

モバイルセンサデータベースのための データ記録保証機構

Guaranteed Data Storing Mechanism for Mobile Sensor Database

由良 淳一 ♡
徳田 英幸 ♠

中澤 仁 ◆

Jun'ichi YURA
Hideyuki TOKUDA

Jin NAKAZAWA

本論文では、携帯型小型ノード内に近傍のセンサデータを記録していくモバイルセンサデータベースを提案する。この際、ノードの記憶容量の制約から最大記録可能時間が限られるため、必要記録時間と許容する記録データの必要品質を利用者に指定させ、それを複数のフィルタによりセンサデータを適応的に取捨選択して記録することで保証する。また、本機構の実機への実装と、実データを用いた実験による評価を示す。

We propose a sensor database for tiny mobile nodes. Due to the storage capacity limitation such nodes, the sensor database needs to reduce the amount of data to service for a long period of time. This paper shows an implementation and an evaluation of multiple filters and a scheme to adjust their threshold adaptively to the storage capacity, guaranteeing the quality of storing data and the service time requested by the user.

1. はじめに

現在、屋内外の物理的な環境内に無線センサネットワークを敷設して、センサ情報を収集する技術が盛んに研究されている。同技術の応用としては、以下のような例が考えられている。

- (1) ユーザのライフログ ユーザの接した物を RFID などの電子タグ技術を用いて記録し、同時に周囲の環境情報や位置情報をセンサを用いて記録する。
- (2) 商品のトレーサビリティ 製品が作られてから消費者に渡るまで、その周囲の状況を記録する。例えば、ある食品の周囲の温度を 10 分ごとに記録することで、流通の過程を確認できる。

一般的な無線センサネットワークでは、無線センサノードで

検知したセンサデータをインターネット上のサーバに集中させ、データベースへ記録して、複数のアプリケーションから利用可能とする。既存の広域センサネットワークに関する研究 [3] は、こうした記録方式を地理的に分散したセンサネットワーク群において可能とする。しかし、このような集中記録方式は耐故障性が低い上に、超広域をカバーするためには複数サーバの連携をはじめとする複雑な機構が必要となる。

本研究では、利用者が保持する小型ノード内のデータベースに近傍のセンサデータを記録していく、モバイルセンサデータベース (以下 MSDB と呼ぶ) 方式を採用する。同方式はパブリックネットワークを必要としないため、実利用時の制約が少ない。さらに、記録対象ごとにデータベースが構築されるため、対象ごとに異なるデータ記録の要求を指定できる。一方、MSDB を構築する小型ノード (無線センサノードや PDA など) は、記憶容量が小さいため、センサデータを記録できる最大時間が限られる。従って MSDB では、以下の制約を同時に満たす必要がある。

- (1) 必要記録時間の保証 記録時にセンサデータを減少させ、必要記録時間を保証しながら実際の記録時間を最大化する。例えば、無線センサノードである MICA2DOT[1] の内蔵記憶容量は 512KB であり、3 つのセンサノードから 1 秒ごとに 20 バイトのセンサデータを受信すると想定した場合、142 分しか記録できない。この領域に 12 時間分のセンサデータを記録するためには、実際のデータ記録を数秒に 1 回程度に抑制する必要がある。
- (2) 必要データ品質の保証 センサデータの減少をアプリケーション要求を考慮して行うことで、データ品質を最大化する。例えば、短い時間間隔で記録を行うアプリケーションでは記録時間間隔を短くし、詳細な値が必要なアプリケーションでは入力データ値の差が小さい場合にも記録するなど、アプリケーションごとに異なる要求を満たして記録する。

本論文では、利用者に、必要記録時間と許容する記録データの必要品質を指定させ、それを保証しながらセンサデータを取捨選択して記録する方式 (センサデータの凝縮) を提案し、実機への実装と実データを用いた実験による評価を示す。第 2 章で関連研究との差異を示し、第 3 章で提案手法の詳細と実機への実装を示す。第 4 章で提案手法を実験により評価し、第 5 章でまとめる。

2. 関連研究

センサネットワークを抽象化してデータベースとして扱うシステムとして、TinyDB[5] や Cougar[6] などの網内データ処理技術が挙げられる。これらのシステムでは、センサデータの獲得を SQL に準じた問合せ言語を用いて行える。本研究の対象とするアプリケーションでは、記録対象が必要とするセンサデータはその対象が存在する空間に関するセンサデータである。従ってこれらのシステムでは、移動対象を継続的に追跡する必要があり、パブリックネットワークを持たないセンサネットワークでは記録できない。一方本論文では、MSDB に対する問合せ言語については規定せず、既存の組込みデータベースを利用することを想定する。本研究は、そうしたデータベースに効率的にセンサデータ

