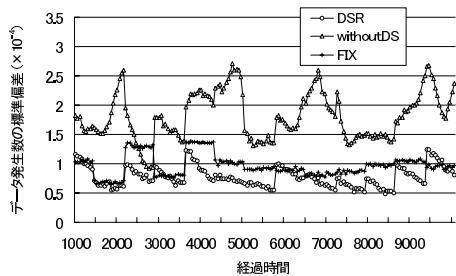


図 8 データ発生数の標準偏差 ($ratio = 0.02$)

Fig. 8 Standard deviation of the number of generated data.

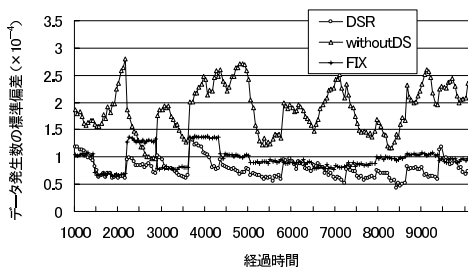


図 9 データ発生数の標準偏差 ($ratio = 0.05$)

Fig. 9 Standard deviation of the number of generated data.

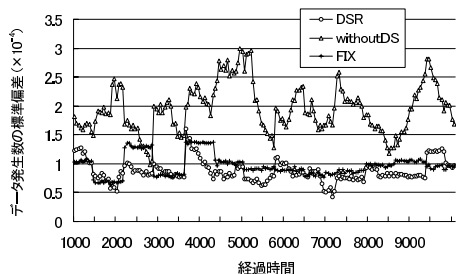


図 10 データ発生数の標準偏差 ($ratio = 0.1$)

Fig. 10 Standard deviation of the number of generated data.

の推移を図 8, 図 9, 図 10 に示す. 横軸は経過時間である. この値が小さいほど, 一次ストレージ間で単位時間単位容量あたりのデータ発生数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動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担当の変更が頻繁に行われるため

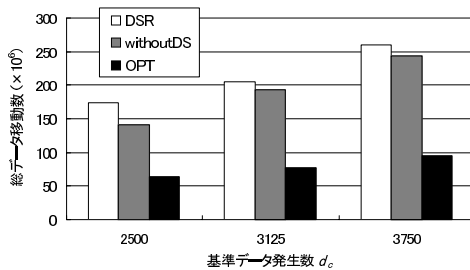


図 11 データ移動数 ($ratio = 0.02$)

Fig. 11 The number of moved data.

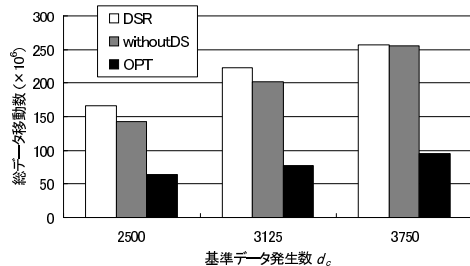


図 12 データ移動数 ($ratio = 0.05$)

Fig. 12 The number of moved data.

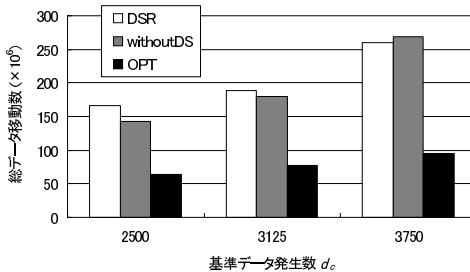


図 13 データ移動数 ($ratio = 0.1$)

Fig. 13 The number of moved data.

あると考えられる.

5.2 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, 図 12, 図 13 に示す. 横軸は d_c である. DSR 法, withoutDS 法における総データ移動数が, 実環境でどれくらいのネットワーク帯域消費量に相当するかを表 1 に示す.

