

## 無線センサを用いた歩行特性解析に適したデータ管理手法の設計

森 雅 智<sup>†</sup> 上 原 雄 一<sup>††</sup>  
齊 藤 裕 樹<sup>†</sup> 戸 辺 義 人<sup>†</sup>

センサネットワークから取得されたデータストリームの解析は、あらゆるセンサネットワークアプリケーションにとって最も重要な問題の一つである。従来の時系列データに対する解析方法は、センサネットワーク特有の問題により、単純にデータ解析手法として用いるのには問題がある。本稿では、人の歩行特性に着目し、複数の時系列データが密接に関連したデータストリームに対する解析手法として、リンク指向の新たなデータモデル LBSM を提案する。

### Design of Managing Data for Analyzing Walking Characteristics Using Wireless Sensors

MASATO MORI,<sup>†</sup> YUICHI UEHARA,<sup>††</sup> HIROKI SAITO<sup>†</sup>  
and YOSHITO TOBE<sup>†</sup>

For every sensor network application, it is one of the most important challenges to analyze a data stream acquired from its own sensor network. Some previous analytical method for time series data has any problems. In this paper, we propose a new link-based data model called LBSM to analyze the data stream with multiple correlated time series of data.

#### 1. はじめに

近年、センサネットワークに関する研究が世界中で活発に行われている。急速な無線通信デバイスの小型軽量化、及び高性能、省電力化に伴い、非常に小さな基板上に無線通信機能とセンサを実装した小型無線センサデバイスを用いた無線センサネットワークが、近い将来のユビキタスコンピューティング環境を支えるものとして期待されている。無線センサネットワークの適用範囲は、自然環境モニタリング、建造物の構造劣化診断、農業プラント、ヘルスケアモニタリングなど多岐に渡り、今後あらゆるアプリケーションが開発されていくと考えられる。

一方で、センサネットワークから取得されたデータの解析については、既存のデータ解析手法をそのまま適用できないという、いくつかのセンサネットワーク独自の問題がある。第1に、センサネットワークから取

得されたデータは連続データストリーム (continuous data stream) であり、絶えずデータが生成され続けるため、全てのデータを有限のストレージに蓄積し続けることは不可能であるという問題が挙げられる。既存のリレーショナルデータベースを用いて全てのデータを管理することは困難であるため、データストリームを専門に扱うシステムが必要となる。第2に、通信品質やセンサデバイスの状態によっては欠損値を含むという問題がある。欠損値については、どのように扱うべきかは解析手法に依存するが、FFT (Fast Fourier Transform) やウェーブレット変換といった多くの解析手法では、欠損値を含むデータを処理することができない。そのため、これらの解析手法を用いるには欠損値の予測計算が必要となる。第3に、アプリケーションによってはデータ解析処理にリアルタイム性が求められるという問題がある。危険検出のようなイベント検出型のアプリケーションにおいては、解析処理にはより軽量なものが求められるが、既存の解析手法には多くの計算時間が必要なものが多い。特に、多数のセンサから得られるデータストリーム間の関係を導き出すには膨大な計算時間が必要となる。

これらの問題を解決するため、我々は歩行特性解析に特化した複数データストリーム間の関係を単純化す

<sup>†</sup> 東京電機大学 情報メディア学科  
Department of Information Systems and Multimedia Design, Tokyo Denki University  
<sup>††</sup> 東京電機大学 工学研究科 情報メディア学専攻  
Department of Information and Media Engineering, Tokyo Denki University

るデータモデルとして LBSM (Link-Based Stream Modeling) を提案する。LBSM により、従来は多くの手順を踏む必要のあった複数時系列間の関係情報を容易に抽出することが可能となる。

本稿では、第 2 章で関連研究について述べ、第 3 章でセンサネットワークにおけるデータストリーム解析の問題を取り挙げる。その後、第 4 章で歩行特性考察のための予備実験について解説し、第 5 章で LBSM の設計について解説する。最後に、第 6 章でまとめとして今後の課題について述べる。