♡ 正会員 慶應義塾大学大学院 政策・メディア研究科
yurayura@ht.sfc.keio.ac.jp

◆ 慶應義塾大学大学院 政策・メディア研究科
jin@ht.sfc.keio.ac.jp

♠ 慶應義塾大学 環境情報学部 hxt@ht.sfc.keio.ac.jp

を記録する際の保証付きデータ量減少機構と位置づけられる。

小型ノード内でのセンサデータの減少方式としては、コンテキスト抽出 [2] と呼ばれる手法がある。この手法では、受信した複数のセンサデータ値の組み合わせが、幅を持ったある条件に合致する場合、データそのものではなく合致した条件を記録する手法である。例えば、温度が 30 度以上で湿度が 75 % 以上のデータが得られた場合「蒸し暑い」というコンテキストを抽出してそれを記録することにより、記録に必要な容量を減少できる可能性がある。こうした手法は、センサデータの値が重要ではない場合には有効である。これに対して本研究ではワインの流過程での温度の記録等のように値自体が重要である応用も想定する。従って本論文では必要なデータ品質を保証しながら記録するデータ量を減少させる手法を提案する。

前後のデータに対して値の関連や類似度を用いることで、データの圧縮を行える [4]。提案方式であるセンサデータ凝縮では、連続する複数のデータの受信間隔や、類似度に基づいてデータの取捨選択を行うことで同様の圧縮を実現している。また、センサデータ全体に対し既存の可逆圧縮を適用することにより、本機構によって出力されたデータをさらに圧縮することも可能である。ただし本論文では、そうした圧縮の前段階でデータを取捨選択する手法にフォーカスする。

3. 保証付きセンサデータ記録機構

本機構では、入力センサデータに対して異なる複数のフィルタを順次適用することでデータ凝縮を行い、結果を MSDB に記録する。本研究では、MICA2DOT 無線センサノード (Atmel ATmega128L 4MHz CPU, 512KB Flash ROM, TinyOS 1.1) に、nesC 1.1.3 (avr-gcc 3.4.6) を用いて本機構を実装した。

3.1 データフィルタ

データフィルタは、入力データに対して単純なアルゴリズムを適用してデータの破棄を決定する。各入力データには、温度や照度といったデータの種類 (以下タイプと呼ぶ) が付随しており、フィルタは全てのデータをタイプごとに処理する。破棄されないデータは次のフィルタに渡され、最後のフィルタから出力されたデータは MSDB に記録される。閾値を持つフィルタは、その値を動的に変化させることで、本機構の提供する動的なデータ凝縮を実現する。以下に、本機構で利用するフィルタを示す。

平均化フィルタ 同時刻に異なるセンサノードから入力したデータの平均値を求め単一の値として出力する。

時間フィルタ センサデータが閾値時間内に複数入力された場合に破棄する。閾値を大きくすると、記録データ数が減少する。

距離フィルタ 発信元からのネットワークホップ数が閾値よりも大きいデータを破棄する。閾値を大きくするとより遠くのセンサデータまで記録され、記録データ数が増加する。

類似フィルタ 入力データと直前のデータとのユークリッド距離が閾値以下の場合、入力データを類似データとみなして破棄する。閾値を大きくすると記録データ数が減少する。

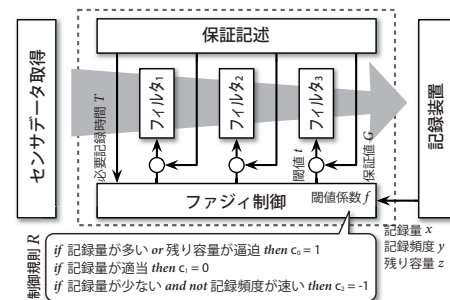


図1 保証機構 概要図

Fig. 1 Overview of Guaranteed Data Reduction System

3.2 保証機構

前節のフィルタを、閾値を固定して適用した場合、必要記録時間を達成できない場合がある。そのため本機構では、ユーザに必要記録時間に関する要求を指定させ、それを保証するようにフィルタ閾値の動的適応を行う。またユーザは、必要記録時間と同時に必要データ品質を指定できる。必要データ品質は、記録データをそれ以上減少させない各フィルタの閾値の限界値 (以下保証値と呼ぶ) として指定する。例えば、必要記録時間が 48 時間であり、最長データ記録間隔 (時間フィルタの保証値) が 10 秒である場合を想定する。この場合、MSDB は常に 10 秒以下の閾値を選択しながら記録データを凝縮する。

