

TomuDB: 都市空間センサ情報を扱うデータベースシステム

岩井 将行[†] NiwatThepvilojanapong[†] 石塚 宏紀^{††} 中村 陽一^{†††} 金井 圭介^{†††}
白石 陽^{††††} 戸辺 義人[†]

[†] 東京電機大学 未来科学部 〒101-0054 東京都千代田区神田錦町 2-5 大晃ビル 501

^{††} 東京電機大学大学院先端科学技術研究科

^{†††} 東京電機大学大学院工学研究科情報

^{††††} 東京大学 空間情報科学研究センター 〒277-8568 千葉県柏市柏の葉 5-1-5

柏キャンパス 総合研究棟 451 号室

E-mail: †{masa,wat}@osoite.jp

あらまし 近年、センサデバイスの小型化により、実世界の環境情報を安価なデバイスで取得できるようになってきている。特に、都市空間におけるヒートアイランド現状や人や車の流れを検知する必要が強まり、短期間で準備を行い、広範囲のエリアで情報を計測する要望が高まりつつある。しかし、センサは広域に分散して配置しなければならず、全てのセンサデータを一極に収集すると爆発的にトラフィックが増大する問題が発生する。我々は、都市空間の通信網でDB間をオーバーレイネットワークによって接続し効率よくデータの収集と検索を行う TomuDB を開発した。キーワード 都市センシング、分散データベース、センサデータクエリ

TomuDB: A Database Management System for Distributed Urban Sensors

Masayuki IWAI[†], Niwat THEPVILOJANAPONG[†], Hiroki ISHIZUKA^{††}, Yoichi NAKAMURA^{†††},
Keisuke KANAI^{††††}, Yoh SHIRAISHI^{††††}, and Yoshito TOBE[†]

[†] School of Science and Technology for Future Life Tokyo Denki University Researcher of OSOITE,
CREST, JST

^{††} Tokyo Denki University, Graduate School of Advanced Science and Technology

^{†††} Tokyo Denki University, Graduate School of Engineering

^{††††} University of Tokyo, Center for Spatial Information Science

E-mail: †{masa,wat}@osoite.jp

Abstract People in urban city are likely to perform a real-world search; they search things related to their daily lives in the real world ranging from straight-forward queries to complex ones. However, a single sensor network cannot respond to a variety of queries by using its own sensing data. Currently, each organization deploys sensor networks for their own uses. Each sensor network has many different characteristics such as sensor types, network size, sensing resolution, etc. In order to realize the real-world searches in urban cities, a collaboration of such heterogeneous sensor networks is a promising solution. This paper proposes a *TomuDB* architecture for performing distributed queries in heterogeneous sensor networks. In order to cope with the real-world searches, heterogeneous sensor networks work collaboratively by forwarding queries to other networks that are likely to answer the queries. A prototype implementation is presented and the experiments are performed to study the performance of TomuDB architecture in the real-world environment.

Key words Urban Sensing, Distributed Database, Queries for Sensor Data

1. まえがき

センサノードの小型化・多様化、無線ネットワーク技術の発達により、様々な用途のセンサネットワークが設置されつつあり、都市環境をセンシングし、モニタリングするためのセンサネットワークに関するプロジェクトや研究活動が進められつつある[1]~[3]。従来は、気象、地震、大気汚染など比較的広範囲の密度の小さいセンサデータ[4]~[7]しか利用できなかったが、近年では、都市レベルの、より高い密度で計測時間間隔の短いセンサデータを取得できる環境が整いつつある。すなわち、広範囲の粒度の粗いセンサネットワークに加えて、より狭い範囲の粒度の細かいセンサネットワークからも、実世界の情報を取得することによって、都市環境の分析やモニタリングだけでなく、日常生活の様々な支援にも役立てることができるようになると考えられる。

こういった都市部のセンサは現状では各ビル等に設置され独自のアプリケーションに特化した形で、個々の組織の内側でデータを管理されているのが現状である。ネットワークドメインが異なるユーザが容易にアクセスすることは不可能であった。

本研究では、設置した組織、データの種類の異なるセンサネットワークが複数混在する環境（「異種センサネットワーク環境」）を想定する。異種センサネットワーク環境においては、複数のセンサネットワークから種類の異なる様々なデータを収集して、それらの統合処理の結果をユーザに提示する必要がある。