## 2. 関連研究

### 2.1 データ収集

無線センサネットワークを対象としたデータ収集技術については、既にいくつかのプラットフォームが存在する。これらのデータ収集技術は、無線センサネットワークにおけるノード管理と問い合わせ管理を提供する。ノード管理では、スリープ機能を用いた省電力化やマルチホップのネットワーク管理などを行い、センサネットワークの可用性を高める。問い合わせ管理では、アプリケーションからの必要な情報に対する問い合わせを管理する。このように、センサネットワーク依存の問題をプラットフォーム部分で吸収することで、アプリケーションからはセンサネットワークの種類や環境に関わらずデータを収集することができる。

以下に、近年注目されているデータ収集技術についての具体例を示す。

TinyDB<sup>4)</sup> や Cougar<sup>8)</sup> は、無線センサネットワークを一つのデータベースとして捉え、ネットワークに対して宣言的問い合わせ言語を発行することでデータを取得する。TinyDB は、無線センサネットワーク全体を一つの sensors テーブルとして定義し、各列にノード ID やセンサの値を割り当てている。データの問い合わせには TinySQL と呼ばれる SQL 風の言語を用い、必要なデータを取り出すことができる。集約関数として AVG や SUM などに対応しており、センサネットワーク内集約処理 (in-network aggregation) に対応している。また、センサネットワーク独自の問い合わせとして連続的問い合わせ (Continuous Query) に対応しており、データストリームを検索結果として受け取ることができる。

ZigBee<sup>15)</sup> は、アプリケーションフレームワークとしてセンサの値やアプリケーション変数を簡単に取得できる機能を持っている。各アプリケーションは一つ一つがアプリケーション・オブジェクトと呼ばれるオブジェクトに割り当てられ、KVP (Key Value Pair)

コマンド・フレームをアプリケーション・オブジェクトに対して送信することで、指定した属性値を取得、設定できる。また、複雑なコマンドを実行するためにユーザ定義可能な MSG (Message) フレームも用意されている。

### 2.2 データストリーム管理

データストリームの管理技術については、多くの研究が行われている。ここで、データストリームとはセンシングデータなどのデータが時間経過と共にオンラインで生成され、無限長のデータを持つものと定義する。

データストリームの具体例としては、センサネットワークからのセンシングデータの他に、WEB におけるアクセスログや業務システムのトランザクションデータなどが挙げられる。こうしたデータストリームを管理するデータストリーム管理システム (DSMS) としては、STREAM プロジェクト<sup>2)</sup> や Tapestry<sup>6)</sup>, Aurora<sup>1)</sup> などがある。

Aurora は、ユーザの指定した QoS プロファイルに基づいて、オペレータのスケジューリングを行う。データが高い到着率で到着し、処理が追いつかなくなったら、データの一部を読み捨てることで処理の効率化を図る。

多くのデータストリーム管理システムでは、到着データを全て保存するというはせずに、スライディングウィンドウと呼ばれる有限長のデータ領域に一時的に保管する。スライディングウィンドウに保管されたデータは、ウィンドウサイズを超えるデータが到着すると、古いデータから削除されていき、メモリを節約する。このことから、問い合わせに対する回答は近似値 (approximate answer) となる。

### 2.3 データストリーム解析

データストリームから有用な規則やパターンを探し出すデータストリームマイニング技術として、頻出値や異なり数の計算手法や、近似値を集約するといった手法が考案されている。<sup>3)9)</sup> いずれも、大量のデータストリームを処理するために、ランダムアクセスを行わない 1 パス (1-Pass) での解析に主眼が置かれている。

また、異常検出の分野では、センサデータストリームにおける周期性に着目し、サポート率と確信度を用いてデータストリームをモデル化し、非日常的なデータの検出を行うアルゴリズムが開発されている。<sup>11)</sup>

### 2.4 センサによる歩行解析

履き物に対してセンサを付け、位置情報や生体情報を取得するという試みが行われている。

Adidas 1 smart シューズでは、スニーカーにマイ

クロコンピュータおよびセンサを埋め込み、使用者が歩くことで足圧力データを取得する。<sup>5)</sup> 足圧力データは秒間数百回内蔵のマイクロコンピュータに送信され、疲労のレベル、地形、ペース、使用者の体格の大きさから使用者への影響を計算する。その計算結果を基にかかとのクッションを調節し、走りやすい環境をつくることを目的としている。