しかし、記録可能時間と必要記録時間の大小による 2 値の述語論理により閾値を決定した場合、閾値が過度に変化し、記録データの品質が時間によりばらつく可能性がある。そのため本保証機構では、現在の記録量、記録頻度、残り容量の値をファジィ制御により処理し、フィルタ閾値を決定する。図1に保証機構の概要図を示す。無線センサネットワークを介して取得したセンサデータは、複数のフィルタを経由して記録装置に保存される。次に、記録量 x 、記録頻度 y 、残り容量 z からファジィ集合を求め、制御規則 R の前件部に適用することで、各規則の適合度 w_i (ただし i は規則番号) を算出する。制御規則 R の後件部は、閾値係数算出のための実数値 c_i で指定されているため、加重平均をとることで、最終的な閾値係数 f が得られる。ここで、各フィルタごとに持つ閾値 t に重み付けされた閾値係数を加算することで、新しい閾値となる。閾値がユーザによって指定された保証値 G 外に存在する場合は、閾値を保証値にすることで、記録データの必要品質を保証する。以上により、過度あるいは急激な閾値変化を抑制し、記録データ品質を安定させる。

3.3 保証記述と保証選択

保証値に対する要求は、アプリケーションごとに異なる。例えばライフログアプリケーションでは、RFID や加速度、温度や湿度、照度など、多くの種類のセンサデータが必要となり、また、時間粒度を細かくとる必要がある。一方、食品のトレーサビリティを実現する場合、温度と湿度のみの限られたセンサデータを記録すればよいが、必要記録時間は 1 週間など長期となる。このように各アプリケーションごとに異なる必要記録時間や保証値を適用するため、本研究ではアプリケーションごとに保証記述

Profile: LifeLog	Profile: Tracerbility
GuaranteedTime: 172800	GuaranteedTime: 604800
TimeFilter: 0, 10	TimeFilter: 0, 120
ProximityFilter: 0, 20	ProximityFilter: 0, 5

図2 アプリケーションごとの保証記述例

Fig. 2 Guaranteed Descriptions of Applications

を指定し、その記述を元にデータ凝縮を行う。図2左および図2右にライフログアプリケーションおよびトレーサビリティアプリケーションの保証記述例を示す。

図2左では、48時間(172,800秒)の必要記録時間を要求している。最低10秒ごとに記録するため時間フィルタの閾値は10、また20%以下の類似データは必要ないため類似フィルタの閾値を20とした。一方図2右では、1週間(604,800秒)のセンサデータを記録する。120秒ごとに詳細なデータを記録するため、時間フィルタの閾値を120、類似フィルタの閾値を5に設定している。平均化フィルタは閾値が常に一定のため、両記述において保証値の指定はない。なお、取得したセンサデータの類似度が類似フィルタの保証値を常に超える場合など、保証記述を満足させつつ必要時間の記録を行えない場合がある。その場合は、保証値の変更または必要記録時間の短縮をユーザに提示し、アプリケーションの保証記述の内容を変更することで対応する。また記憶容量を使い切った場合においても、記録の終了や上書きをユーザに提示することで、記録の継続を選択させる。

4. 評価

本節ではセンサデータ記録保証機構の評価を、実データを用いた実験により述べる。本実験は無線ネットワーク状態による取得されるセンサデータの差異を除外するとともに、異なる条件で複数回の同じセンサデータを適用する必要がある。そのため、無線センサノードが送信したセンサデータを、ゲートウェイとして動作する無線センサノードが接続されたPC上に記録し、PC上に実装した記録保証機構を用いて行った。実験は、15m四方の空間にMICA2無線センサノードを30個配置し、1週間分の照度データを記録することで得られた1,695,719個のセンサデータを用いて行った。1データの大きさはタイムスタンプ等を合わせて20バイトであるため、前述の小型センサノードの持つ512KBの記憶装置には26,214個のセンサデータを記録できる。そのため本実験では、データの記録容量を26,214ブロックとし、記録終了時にセンサデータをこの数まで減少させることを目標とする。