ほとんどすべての場合において, 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, 図 6, 図 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,

表 1 ネットワーク帯域消費量 (Mbps)
Table 1 The amount of consumption of bandwidth.

			d_c		
			2500	3125	3750
ratio	0.02	DSR	114.7	135.3	172.1
		withoutDS	93.8	127.3	160.9
	0.05	DSR	109.8	147.1	169.8
		withoutDS	94.6	133.3	168.7
	0.1	DSR	110.2	124.5	172.2
		withoutDS	94.5	118.8	178.1

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分かっているならば最適なセンサ担当の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のセンサネットワーク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 では、ネットワーク障害によるデータ送信時のエラーについて考慮していないが、システムを安定して長期的に動作させるためには何らかの対策をとる必要がある。グローバル規模の分散ストレージを対象とした研究に、分散ストレージのプラットフォームを提供する OceanStore¹⁵⁾ や、一般的な分散アプリケーションの構築のための基盤を提供する Tapestry²³⁾ があるが、これらのシステムでは、実際のネットワーク環境下におけるメッセージ送信エラーなどに対する高い耐障害性を有しているため、こうした研究に用いられている技術は DSUSN におけるネットワーク経由でのデータ送信に利用できると考えられる。

DSUSN は、従来の SAN (storage area network)^{8),9)} などの分散ストレージシステムや、定期的に収集した数バイト程度の降雨データを巨大な 1 つのストレージに蓄積する東京アメッシュ¹⁸⁾ などの広

域センサネットワークと異なり、広域にわたって非常に密に設置されたセンサから不定期かつ頻繁にデータが発生し、かつ 1 つのデータサイズが数十 KB 程度という、ストレージ容量やネットワーク帯域に比べて発生するデータの量が圧倒的に大きな環境を想定している。そうした環境では、ストレージにデータを蓄積できる期間が短くなるため、必要とするデータにアクセスできなくなったり、データをストレージに送信する際に十分なネットワーク帯域を確保できなくなったりするといった問題が生じる。したがって、データの送信や蓄積において、SAN や東京アメッシュなどで用いられている手法をそのまま DSUSN に適用することは難しい。そのため、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^{6),13)} に加えて、ストレージが溢れることを前提とした、DSR 法のようなデータの格納に対する新しい負荷分散戦略や、センサとストレージを有線で接続するといったアプローチが必要となる。

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。DSR 法では、センサ担当を変更する際、ストレージの担当するセンサのうち、それが自分と有線で接続されているかどうかにかかわらず、相手ストレージに地理的に最も近いものから順に担当を変更する。そのため、OPT 法に比べ、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る。すなわち、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ためにネットワークの帯域を消費する機会が多くなっていると考えられる。5.2 節に示したように、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、DSR 法におけるネットワーク帯域の消費は、主にこのデータ転送に起因するものと考えられる。

本シミュレーションのように、画像データなど、数十 KB 程度のデータを収集することを想定した場合は、表 1 に示されたように DSR 法によって帯域を消費してしまう。都市部における区程度の地域に街角センサネットワークを構築する場合など、センシングエリアがシミュレーション領域の数倍の広さとなれば、ネットワーク帯域消費量はさらに増加することになる。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。各一次ストレージが、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、たとえば、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いたりするといったようなシステムのチューニングが考えられる。

DSR 法におけるネットワーク帯域消費量は、ストレージと有線で接続されていないセンサから優先的にセンサ担当を変更するようにアルゴリズムを改良することで削減できると考えられる。ただし、このアプローチでは、発生したデータを地域ごとに管理できなくなるため、データアクセス性能の低下につながる恐れがあるほか、地理的なデータ発生頻度の偏りを考慮したセンサ担当の変更が行われなくなるため、データ滞留時間の均一化性能の低下にもつながるなど、トレードオフの検討が必要となると考えられる。

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。

今回のシミュレーションでは、DSR 法によるデータ系列移動に要する時間を考慮していない。一次ストレージの容量は 100 GB から 500 GB を想定しているため、 $ratio = 0.05$ のとき、移動するデータの量は 5 GB から 25 GB に相当する。1 Gbps の帯域を用いた場合でも、これだけのデータの移動に要する時間は 40 秒から 200 秒となるため、DSUSN のデータアクセス性能にも影響を及ぼすと考えられる。そこで、今後はデータ移動の際のネットワーク遅延を考慮した手法の検討が必要であろう。

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、DSR 法によってデータ発生頻度の偏りに応じて動的にストレージ間でセンサ担当を変更し、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設計が容易になっている。

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発的に対応し、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。

また DSUSN は、利用するセンサを変えることにより、人の動きのセンシングに限らずさまざまなデータを収集できるように汎用性を考慮して設計されている。

今後の課題として、DSUSN のデータアクセス性能をさらに向上させるため、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する手法や、実環境でのデータ移動にともなうネットワーク遅延を考慮して、センサ担当変更の際に移動するデータの量を調整する手法、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あげられる。

