

センシングデータの分散管理における プロアクティブデータ予測手法

井邊 研吾^{†1} 横田 裕介^{†2} 大久保 英嗣^{†2}

我々は、無線アドホック通信を用いて、センシングデータを分散管理するネットワークストレージシステムである P2P データポットシステムの開発を進めている。本システムでは、各端末間でアドホック通信を行うことにより、インフラレスな環境で柔軟にネットワークを構成することが可能である。しかし、アドホック通信を用いているため、遠方の端末に問合せを発行した場合に、ホップ数が増加し応答性が低下する。この問題を解決するため、センシングデータの利用を予測し、事前に当該の端末へ転送するプロアクティブデータ転送・検索機構を提案している。転送されたデータを利用して問合せに回答することで、応答性の向上を図る。本稿では、本機構におけるセンシングデータの利用予測手法について述べる。

A Proactive Data Prediction Technique for a Distributed Sensor Data Management System

KENGO IBE,^{†1} YUSUKE YOKOTA^{†2} and EIJI OKUBO^{†2}

We have been developing P2P Data Pot system, which is a kind of a distributed network storage system that manages sensing data on MANET environments. It organizes a flexible network using ad hoc communications among terminals without existing network infrastructure. However queries to distant terminals require more hop counts to communicate and decrease responsiveness because of characteristics of MANET. To solve this problem, we propose a proactive data transfer and search mechanism. It gains responsiveness by making responses to queries using transferred data in advance. In this paper, we describe a prediction mechanism of sensing data.

^{†1} 立命館大学大学院理工学研究科
Graduate School of Science and Engineering, Ritsumeikan University
^{†2} 立命館大学情報理工学部

1. はじめに

我々は、マイクロストレージを用いたセンシングデータの分散管理システムである P2P データポットシステムの開発を進めている¹⁾。本システムは、環境観測を行うセンサネットワーク、センサネットワークで得られたセンシングデータを蓄積および管理する P2P データポット、P2P データポットに対してセンシングデータの問合せを行うアプリケーションで構成される。センサネットワークと P2P データポットは 1 対 1 で対応し、センサネットワークで取得したデータは、対応する P2P データポットに蓄積される。また、P2P データポット間は無線アドホックネットワークで接続され、アプリケーションは直近の P2P データポットに問合せを発行する。アプリケーションから問合せを受けた P2P データポットは、対象となるデータを保持する各々の P2P データポットに対して問合せを転送し、結果の集約を行う。この際、一連の通信処理は、無線アドホックネットワーク上で行われる。このため、ネットワーク上の離れた P2P データポットに対する問合せが発生した場合、応答性の低下が問題となる。

このため、アプリケーションが要求する問合せの予測を行い、その問合せに対応するデータを事前に転送することで、応答性の向上を図る。ここで、事前に転送するデータをプロアクティブデータと呼び、事前に転送するデータを決定することをプロアクティブデータの予測と呼ぶ。これを実現する機構として、我々は、プロアクティブデータ転送・検索機構を提案してきた²⁾³⁾。提案機構では、これまでに発行された問合せの履歴を基に、今後利用されるセンシングデータの予測を行い、事前に対応するデータの転送を行う。アプリケーションからの問合せに対し、事前に転送したデータによって対処することができた場合、問合せ結果の転送に要する時間が短縮され、応答性の向上が可能となる。本稿では、プロアクティブデータ転送・検索機構が必要となるセンシングデータの利用予測手法について述べる。

以下、2 章でプロアクティブデータ転送・検索機構について述べ、3 章で、センシングデータの予測手法について述べる。4 章でアルゴリズムについて述べ、5 章で、関連研究について述べる。6 章で本稿のまとめを行い、今後の予定について述べる。

2. プロアクティブデータ転送・検索機構

本機構では、これまでに発行された問合せを基に、今後発行される問合せの予測を行い、

College of Information Science and Engineering, Ritsumeikan University

事前に問合せ結果の転送を行う。予測が一致し、アプリケーションからの問合せが事前に転送したプロアクティブデータで対応可能な場合は、データ転送のホップ数が減少し応答性が向上する。本章では、P2P データポットの役割の分類、データ転送および検索機構の概要、およびアプリケーション例について述べる。