地理情報システム（Geographic Information Systems: GIS）の分野では、空間分析に関する様々な手法が開発されており、異なる種類の空間データを分析・統合する方法として、オーバーレイ手法[13]がある。GISのオーバーレイ手法では、異なる種類のデータに対して、同一のメッシュ構造を生成することで、対応するメッシュのデータを重ね合わせることで、容易に比較したり、組み合わせたりすることが可能である。本研究では、アプリケーションにおいて、GISにおけるオーバーレイ手法を用いて異種センサネットワークからの多様なデータを処理することを想定して、分散型のDBを提案する。

全体的な手順としては、まず各組織が管理するDBを外部に公開可能なDBMSを有するセンサノードとして起動する。我々はDBSN(Data Base management system as a Sensor Node)とよぶ。各DBSN peerは、管理エリア（クエリ処理可能領域）を有しており、より、自らを包含するより上位のDBSN peerと常時通信を行うようにネットワークを構築する。DBSN peerは、重なる部分のメッシュデータを作成するための補間処理を行い、その上で、各DBSN peerの補間結果を用いて、クエリ領域全体の補間処理を行い、クエリ領域全体のメッシュを生成する。本手法は、データの種類の異なるセンサネットワークが混在する環境においても、指定したセンサデータを生成し、アプリケーションに提供することが可能になる。

本論文の構成は、以下の通りである。2章で、問題点と関連研究について述べる。3章で、TomuDBのアーキテクチャについて説明を行い、4章では、実験結果に基づいてTomuDBの

性能評価を行う。5章で、結論を述べる。

2. 問題意識と関連研究

2.1 問題意識

現在、センサの値段が安価になり個人や小規模オフィスでもセンサを購入することが可能になった。これらのセンサはうまく相互利用されて以内ため、以下のようなシナリオの利用ができていない。(1) 高齢者が真夏の熱中病になる恐れがあるためにできるだけ涼しい道を選んで目的地まで行きたい。各道の温度を集めてきたい。(2) 暗い夜道を避けてなるべく明るい道を歩いて帰宅したい。現時点で暗く、人の気配のない道を避けるルートを表示する。このようなシナリオがあった場合現状では個々のセンサデータベースが接続されていないためアプリケーションを開発することが困難である。

また、我々の想定する都市においては、センサの管理者は個々の組織中で運用するセンサを一般向けにボランティアとして提供することを協力してくれることを想定している。扱うデータベースやデータの種類の異なるエリアやセンサの精度(Resolution)も異なってくる。また個々のセンサはボランティアによって運営されるためネットワーク上で常時利用可能とは限らない。そのような不確実性に対応できなければならない。

更に、センサは広域に分散して配置しなければならず、すべてのセンサデータを一極に収集すると爆発的にトラフィックが増大する問題が発生する。

2.2 関連研究

既存の分散データベースシステムにおいて、クエリ内のデータタイプ、データ取得領域、空間解像度の条件をすべて満たす1つのDBピアがあれば、そのピアにクエリを転送し、データ処理可能である。また、データ取得領域が複数のDBピアに跨っていても、その全域を複数ピアが完全包含していれば複数のDBピア内のデータを統合してクエリ領域に関わるデータを提供することができる。しかし、上記の問題の要に都市部のセンサの相互運用においてこのようなDBは障害が多い。

センサデータベースに関する研究は、大きく2つに分けることができる。一つは、ネットワーク内問合せ処理(in-network query processing)、もう一つは、ストリームデータ処理(stream data processing)である。ネットワーク内問合せ処理では、バッテリーの寿命や通信帯域などの物理的な制約の中で、センサネットワーク内でのデータ収集やデータ集約(agggregation)のための問合せを効率的に処理するための様々な手法が提案されている[9],[10]。ストリーム処理に関する研究では、データソースから時々刻々と到着するデータに対して、統計処理、結合演算、マイニングのための沢山の手法が提案されている[15]~[17]。ネットワーク内問合せ処理に関する研究は、結合演算や単純な集約処理(最小値、最大値、平均など)を対象としたもの[14]から、近年、点的なセンサデータ分布から等高線や連続的な面分布を計算するなど、より高度な処理を対象とした研究へと進んでいる[11]。Nittelらの研究[11]では、地理情報システム(GIS: Geographic Information System)の空間補間処理をセンサネットワーク内で実行するための手法を提案しているが、