東京大学では、高齢者介護のための歩行分析用 GPS シューズの開発を行っている。<sup>12)</sup> GPS シューズは、GPS、足圧力センサ、無接点充電、無線信号伝送の機能をもっており、このシューズを用いることで、転倒防止や痴呆高齢者の徘徊行動追跡・痴呆と徘徊の関連性を発見することを目指している。

NaviGeta は ID カーペットと RFID タグを用いた、ユーザ位置検出システムである。<sup>10)</sup> ID カーペットとは RFID タグを埋め込んだカーペットである。そのカーペット上を RFID リーダを備えた NaviGeta を履き、歩くことでそのカーペットの RFID のタグ情報から使用者の位置検出を行う。

あしナビでは、超音波センサを備えたスリッパを使用し、利用者を誘導するシステムを提案している。<sup>13)</sup> 天井に設置された超音波センサと利用者の履いているスリッパの超音波センサを用いて、利用者の位置および身体の向きを測定する。測定した位置と向き情報を利用し、利用者の目的地までの移動方向を示していく。美術館などのナビゲーションをすることを目的としている。

### 3. 課 題

我々は無線センサネットワークから取得されたデータストリームを解析する上での課題を、データ取得に関するものとデータ解析処理に関するものの2つに分けて議論する。

#### 3.1 データ取得における課題

データ取得時において問題となるのは、センシングデータの時刻同期と欠損データの扱いについての2点である。

時刻同期について、一般的なアプリケーションでは全てのセンサが同じタイミングでセンシングしたデータを受信することを期待するが、センサノードは各々が省電力化のためにスリープタイマを持って動作するため、データの取得タイミングを完全に合わせることは困難である。データのサンプリング周期が30分や1時間といった長い周期であればこの点は無視できるが、ミリ秒単位の周期である場合、この点は無視できないずれとなる。

加えて、無線センサネットワークでは、電波状態や端末の電池切れなどの理由により、欠損データの発生が避けられない。時刻のずれたデータや欠損を含むデータは、解析処理に大きな制限を加えてしまう。

時刻のずれたデータや欠損を含むデータは、解析処理に大きな制限を加える。例えばFFTやウェーブレット変換は、等間隔の欠損の無い有限長のデータを必要とするため、得られたデータストリームにそのまま適用することはできない。

#### 3.2 データ解析における課題

無線センサネットワークから得られるデータストリーム解析については、一般的なデータマイニングと異なる点が3点存在する。

1つはデータアクセスの問題である。前章で述べた通り、データストリームの解析においては大量のデータを扱うため、ランダムアクセスを避け、1パスで行うことができる解析手法が求められる。

2つ目に、データ解析のリアルタイム性における問題である。既存の時系列解析手法では、計算に数時間かかることもあるため、危険察知や災害予測など、緊急性の高いセンサネットワークアプリケーションではそのような手法を用いることができない。

最後に、複数のデータストリーム間の関係情報検出の問題が挙げられる。センサネットワークでは環境などに多数のセンサを配置してデータをセンシングするため、現実世界でのある事象が複数のセンサに影響を及ぼすことが普通である。一方、各センサノードはそれぞれがデータストリームを出力するので、一つのデータストリームにのみ着目することは妥当ではなく、複数のデータストリームを監視し、互いの関係に注目する必要がある。

### 4. 予 備 実 験

我々は、歩行時における足圧力データの特徴を掴むため、予備実験を行った。

#### 4.1 実験概要

足圧力センサを実装したスリッパを用い、数種類の歩き方でデータを取り、データの特性を分析する。

実験に用いたスリッパは、市販のスリッパにセンサノードとしてMICA2 MOTE<sup>14)</sup>を実装したものである(図1)。センサスリッパには2つの圧力センサが接続しており、前部と後部の足圧力データを取得することができる。