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」ならびに、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)、科学研究費補助金(基盤研究 (B)(2))「大規模な仮想空間システムを構築する放送型サイバースペースに関する研究」(プロジェクト番号: 15300033) によっている。ここに記して謝意を表す。また、研究を進めるうえで貴重なご意見をいただいたサイバーメディアセンターの秋山豊和講師に深く感謝する。

参 考 文 献

- 1) Gupta, H., Das, S.R. and Gu, Q.: 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) 早坂光晴, 富永英義: 動画像からの背景画像生成を用いた移動物体抽出方法に関する一検討, 情報処理学会研究報告, オーディオビジュアル複合情報処理, IPSJ-AVM29-1, pp.1–6 (June 2000).
- 3) 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).

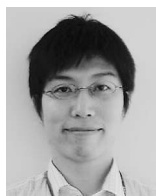
- 4) 楠本晶彦, 中澤 仁, 戸辺義人, 徳田英幸: A Location-Adaptive Virtual Networked Appli-
cance, 日本ソフトウェア科学会, Systems for Pro-
gramming and Applications (SPA2001) (Mar.
2001).
- 5) Madden, S., Franklin, M.J., Hellerstein, J.M.
and Hong, W.: The Design of an Acquisitional
Query Processor For Sensor Networks, *Proc.
Special Interest Group on Management of Data
(SIGMOD2003)*, San Diego, CA. pp.491-502
(June 2003).
- 6) Matthijs, F., Berbers, Y. and Verbaeten, P.:
A Flexible I/O Framework for Parallel and
Distributed Systems, *Proc. 4th International
Workshop on Object-Oriented in Operat-
ing Systems (IWOOS1995)*, Lund, Sweden,
pp.187-190 (Aug. 1995).
- 7) 南 正輝, 杉田 馨, 森川博之, 青山友紀: コピ
キタス環境に向けたインターネットアプリケーシ
ョンプラットフォーム, 電子情報通信学会論文誌,
Vol.J85-B, No.12, pp.2313-2330 (Dec. 2002).
- 8) Moleró, X., Silla, F., Santonja, V. and Duato,
J.: Modeling and Simulation of Storage Area
Networks, *Proc. 8th International Symposium
on Modeling, Analysis and Simulation of Com-
puter and Telecommunication Systems (MAS-
COTS2000)*, IEEE, San Francisco, pp.307-314
(Aug. 2000).
- 9) Moleró, X., Silla, F., Santonja, V. and Duato,
J.: On the Interconnection Topology for Stor-
age Area Networks, *Proc. 15th International
Parallel and Distributed Processing Symposium
(IPDPS2001)*, IEEE, San Francisco, p.162
(Apr. 2001).
- 10) 森川博之, 南 正輝, 青山友紀: コピキタスネッ
トワーキングへの道, 情報処理学会誌, Vol.43,
No.6, pp.631-638 (June 2002).
- 11) 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
(Mar. 2001).
(平成 16 年 9 月 20 日受付)
(平成 17 年 1 月 13 日採録)
- 12) Nakamoto, T., Suzuki, Y. and Moriizumi, T.:
Study of VHF-band QCM gas sensor, *Sensors
and Actuators B, Chemical*, Vol.84, Issues 2-3,
pp.98-105 (May 2002).
- 13) Qin, X., Jiang, H., Zhu, Y. and Swanson,
D.R.: A Dynamic Load Balancing Scheme for
I/O-Intensive Applications in Distributed Sys-
tems, *Proc. 32nd International Conference on
Parallel Processing (ICPP2003)*, Kaohsiung,
Taiwan, pp.79-86 (Oct. 2003).
- 14) Ratnasamy, S., Estrin, D., Govindan, R.,
Karp, B., Shenker, S., Yin, L. and Yu, F.:
Data-Centric Storage in Sensornets, *Proc. 1st
Workshop on Sensor Networks and Applica-
tions (WSNA2002)*, Atlanta (Sep. 2002).
- 15) Rhea, S., Wells, C., Eaton, P., Geels, D.,
Zhao, B., Weatherspoon, H. and Kubiawicz,
J.: Maintenance-Free Global Data Storage,
IEEE Internet Computing, Vol.5, No.5, pp.40-
49 (Sep./Oct. 2001).
- 16) 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, p.252 (Sep. 2003).
- 17) 白石 陽, 安西祐一郎: センサネットワークア
プリケーションのためのセンサデータエージェントの
設計と実装, *Workshop on Multi-Agent and Co-
operative Computation 98*, online (Dec. 1998).
- 18) 東京アメッシュ.
<http://tokyo-ame.jwa.or.jp/>
- 19) 浮田宗伯, 松山隆司: 移動対象の協調的追跡のた
めの観察可能領域モデル生成・更新法, 情報処理学
会論文誌, Vol.42, No.7, pp.1902-1913 (2001).
- 20) 梅澤 猛, 白石 陽, 安西祐一郎: センサネッ
トワーク利用のためのモバイルエージェントプラッ
トフォームの設計と実装, 日本ソフトウェア科学
会第 16 回大会, pp.417-420 (Sep. 1999).
- 21) Warneke, B., Last, M., Liebowitz, B. and
Pister, K.S.J.: Smart Dust: Communicating
with a Cubic-Millimeter Computer, *Computer
Magazine*, IEEE, Piscataway, NJ, pp. 44-51
(Jan. 2001).
- 22) 米山華子, 白石 陽, 安西祐一郎: センサネッ
トワーク利用のためのエージェントプラットフォ
ームの実装, *Workshop on Multi-Agent and Co-
operative Computation 98*, online (Dec. 1998).
- 23) Zhao, B.Y., Huang, L., Stribling, J., Rhea,
S.C., Joseph, A.D. and Kubiawicz, J.D.:
Tapestry: A Resilient Global-Scale Overlay for
Service Deployment, *IEEE Journal on Selected
Areas in Communications*, Vol.22, No.1, pp.41-
53 (Jan. 2004).