地理情報システムとセンサネットワーク・データベースの分野横断的な研究領域として、地理センサネットワーク (Geo-sensor network) [18] に関する研究も行われている。これらの研究では、センサノード単位で処理が行われているのに対し、我々の提案手法では、センサノードからデータを収集したデータベース管理システム群によって「ネットワーク内間合せ処理」が行われる点が異なる。白石らは、異種センサネットワーク環境を想定したセンサデータの検索・統合手法を提案した [12] が、本提案と異なり、アプリケーションが存在するクライアント側でのデータ統合を対象としている。

ストリーム処理に関する研究では、到着したデータを逐次的に処理する、あるいは、別のデータベースのデータと統合するといった研究が見られるが、近年では、センサからのデータを時々刻々とアーカイブとして蓄積する方向性も見られる [16]。過去のデータとの統合を考えた場合、センサデータの蓄積は重要である。Kraft では、鮮度とリアルタイム性を考慮して、データを蓄積するための手法を提案しており、時系列の類似度計算など高度な問合せ処理もサポートしている [19]。MauveDB および Kraft では、こうした高度な処理が可能であるが、分散アーキテクチャとなっていない点で、我々の提案手法と異なる。

3. TomuDB の概要

TomuDB アーキテクチャの詳細を述べる。Tomu DB は、複数のセンサデータを持つデータベース (DBSN peer) と、クライアントと DBSN peer の通信を変換、統合、要約、補間など、さまざまな処理をとりもつゲートウェイ (DBSN requester) によって、センサデータを効率的に分散管理、検索する分散センサデータベースである。Tomu DB は、DBSN requester と複数の DBSN peer で構成される。DBSN requester は、Tomu DB におけるクライアントとの通信部分であり、DBSN peer はセンサデータの管理、蓄積、検索部分である。DBSN requester は、クライアントからのクエリの種類に対して、適切な形で DBSN peer との通信を行い、DBSN peer から受信したデータをクライアントの要求する形へ変換し、クライアントへ送信する。DBSN peer はセンサデータを管理、蓄積し、DBSN requester から他の DBSN peer を経由して送信されてきたクエリに対して、自己の管理領域、センサデータとの照合を行い、必要であればデータの送信とクエリの転送を行う。DBSN peer はスーパーノード DBSN peer と下位 DBSN peer を持つ木構造のネットワークとなっており、個々の DBSN peer は木構造を構成する一部である。各 DBSN peer がどのような親子関係を持つかは、初期時に設定として与えられる。

TomuDB は以下の構成からなる (図 1)。

- **Web API Server:** Web API server は、DBSN に対してクエリを出す処理を行う。ユーザは Web API server を経由してデータ転送を行うことが可能になる (図 2)。Web API の利用例として、温度データを問い合わせることによる「猛暑時における涼しい道の検索」アプリケーションなどが想定される。XML query の例を 3 に示す。

- **DBSN Requester:** The DBSN requester は、XML

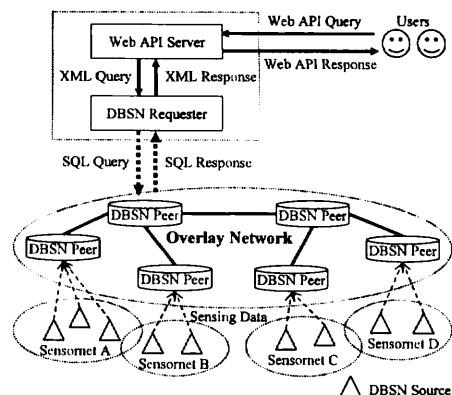


図 1 TomuDB アーキテクチャ

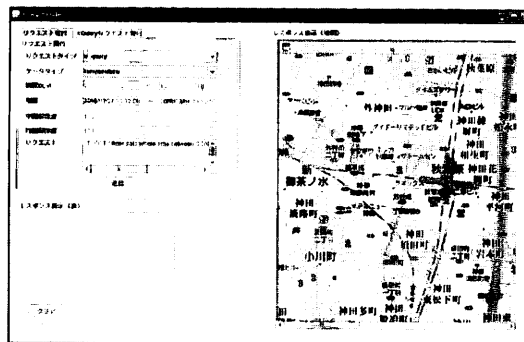


図 2 TomuDB web API

query から SQL query を効率的に生成する。また SQL query から XML query の逆変換も行う。SQL query は dQuery と呼んでおり、DBSN peer へ転送される。

- **DBSN Peer:** DBSN peer の処理はセンシングデータを収集し DB で管理することである。構成要素として、DBSN ピア情報、ルーティングテーブル、DBSN ソースから集めたセンシングデータを格納するデータベースがある。さらに DBSN peer は、他の DBSN Peer や DBSN Requester との通信も行う機能を有する。