評価1: 閾値の動的変化によるデータ量の推移

フィルタ閾値を動的に変化させることの有効性を、記録データ量の推移より評価する。前述のトレーサビリティアプリケーションの必要データ品質は、記録間隔が120秒以下、類似度は5%以下、また必要記録時間は1週間である。これを用いて、時間フィルタの閾値を120、類似フィルタの閾値を5に固定してデータ記録を行った結果と、図2右に示される保証値を用いて動的に変化させてデータ記録を行った結果を比較する。

図3に両結果を示す。グラフ中の点は記録された照度データを、線は記録量の推移を表す。固定閾値によるデータ記録では必要以上にデータが減少した。最終的に記録されたデータを記憶容量で割った値(以下、充足率と呼ぶ)は7.6%であった。一方提案方式では、ほぼ一定の記録頻度でセンサデータが記録された。必要記録時間を満たしたとともに、データの充足率は98.18%であり、記憶容量を十分に使い切った。閾値を固定して同様の充足率を得るためには、受信データ数および頻度、データ値の予測が必要であり、ほぼ不可能である。本実験により、閾値を動的に変化させてデータ記録を行う方式が有効であることが示された。

評価2: 閾値決定方式の違いによるデータ量の推移

本機構で用いたファジィ制御による閾値変化の有効性を、2値の述語論理による単純な制御を用いたものと比較し、評価する。前項で用いたデータおよび図2右の保証記述を利用して、2値の述語論理によって閾値を変化させる方式(以下2値方式と呼ぶ)を用いて記録を行った結果と、提案方式であるファジィ推論を元に閾値を変化させる方式を用いて記録を行った結果を比較する。

図4に両結果を示す。2値方式では、記録可能時間が必要記録時間を下回ると閾値を上げ、逆の場合は閾値を下げる。そのため、閾値が大きく上下して安定せず、保証値と最小値の間を繰り返し。さらに、記録終了40分前に受信データ6,024個を残して記録が終了する。これは、残り容量が小さい時点で記録頻度が低いために閾値を大きくしたが、その後記録量が増えたため、閾値を小さくしても十分なデータ減少が行えなかったためである。一方提案方式では、閾値が保証値と最小値の間の適当な値に収束し、記録推移がほぼ直線となった。また、残り容量が少なくなった150時間以降では、記録頻度が落ちて閾値を過剰に上げることはない。そのため、記録終了まで全てのセンサデータを記録できた。本実験の結果より、ファジィ制御による動的適応が、2値方式によるものと比べて有効に記録を行えたとともに、複数の制御要因を用いて正しく閾値設定を行えたことが示された。

評価3: アプリケーションの違いによるデータ量の推移

保証記述が異なるアプリケーションに適用し、記憶容量を十分に使用して記録できることを、図2左(ライフログ)と図2右(トレーサビリティ)の保証記述を用いたデータ記録実験により示す。ライフログの保証記述では類似フィルタの適応範囲が広く、トレーサビリティの保証記述では時間フィルタの適応範囲が広い。またライフログの保証記述では必要記録時間が2日間となっており、トレーサビリティのそれ(7日間)よりも短い。

図5上に、ライフログの保証記述を用いた場合の結果を、同図下に、トレーサビリティの保証記述を用いた場合の閾値変化と記録推移を示す。図5上では、主に類似フィルタの閾値が動的適応処理により変化しているのに対して、同図下では、主に時間フィルタの閾値が変化している。これらは保証記述における適応範囲を反映していると言える。表1に、本実験における充足率とフィルタごとのデータ減少率を示す。ライフログでは時間フィルタよりも類似フィルタでのデータ減少率が高く、トレーサビリティではその逆となった。保証記述において制約の緩い適応範囲が指定されたフィルタにおいて、データ減少率が高くなっている。これ