2.1 P2P データポットの役割

プロアクティブデータ転送・検索機構における P2P データポットの役割は、次の 3 つに分類される。

ゲートウェイデータポット

アプリケーションから発行された問合せを受信するデータポットである。アプリケーションが発行する問合せは、「自身の周囲の半径 100m の区域の平均温度」などのように、主に位置や範囲情報を用いて指定することを想定している。従って、ユーザは P2P データポットシステム内部のネットワーク構成を意識せず、システム全体を一つの仮想的なデータベースとみなして問合せを発行することが可能である。ゲートウェイデータポットは、このようなアプリケーションからの問合せを対象データを持つデータポット群が直接処理可能な問合せに変換する。変換した問合せを各データポットに転送し、戻ってきた問合せ結果を受信・集約してアプリケーションに返送する。

データソースポット

アプリケーションの問合せ対象となるデータを保持するデータポットである。ゲートウェイデータポットから問合せを受信し、結果を返送する。また、自身に対して発行される問合せの予測、および自身が転送したプロアクティブデータの削除といった制御を行う。

プロアクティブデータポット

事前にデータソースポットより転送されたプロアクティブデータを保持するデータポットである。ゲートウェイデータポットからの問合せに、データソースポットの代わりに返答することで、応答性の向上を実現する。

本機構では、まずアプリケーションからの問合せをゲートウェイデータポットが受信する。ゲートウェイデータポットからデータソースポットに要求が転送され、問合せの履歴が保存される。プロアクティブデータの予測はデータソースポットが行い、プロアクティブデータポットに事前に転送する。1 つの P2P データポットが複数の役割を担う場合もある。

2.2 データ転送機構

プロアクティブデータ転送機構は、クエリの履歴保存、クエリの予測、予測したデータを

表 1 問合せ履歴の例

Table 1 Example of a query history record.

Query ID	Q2-5
Query	SELECT * FROM mts310_sample LIMIT 10
Gateway ID	DP8
Application ID	AP8-3
Date	2010-01-19 17:11:16

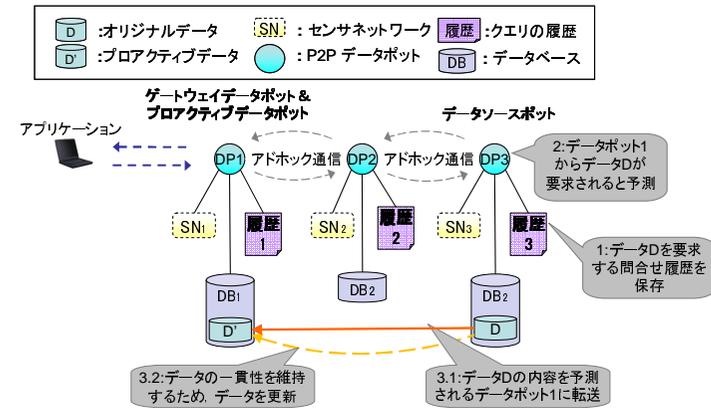


図 1 プロアクティブデータ転送機構の動作例

Fig. 1 Operation example of a proactive data transfer mechanism.

事前に転送するプロアクティブデータ転送の三つの機能から構成される。図 1 に動作例を示す。

まず、図 1 の 1 では、ゲートウェイデータポットから受けたクエリの内容を保存する。保存内容は、各クエリを識別する Query ID、ゲートウェイデータポットを識別する Gateway ID、問合せ内容である Query、アプリケーションを識別する Application ID、問合せを受けた時刻である Date で構成される。図 1 の履歴 3 に保持される問合せ履歴の例を表 1 に示す。これは、P2P データポット DP1 がアプリケーションからの問合せを受け、DP3 に問合せを転送したときの例である。問合せごとに生成される Query ID は、システム全体でユニークな ID である。Query は、P2P データポットが受信した SQL 文の内容である。Application ID は、各アプリケーションインスタンスに発行される ID であり、ゲートウェ

表 2 プロアクティブデータのメタデータの例
Table 2 Example metadata of proactive data.