- **DBSN Source:** センサで計測し、DBSN ピアにセンサデータを送信するセンサノードの総称である。Particle Computer 社製無線センサモジュールの uPart [8] と受信機の USB bridge を想定する。複数の uPart は対象となる環境に設置され、USB bridge によってセンサデータを受信し、PC にデータを送る。PC では受信したデータを USB ポートから読み取り、DBSN ピアにデータを転送し保存する。uPart には温度・照度・振動センサが組み込まれており、定期的に無線モジュールによって各センサデータが送信される。送信データ間隔は利用者が設定可能である。(ただし、送信間隔は 144[msec] の整数倍の間隔で設定可能である) 設置箇所のセンシングを行うことが可能であり、対象環境に対して複数の uPart を設置することで温度分布などの測定が可能となる。

- **dQuery:** DBSN ピアに問い合わせを行うためのクエ

```

<query xmlns="http://osoite.jp/drops/datatype/
xml/query" xmlns:geo="http://osoite.jp/drops/
datatype/xml/geometry">
  <requester>
    <host>requester.test.osoite.jp</host>
    <port>5555</port>
  </requester>
  <type>temperature</type>
  <time>
    <start>2008/4/28 10:00:00.0000</start>
    <end>2008/4/28 15:00:00.0000</end>
  </time>
  <space type="2d">
    <region type="polygon">
      <point><x>0</x><y>0</y></point>
      <point><x>0</x><y>10</y></point>
      <point><x>10</x><y>10</y></point>
      <point><x>10</x><y>0</y></point>
      <point><x>0</x><y>0</y></point>
    </region>
    <resolution>1.0</resolution>
  </space>
</query>

```

図3 XML クエリのサンプル

表1 クエリのパラメータ

Table 1 Parameters of a query

パラメータ	説明
$\hat{A}(Q)$	クエリの指定する矩形領域
$\hat{D}T(Q)$	データの種類
$\hat{T}I(Q)$	時系列を要求するための時区間
dW	メッシュ幅 (セルのサイズ)
$\hat{R}_s(Q)$	空間解像度
$SQL(Q)$	SQL スレートメント

りである。DBSN ピア間のルーティング制御を行うヘッダと、DBSN ピアのデータベースへの問い合わせを行う SQL で構成される。利用する SQL は、DBMS に依存しない標準仕様に準ずる。dQuery のヘッダに含まれる要素を表 1 に示す。

4. 評価実験

我々は TomuDB を実装し、そのパフォーマンスを測定した。実験では、TomuDB の中で DBSN requester, DBSN peer, DBSN source を用いた。各実装は、Linux 上で C 言語で開発されており、DBSN peer の台数のみ 3 台起動した。The DBSN source は、リージョンの温度情報と照度データを取得できる uPart [8] を利用している。DBSN peer は PostgreSQL を拡張する形で利用している。また、リージョンの形は本実験の処理手続きを簡素化するために四角形のみを用いて行う。リージョン情報は、座標で表し $[(x_u, y_u), (x_{ur}, y_{ur})]$ と記述してい

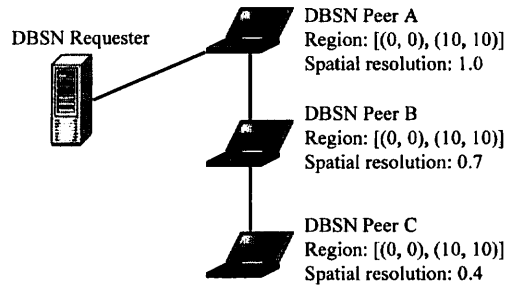


図4 実験 I における DBSN peer のネットワーク構成図

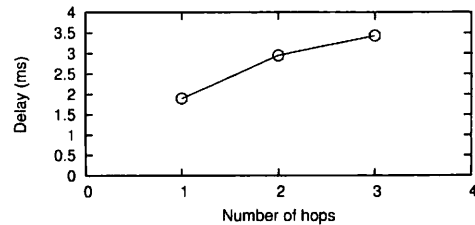


図5 実験 I における遅延時間の推移

る。実環境での評価を想定し我々は、クエリ伝播とその処理時間想定し、クエリ送信時から、いくつかの DBSN ノードを経由して返答を受け取るまでの時間に着目した。

4.1 実験 I: Resolution-based Queries

実験の環境を、図 4 に示す。1 台の DBSN requester が常に DBSN peer(A, B, C) に対してクエリをおくる。A, B, C は直列に組まれており管理されるリージョンは $[(0, 0), (10, 10)]$ のようにすべて等しいものとする。一方、センシングリージョンの細かさ (Resolution) は、相対的に 1.0, 0.7, 0.4 となっており C がもっとも細かい。そのためより細かい Resolution を希望するクエリに対しては A- \rightarrow B-C とマルチホップされ C が返答を行う。