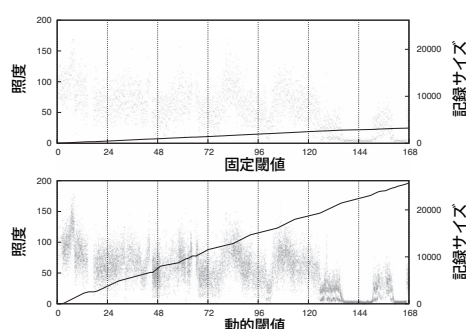


図3 固定閾値(上)と動的閾値(下)によるデータ分布と記録推移

Fig. 3 Static Threshold vs. Dynamic Threshold

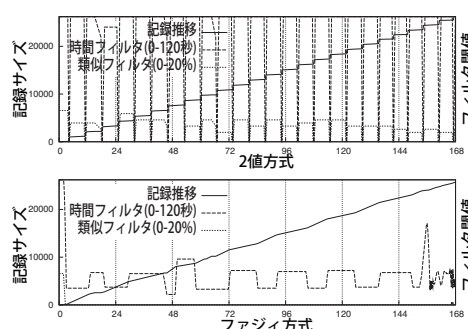


図4 2値方式(上)とファジィ方式(下)による閾値変化と記録推移

Fig. 4 Simple Method vs. Fuzzy Method

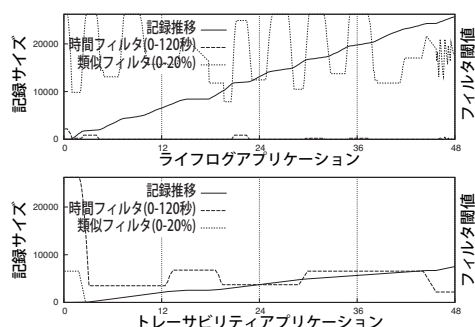


図5 異なるアプリケーションによる閾値変化と記録推移

Fig. 5 Comparison of Different Applications

表1 フィルタごとのデータ減少率

Table 1 Data Reduction Rate on Each Filters

	ライフログ		トレサビリティ	
	保証値	減少率	保証値	減少率
時間フィルタ	10 秒	8.43%	120 秒	95.38%
類似フィルタ	20%	81.71%	5%	8.93%
記録量(充足率)	25723 (98.13%)		25738 (98.18%)	

に対して、充足率は両実験とも 100% に近く、保証記録の違いに関係なく記憶容量一杯にデータが記録された。これらにより、提案方式における動的適応処理は、アプリケーションごとの保証記録を正しく処理したと言える。

5. おわりに

本稿では、必要記録時間と許容する記録データの必要品質を利用者が指定し、それを保証しながらセンサデータを記録する、保証付きデータ記録機構を提案し、実機上に実装した。小型で記憶容量が限られる MSDB において、大量のセンサデータを全て記録し続けることは難しい。そのため、ユーザが保証値と記録時間を指定し、それらの情報から記録機構が動的に記録品質を変化させることで、必要記録時間を満たすと同時に記録データを最大化させる。また、本機構を PC 上にも実装し、実データを用いた実験により、複数のフィルタを適用し閾値を動的に変化させて、記

憶容量が限られた場合でも保証値を満たしながらデータを記録できることを示した。

[文献]

- [1] Crossbow Technology, Inc. MPR500/510/520 MICA2-DOT Mote. <http://www.xbow.com/>.
- [2] A. K. Dey. Understanding and using context. *Personal Ubiquitous Computing*, 5(1):4–7, 2001.
- [3] P. B. Gibbons, B. Karp, Y. Ke, S. Nath, and S. Seshan. Irisnet: An architecture for a world-wide sensor web. *IEEE Pervasive Computing*, 2(4):22–33, 2003.
- [4] B. G. Haskell. Entropy measurements for nonadaptive and adaptive frame-to-frame, linear-predictive coding of videotelephone signals. *The Bell System Technical Journal*, 54(6):1155–1174, 1975.
- [5] S. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong. Tag: A tiny aggregation service for ad-hoc sensor networks. In *5th Symposium on Operating Systems Design and Implementation (OSDI'02)*, 2002.
- [6] Y. Yao and J. Gehrke. The cougar approach to in-network query processing in sensor networks. *SIGMOD Rec.*, 31(3):9–18, 2002.

由良 淳一 Jun'ichi YURA

慶應義塾大学大学院 政策・メディア研究科 特別研究助教。2002 年同大学同研究科修士課程修了。ユビキタスシステムの研究に従事。情報処理学会、日本データベース学会 各正会員。

中澤 仁 Jin NAKAZAWA

慶應義塾大学大学院 政策・メディア研究科 講師。2003 年 同大学同研究科博士課程修了、政策・メディア博士。ユビキタスシステム、センサネットワークの研究に従事。情報処理学会正会員。

徳田 英幸 Hideyuki TOKUDA

慶應義塾大学 環境情報学部 学部長、教授。1983 年 ウォータールー大学数学部計算機科学科博士課程修了、Ph.D. in CS。オペレーティングシステム、分散システム、ユビキタスシステムの研究に従事。ACM、IEEE、情報処理学会 各会員。