Table Name	proactive-<1>
Proactive ID	DP8-1
DataSourcePot ID	DP8
Query	SELECT * FROM mts310_sample LIMIT 10;

イデータポットがアプリケーションを識別するために用いられる。Date は、年月日と時刻を保持している。

次に、図 1 の 2 ではこの履歴から今後発行される問合せの予測を行う。予測内容は、対象となるデータと転送先となるプロアクティブデータポットである。予測を基に、事前にデータの転送を行う。図 1 では、アプリケーションから発行されたデータ D を要求するような問合せが繰り返し発行されており、これがゲートウェイデータポットである DP1 を経由してデータソースポットである DP3 に転送されている。DP3 上ではこの D を要求する問合せの履歴が保存されている。この場合、DP3 は将来 D を要求するような問合せが DP1 に発行されると予測する。そして、図 1 の 3.1 のように、DP1 をプロアクティブデータポットとして、事前にデータの転送を行う。

プロアクティブデータを受け取ったプロアクティブデータポットは、これを格納し、そのメタデータを生成する。これは、プロアクティブデータの保存場所と予測結果で構成される。メタデータの例を表 2 に示す。プロアクティブデータは、問合せの予測ごとに問合せ結果のデータベーススキーマが異なる。従って、予測とプロアクティブデータを保存するテーブルは一对一に作成する必要がある。このためメタデータには、プロアクティブデータが保存されているテーブル名 (Table Name) が含まれる。メタデータに記述されるプロアクティブデータの予測結果は、プロアクティブデータを識別する Proactive ID と、データソースポットを識別する DataSourcePot ID、予測された問合せ内容 Query により構成される。このメタデータを他の P2P データポットに配布し、プロアクティブデータの検索に利用する。

プロアクティブに転送するデータは、単なるキャッシュデータとは異なり、時間により変化する可能性がある。このため、データの一貫性維持が必要となる。例えば、図 1 にある問合せ対象として予測されたデータ D に対し、 SN_3 から新たにデータが追加される場合、データ D は変化する。この場合、事前に転送したデータ D' とのデータの一貫性が失われてしまう。そこで、図 1 の 3.2 のように、新たにデータ D に追加されたデータも転送することでデータの一貫性を維持する。これにより、問合せ結果が時間の経過につれ変化するよ

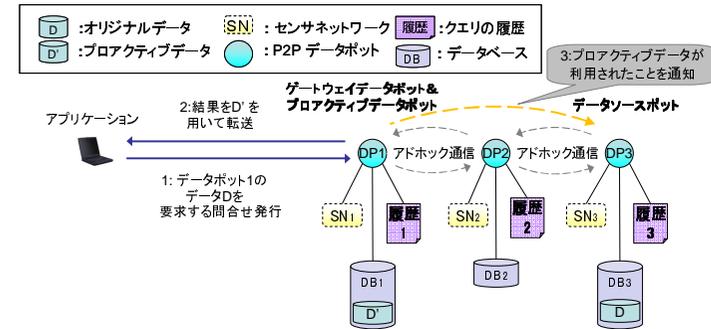


図 2 プロアクティブデータ検索機構の動作例
Fig. 2 Operation example of a proactive data search mechanism.

うな場合でも、転送したデータ D' を用いて問合せに対応することが可能になる。

プロアクティブデータには予測が失敗する場合や、予測には成功した場合でも時間経過により利用されなくなる場合が発生する。このため、一定の期間プロアクティブデータが使われていない場合には、これを削除し、転送も停止する必要がある。そこで、プロアクティブデータに有効期限を設ける。プロアクティブデータを利用して問合せに対応した P2P データポットは、データを利用したことを、プロアクティブデータの転送元の P2P データポットに通知する。これにより、通知を受けた P2P データポットは、そのプロアクティブデータが実際に利用されていることがわかる。もしこの通知が有効期限までに来なければ、プロアクティブデータを削除し、転送も停止する。有効期限は、プロアクティブデータの予測を行った P2P データポットが設定する。