(担当編集委員 今井 正和)



田中 孝浩

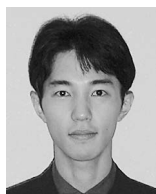
2003 年大阪大学工学部電子情報
エネルギー工学科卒業。同大学院情
報科学研究科マルチメディア工学専
攻博士前期課程在学中。



中尾 太郎 (正会員)

1999年大阪大学工学部情報システム工学科卒業。2003年同大学院工学研究科博士後期課程修了。工学博士。同大学院情報科学研究科マルチメディア工学専攻特任助手を経て、

現在株式会社NTTデータ。仮想空間、モバイルコンピューティングに興味を持つ。日本バーチャルリアリティ学会会員。



小川 剛史 (正会員)

1997年大阪大学工学部情報システム工学科卒業。1999年同大学院工学研究科博士前期課程修了。2000年同研究科博士後期課程中退後、大阪大学サイバーメディアセンター情報メディア教育研究部門助手となり、現在に至る。博士(情報科学)。グループウェア、ヒューマンインタフェース、バーチャルリアリティ、オーグメンティッドリアリティに興味を持つ。ACM、電子情報通信学会、日本バーチャルリアリティ学会会員。

現在に至る。博士(情報科学)。グループウェア、ヒューマンインタフェース、バーチャルリアリティ、オーグメンティッドリアリティに興味を持つ。ACM、電子情報通信学会、日本バーチャルリアリティ学会会員。



塚本 昌彦 (正会員)

1987年京都大学工学部数理工学科卒業。1989年同大学院工学研究科博士前期課程修了。同年シャープ株式会社入社。1995年大阪大学大学院工学研究科情報システム工学専

攻講師、1996年同専攻助教授。2002年同大学院情報科学研究科マルチメディア工学専攻助教授。2004年神戸大学工学部電気電子工学科教授となり、現在に至る。工学博士。ウェアラブルコンピューティングとユビキタスコンピューティングの研究に従事。ACM、IEEE等8学会の会員。



西尾章治郎 (フェロー)

1975年京都大学工学部数理工学科卒業。1980年同大学院工学研究科博士後期課程修了。工学博士。京都大学工学部助手、大阪大学基礎工学部および情報処理教育センター助

教授、大阪大学大学院工学研究科情報システム工学専攻教授を経て、2002年より同大学院情報科学研究科マルチメディア工学専攻教授となり、現在に至る。2000年より大阪大学サイバーメディアセンター長、2003年より大阪大学大学院情報科学研究科長を併任。この間、カナダ・ウォータールー大学、ビクトリア大学客員。データベース、マルチメディアシステムの研究に従事。現在、Data & Knowledge Engineering等の論文誌編集委員。本会理事を歴任。電子情報通信学会フェローを含め、ACM、IEEE等9学会の会員。