DBSN requester は event ($E = \text{all measured temperature}$), region ($L = [(0, 0), (10, 10)]$), the period of time ($T = [18:50, 18:55]$) として、Resolution 以外は同一となっている。

クエリの Resolution 部分は 1.0, 0.7, 0.4 としたため、各クエリは最適な DBSN peer へと転送される回数が変わり 1, 2, 3 ホップすることになる。例えば Resolution を 0.4 とした場合は、ピア A で $1.0 \geq 0.4$ と判断され、ピア B に転送される。さらにピア B において $0.7 \geq 0.4$ と判断されピア C まで転送され、ピア C で初めて処理される。

平均遅延時間は、図 5 の結果になった。平均遅延時間の増加は予想のとおり、リニアにあがっておりホップ数以上のオーバーヘッドはないためある程度の階層まで DBSN peer を構成することが可能である。

4.2 実験 II: Region-based Queries

図 6 は、実験 II における DBSN peer のネットワーク構成である。温度データは、図のように異なったリージョンで管理される。DBSN peer A, B, C の管理リージョンは $[(0, 0), (10, 10)]$,

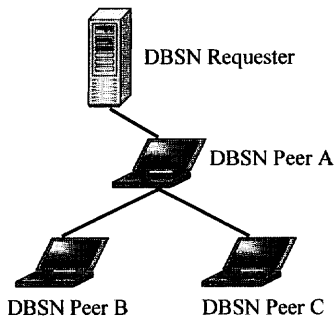
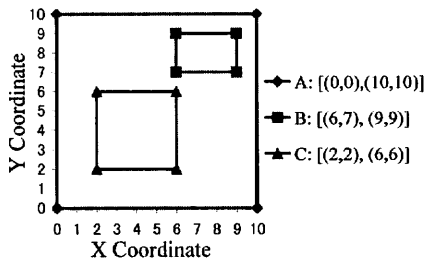
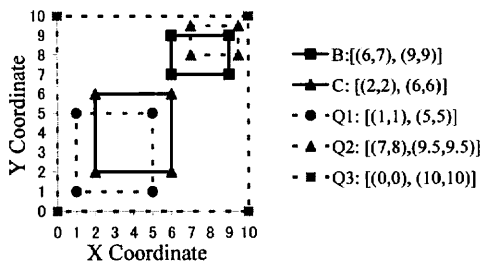


図 6 実験 II における DBSN peer のネットワーク構成図



(a) DBSN peer の Region 構成.



(b) Region 別のクエリ (Q1, Q2, and Q3) と DBSN peer で管理する Region との関係

図 7 Regions of DBSN peers and queries.

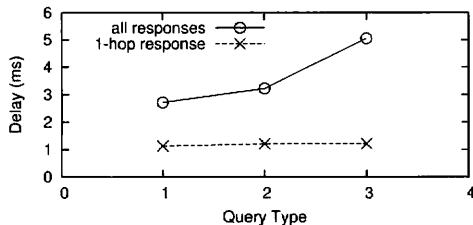


図 8 実験 II での遅延時間測定

[(6, 7), (9, 9)], [(2, 2), (6, 6)] となっている。

図 7(a) に視覚的に関係を示す。DBSN requester の出すクエリは図 7(b) にあるように 3 種 Q1, Q2, Q3 となっておりそれぞれのリージョンは [(1, 1), (5, 5)], [(7, 8), (9.5, 9.5)], [(0, 0), (10, 10)] となっている。

クエリは、以下のように処理される。まず DBSN peer A and C は Q1 が要求する温度データを有する。そのため DBSN peer A は DBSN requester に対して返答すると共に、

DBSN peer C に Q1 を転送する。DBSN peer C は the DBSN requester に返答を返す。

クエリ Q2 は DBSN peers A, B を含むようなエリアを要求している。DBSN peers A, B はクエリ受信後それぞれ DBSN requester に返事をする。