2.3 データ検索機構

プロアクティブデータ検索機構は、配布されたメタデータを利用しプロアクティブデータの検索を行う。メタデータは、転送元の P2P データポット、問合せ内容、データ格納先の情報で構成される。転送機構によってメタデータは配布され、他の P2P データポットと共有することになる。

プロアクティブデータの検索は、次のように行われる。まず、ゲートウェイデータポットは、自身が保持しているメタデータを用いて、自身または近辺の P2P データポット上にアプリケーションから要求されたデータが存在するか否かを検索する。もし、存在すれば、プロアクティブデータを取得するための問合せを発行する。もし存在しなければ、データソー



図3 P2P データポットを利用するアプリケーション例
Fig. 3 Application example using P2P Data Pot.

スポットに問合せを転送する。図2にプロアクティブデータの利用例を示す。事前に、DP3が予測を行い、データ D をプロアクティブデータとして、DP1 に転送している場合を想定する。アプリケーションから、データ D を要求する問合せが発行された場合、ゲートウェイデータポットは、問合せがプロアクティブデータで対応可能かを判断するため、メタデータの検索を行う。この例では、対応可能であるため、自身が持つプロアクティブデータ D' を用いて、アプリケーションに問合せ結果を送信する。その後、データソースポットに対し、プロアクティブデータ D' が利用されたことを通知する。この通知によって、プロアクティブデータで問合せに回答した場合でも、データソースポットに問合せ履歴を残すことが可能となり、プロアクティブデータの利用率を把握することが可能となる。

2.4 適用するアプリケーション例

P2P データポットシステムを利用するアプリケーションの1つとして、環境観測システムがある。図3のように、山間部にあるP2P データポットは、地滑りの予測のために、定期的に地中水分量取得の問合せを受ける。また、登山者がこれから向かう山頂部の気温を知るため、山頂部の気温取得の問合せを受ける。このように、P2P データポットは、複数の目的や性質の異なるアプリケーションから問合せを受ける状況を想定している。

このような環境観測システムの特徴として、周期的にデータ取得の問合せが発行される点あげられる。例えば地滑りを予測するアプリケーションは、数分毎にP2P データポットに対し、地中水分量取得の問合せを発行し、取得したデータを基に地滑りの予測を行い、地滑りの危険がある場合には周囲に知らせるといったことが考えられる。また、登山者が向かう

場所の気温を通知するアプリケーションでは、登山者は山頂に向かって移動しながらこれから向かう場所の気温を取得し、天候の変化を確認することが考えられる。この場合、登山者が問合せを発行する位置が変化しながら移動先の気温情報を繰り返し要求するという振舞いが想定される。

本手法では、このような繰り返し行われる問合せに注目し、プロアクティブデータの予測を行う。

3. センシングデータの予測手法

P2P データポットシステムでは、基本的にすべてのデータポットがデータベースを保持している。応答性向上のためにすべてのデータベースのレプリケーションをすべてのデータポットに配置した場合、膨大な数のレプリケーションを管理する必要が生じるため、無線ネットワークが破たんすると考えられる。そのため、本システムでは、将来的に要求される可能性があるデータを予測し、必要なものだけを転送する機構が必須となる。

本手法では、各データソースポットが所持している問合せの履歴を用いて、繰り返し行われる問合せを発見し、プロアクティブデータの予測を行う。プロアクティブデータには、繰り返し発行される問合せの間隔、問合せ結果のデータサイズ、転送先のデータポットまでの距離、ネットワーク負荷などから決定される重要度が設定される。これを用いて、プロアクティブデータの転送や各P2P データポットがプロアクティブデータの置換を行うか否かを判断する。また、プロアクティブデータを用いて問合せに回答した場合には、転送元のデータソースポットにそれを通知することで重要度を変化させることが可能である。重要度を変化させることで、利用率の低いプロアクティブデータの転送の停止や削除を行うことができる。