データ取得にはTinyDBを用い、サンプリング周期128msでセンシングを行った。各センサノードは1ホップでベースステーションと通信を行い、サンブ

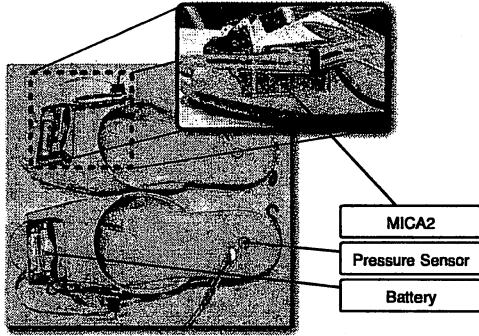


図 1 センサスリッパ

リング周期毎にベースステーションへセンシングデータを送信する。TinyDB への問い合わせ結果はベースステーションのローカル PostgreSQL データベースへと格納した。

実験環境では、通常歩行、すり足歩行、前傾歩行、片足引きずり歩行の 4 種類の歩行状態のデータを取り、解析を行った。片足引きずり歩行とは、足を捻挫したときのような歩き方で、本実験では左足を引きずって歩いたデータを用いた。

## 4.2 実験結果

実験の結果、いくつかの知見が得られた。以降、説明のために左前のセンサを LF とし、以下左後を LB、右前を RF、右後を RB と記述する。

### 4.2.1 各圧力センサ同士の関係

通常歩行において、各圧力センサにかかる力の極大値の山は LB → LF → RB → RF → LB → ... という順序で現れ、前傾歩行時には LF → RF → LF → ... という順序が見られた。対して、すり足歩行時には、一見規則正しいデータの流れは見つけられなかった。片足引きずり歩行の際は、引きずっている側の足のデータには規則性が見られなかったが、正常な側の足のデータは通常歩行と同じ順序の規則性が見られた。

### 4.2.2 圧力センサの生データ

圧力センサから取得された生データの値に着目すると、同じ通常歩行のデータの中でも、重心がかかった際の値のピーク値には最大 2 倍程度の大きなばらつきがあることが確認できた。

原因としては、実験環境がスリッパという足位置が確実に固定されないものを使ったために、毎回の踏み出しごとに微妙に足の接地点がずれ、圧力センサへの重心のかかり具合が変化したということが考えられる。

また、片足だけに着目すると、爪先部分が離れてももう片方の足に重心がかかっている間は足圧力値がほぼ

表 1 データ欠損の連続性

連続数	発生回数
1	44
2	6
3	8
4	3
5	2
10	2
12	1

ゼロになっていることが確認できた。

### 4.2.3 データ欠損率

TinyDB の自動生成するシーケンス番号を元に、パケットの欠損数をカウントしたところ、500 サンプル中 120 から 160 サンプルが何らかのエラーにより取得できず、パケットを取りこぼしていることが確認できた。

欠損発生連続性に着目し、連続欠損数をカウントしたものが表 1 である。少数の欠損が頻繁に発生しているのが確認できる一方で、連続 10 回を超える長い欠損も数度発生していることが確認できた。

データ欠損の原因について、サンプリング周期が短いことが原因であると考え、サンプリング周期を 512 ms に設定して欠損発生率を分析したところ、パケット落ちの回数は減少した。特に、連続した欠損はほとんど見られなくなった。

## 4.3 考察

実験結果より、足圧力センサを用いた歩行特性情報の抽出には、各圧力センサ間の足圧力の極大値の動きを見れば良いのではないかとこの予測を立てることができた。一方、生データの値はばらつきが大きく、生データの値から細かい情報を取り出すのは困難であると予想される。

データ欠損については、サンプリング周期を短く設定した場合に多くの欠損が発生していることから、通信路上で衝突が発生していることが予想される。欠損率を下げるためにはサンプリング周期を長くすることが簡単な解決策であるが、歩行時における足圧力データの動きは短い変化の中で変化するため、単純にサンプリング周期を長くするだけでは意味のあるデータが取れなくなる恐れがある。