Q3 はすべての DBSN peer に対して包含しているため、各 DBSN peer は DBSN requester に返答することになる。図 8 に 1hop の場合の参考遅延時間と適切な処理をされてきた場合の遅延時間を記載した。

one-hop の遅延時間は DBSN peer から DBSN requester へのクエリ回答時間である。one-hop の遅延時間と比較して 3 角クエリ処理には時間がかかる。理由としては DBSN peer はリージョン比較とクエリの転送処理を行うためである。時間の差はわずかではあるが転送によるオーバーヘッドが見られるため、アプリケーション側からはクエリ転送を極力抑えるようなクエリの生成が必要とされる。

5. まとめ

センサを用いた都市における実世界検索アプリケーションのニーズは高まりつつあるが、一方で、種類も複雑であり管理が個別の行われおり、一般の人が利用できる枠組みが今までなかった。TomuDB はそのような実世界におけるアプリケーションを実現するため、異種センサを扱う DB をつなぎ、一般の人からも検索可能にする枠組みである。各 DB は協調し、別々にローカルで管理されていた異種データをシンプルに接続する。クエリ転送の Reason や Resolution の属性で最適化し全体としてのデータ転送の負荷を軽減している。性能評価をおこない複数の DBSN ピアを用いて時間的なシステムの妥当性を検証した。

謝辞 本研究は、独立行政法人科学技術振興機構 先進的統合センシング技術 JST CREST OSOITE プロジェクトの一部で行われたものです。

文 献

- [1] Microsoft, <http://atom.research.microsoft.com/sensormap/>.
- [2] Live E! ~活きた地球の環境情報~, <http://www.live-e.org/>.
- [3] Ito, M., Katagiri, Y., Ishikawa, M. and Tokuda, H.: Airy Notes: An Experiment of Microclimate Monitoring in Shinjuku Gyoen Garden, Fourth International Conference on Networked Sensing Systems (INSS 2007), pp.260-266 (2007)
- [4] 日本気象協会: ワンクリック気象情報サイト, <http://tenki.jp/>.
- [5] 環境省: 大気汚染物質広域監視システム, <http://w-soramame.nies.go.jp/>.
- [6] 日本道路交通情報センター, <http://www.jartic.or.jp/>.
- [7] 防災科学技術研究所, 高感度地震観測網 (Hi-net), <http://www.hinet.bosai.go.jp/>.
- [8] TecO/University of Karlsruhe, "Particle Computer," available from <http://particle.teco.edu/>.
- [9] Gehrke, J. and Madden, S.: Query Processing in Sensor Networks, IEEE Pervasive Computing, Vol.3, No.1, pp.46-55 (2004).
- [10] Madden, S., Franklin, M.J., Hellerstein, J.M. and Hong, W.: TinyDB: an acquisitional query processing system for sensor networks, ACM Transactions on Database Systems, Vol.30,

- No.1, pp.122-173 (2005).
- [11] シルビアニッテル, グァンジン, 白石陽: センサネットワークにおけるネットワーク内空間推定, 電子情報通信学会論文誌 A, Vol.J88-A, No.12, pp.1413-1421 (2005).
 - [12] 白石陽, 安西祐一郎: センサデータの視覚化のためのインクリメンタルな空間集約手法, 情報処理学会論文誌: データベース, Vol. 45, No.SIG7 (TOD 22), pp.63-76 (2004).
 - [13] Longley, P. A., Goodchild, M. F., Maguire, D. J. and Rhind, D. W.: *Geographic Information Systems and Science*, John Wiley & Sons, Ltd (2001).
 - [14] Madden, S. R., Franklin, M. J., Hellerstein, J. M. and Hong, W.: TAG: a Tiny AGgregation Service for Ad-Hoc Sensor Networks, Proc. 5th Symposium on Operating System Design and Implementation (OSDI), (2002).
 - [15] The STREAM Project,
<http://www-db.stanford.edu/stream/>.
 - [16] Abadi, D., Carney, D., Cetintemel, U., Cherniack, M., Conway, C., Lee, S., Stonebraker, M., Tatbul, N. and Zdonik, S.: Aurora: A New Model and Architecture for Data Stream Management, The VLDB Journal, Vol.12, No.2, pp.120-139 (2003).
 - [17] The Borealis Project,
<http://nms.lcs.mit.edu/projects/borealis/>.
 - [18] A. Stefanidis and S. Nittel, "Geo Sensor Networks", CRC Press, 2004.
 - [19] 川島英之, 遺山元道, 今井倫太, 安西祐一郎, 「センサデータベースシステム KRAFT の設計と実装」, 情報処理学会論文誌: データベース (TOD 24) , (2004).