## 5. LBSM の設計

予備実験を通して、歩行特性解析におけるデータ管理においては、主に下記の 2 つの問題があることが判明した。

- 生データ値のばらつきの問題

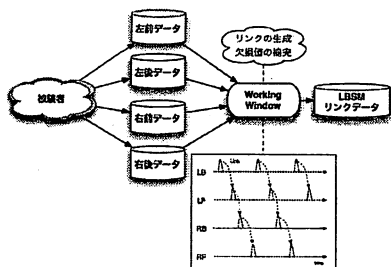


図2 LBSM システム構成

### ● データ欠損の問題

一方で、歩行特性情報を抽出するには各足圧力センサ値の極大値の動きに注目すれば良いことが分かっている。

我々は、これらの事実を踏まえ、歩行特性に特化したデータ管理手法として LBSM を提案する。

### 5.1 特徴

LBSM は下記の特徴を持つ。

- 生データをセンサ値の極大値から極大値へのリンクに要約し、歩行特性の抽出を容易にする
- 生データを必要としないため、送信するデータ量が少ない
- データ欠損時における他のリンクからのデータ補完が可能である
- FFT などの既存手法に比べ、少量データから歩行特性を抽出できるため、リアルタイム処理が可能である

圧力センサの生データではなく極大値を利用することで、生データのばらつきを無視することができる。さらに、センサノードから極大値情報だけを送信させることにより、データ送信の回数を減らした。また、データ欠損時には、過去のリンク情報を参照することで欠損値を補完することができる。

LBSM のシステム構成を図 2 に示す。LBSM は、各圧力センサからのデータストリームを作業ウィンドウ (Working Window) へ入力する。作業ウィンドウは、直近数秒から数分のデータを保持しており、リンクの生成や欠損値の補完に必要なデータを保持する。作業ウィンドウ内で確定したリンクは、リンクデータとして出力される。また、規定の時間を過ぎたデータは作業ウィンドウから破棄される。

### 5.2 リンクの生成

LBSM におけるリンクは、あるセンサの極大値から他のセンサの極大値への繋がり情報を表す。図 3 はリンクのイメージ図である。各センサには、LF や LB

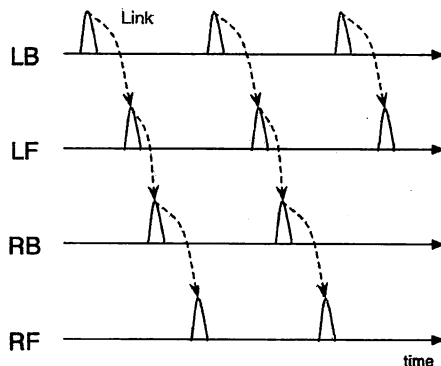


図3 LBSM リンク (通常歩行)

といった ID が定義されており、リンクはそれらの ID の組によって表され、各極大値での時刻を持つ。

以下に LBSM 生成の擬似コードを示す。

```

A pseudocode of the LBSM algorithm.
type SensorEvent = class
  ID : Longint;
  Sequence : Longint;
  Time : Timestamp;
end;

type Link = class
  List : List;
  procedure add;
end;

procedure Link.add;
begin
  /* Link Add */
end;

procedure LBSMProcessMain(
  eventStream : EventStream ;
  var LinkOutputStream
);
begin
  var aLink : Link;
  var aEvent : SensorEvent;

  aLink := Link.Create;
  while eventStream.hasNext do
    begin

```

```

aEvent := eventStream.Next;
if aEvent.ID = EVENT_FIRST then
begin
  LinkOutputStream.write( aLink );
  aLink := Link.Create;
end;
aLink.add( aEvent );
end;
end;

```

各極大値は SensorEvent として EventStream から入力され、リンクは規定の SensorEvent(EVENT\_FIRST) から SensorEvent までの連続したリンク情報を持つ。EVENT\_FIRST の SensorEvent を受け取った時点で、構築中のリンク情報を出力し、新しいリンクを生成し直す。

### 5.2.1 リンクの種類

リンクは、片足内リンクと両足間リンクに分けられる。片足内リンクとは、片足のセンサだけで作られるリンクである。片足内リンクの状態によって、正常に足が着地し、踏み出されたかどうかを検出することができる。両足間リンクとは、両足間をまたがるリンクである。両足間リンクは、片足内リンクだけでは得られない情報を表現する。例えば、LF と RF 間のリンクに着目することで、歩調や歩数を検出することができる。

### 5.3 リンクの解析

生成されたリンクを解析することで、歩行特性情報を抽出する。例えば、予備実験で明らかになった通り、LB→LF→RB→RF というリンクが検出されれば、それは通常歩行の特性を示していると判断する。

### 5.4 欠損値の補完

各々のセンサから送信されるデータには、センサノードが割り振った一意の送信シーケンス番号が振られている。シーケンス番号はセンサノードからのデータ送信ごとに1ずつ加算されていく。このシーケンス番号を監視し、シーケンス番号の間隔が2以上開いたとき、データ欠損が検出される。

欠損値が検出されると、作業ウィンドウ内の過去のリンクデータを検索し、欠損部分のデータを予測し、補完する。

## 6. まとめ

我々は、予備実験を元に歩行特性に特化した新たなデータ管理手法 LBSM を提案した。

今後は実装を行い、実験と評価を行っていく。

## 参考文献

- 1) D. Abadi, D. Carney, U. Cetintemel, M. Cherniack, C. Convey, S. Lee, M. Stonebraker, N. Tatbul and S. Zdonik, Aurora : A New Model and Architecture for Data Stream Management, The VLDB Journal, Vol.12, No.2, pp.120-139 (2003)
- 2) B. Babcock, S. Babu, M. Datar, R. Motwani, and J. Widom. Models and Issues in Data Stream Systems, ACM PODS pp.1-16 (2002)
- 3) P. B. Gibbons and Y. Matias. Synopsis Data Structures for Massive Data Sets, ACM SIAM pp.909-910 (1999)
- 4) S. Madden, W. Hong, J. Hellerstein, and M. Franklin. TinyDB: A Declarative Database for Sensor Networks, <http://telegraph.cs.berkeley.edu/tinydb>.
- 5) Z. L. Rubin, Adidas 1, The Smart Shoe of the Future <http://www.personal.psu.edu/users/z/1/zlr102/>
- 6) D. Terry, D. Goldberg, D. Nochols, and B. Oki. Continuous Queries over Append-Only Databases, ACM SIGMOD pp.321-330 (1992)
- 7) Y. Uehara, T. Uchiyama, M. Mori, H. Saito and Y. Tobe. Always-on Karte: A System for Elderly People's Healthcare Using Wireless Sensors, INSS 2006 (2006)
- 8) Y. Yao and J. Gehrke. Query Processing for Sensor Networks, ACM SIGMOD pp.491-502 (2003).
- 9) 有村 博紀, 喜田 拓也. データストリームのためのマイニング技術, IPSJ Magazine Vol.46 No.1 Jan.2005 (2005)
- 10) 椎尾一郎, RFID を利用したユーザ位置検出システム, 情報処理学会研究会報告, 00-HI-88, 2000.5.12, pp.45-50 (2000)
- 11) 倉光 君郎. センサーデータにおける周期性と異常の検出, 情報処理学会研究報告, 2004-UBI-4, Vol.2004, No.39, pp.7-10 (2004)
- 12) 辻昌彦, 臼井理, 板生清, 保坂寛, 佐々木健, 歩行分析用 GPS シューズの開発, 2004 年度精密工学会春季大会学術講演会講演論文集, pp. 1211-1212 (2004.3)
- 13) 山崎俊作, 伊藤友隆, 河田恭兵, 生天目直哉, 小原栄一, 高橋元, 中澤仁, 徳田英幸, あしナビ: 主張するアンビエントナビゲーションシステムの試作, インタラクシオン 2006 (2006.3)
- 14) MICA2 MOTE, <http://www.xbow.jp/motemica.html>
- 15) ZigBee Alliance. <http://www.zigbee.org>