以下、予測対象となるデータの種類の種類、予測で用いる履歴の構成と共有の必要性、問合せパターンの分類、および具体的な予測の方法について述べる。

3.1 予測の対象となるデータ

本手法での予測対象は、アプリケーションに利用される可能性があるセンシングデータの集合である。1つのセンシングデータは、センサから取得した値、取得時刻、取得したノードの情報の組で構成される。アプリケーションは、P2P データポットに対し問合せを発行することで、その問合せ内容に合致するセンシングデータの集合を取得する。問合せ内容が同じであっても、その問合せが発行される時刻や位置、対象データの動的な更新などによって、取得されるデータは毎回異なる内容となる可能性がある。提案するシステムでは、アプ

リケーションの要求を正確に予測することが目標であるため、実データ集合そのものではなく、発行される問合せを予測するものとする。これにより、対象データが動的に変化するような状況への対応が可能になる。この予測された問合せ内容に合致するデータを転送することで、応答性の向上を実現する。

センシングデータの取得方法としては、センサネットワークから得られるストリームデータをリアルタイムで連続的に取得する方法と、P2P データポット上のデータベースにリレショナルデータとして蓄積されたデータを取得する方法の 2 種類が考えられる。提案するシステムでは、ストリームデータの取得における応答性の向上は対象としていないため、後者の場合のみを対象とする。蓄積されるセンシングデータは、更新がなく、新規データの挿入が頻繁に繰り返される。すなわち、過去に取得されたデータの値が変わることはないという特徴がある。

3.2 予測に用いる履歴

問合せの履歴から繰り返し発行されている問合せを見つけることで、問合せの予測を行う。問合せの履歴は、クエリを一意に識別するための Query ID、ゲートウェイデータポットを識別するための Gateway ID、問合せ内容である Query、アプリケーションを一意に識別するための Application ID、問合せを受けた時刻である Date の 5 つの情報からなる (表 1 参照)。履歴は、データソースポットが記録している。

履歴を用いた予測を行う場合は、繰り返し発行される問合せが対象とするデータソースポットが変化するか否かによって予測に必要な履歴が異なる。データソースポットが変化しない場合、単体のデータソースポットの履歴から予測することが可能である。一方、変化する場合、すべてのデータソースポットの履歴を共有し、複数の異なるデータポット間で繰り返し発行されている問合せを発見する必要がある。このような履歴の共有が必要となる問合せは、対象となるデータソースポットが動的に変化する問合せである。

3.3 問合せのパターン

アプリケーションが発行したある問合せは、対象のデータソースポットという観点から次の二つのパターンに分類することができる。

- (1) 単一データソースポットを対象とする問合せ：例えば、山頂の気温データの取得のような、特定の位置のデータを要求するような問合せが該当する。
- (2) 複数のデータソースポットを対象とする問合せ：例えば、山道周辺の気温データすべてを取得するような、ある程度広範囲にわたるデータを要求するような問合せが該当する。

また、アプリケーションが同一内容の問合せを繰り返し発行する場合、その問合せが対象とするデータポットの変化という観点から次の二つのパターンに分類することができる。

- (3) 対象となるデータソースポットが変化しない問合せ：例えば、山頂などの特定の区域の気温データを取得するような、問合せの対象となる区域が固定される問合せ
- (4) 対象となるデータソースポットが動的に変化する問合せ：例えば、現在地から半径 100 メートル以内の区域の気温データを取得するような、問合せを発行する位置に依存して対象区域が変化するような問合せ

このほかに、問合せの結果と問合せ時刻の依存関係の観点から次の二つのパターンに分類することができる。

- (5) 問合せ時刻に依存しない問合せ：特定の期間のデータの参照など、時間経過により結果が変化しないもの
- (6) 問合せ時刻に依存する問合せ：最近の 100 件の生データ参照や、最近の 100 件のデータの最大値や平均値といった集約演算結果の参照など、時間経過により結果が変化するもの

以上から、プロアクティブデータ転送・検索機構では、3 つの観点の問合せパターンに対応する必要がある。(1)~(4) のパターンには、問合せの予測手続きの中で対応する。予測方法については、3.4 節で述べる。

一方、(5)、(6) のパターンには、プロアクティブデータの取り扱い方法によって対応する。(6) のような場合は、一般的なレプリケーション機構と同様に、その元データを管理する P2P データポットが更新を検知し、更新後のデータを継続的に転送先データポットに送信する。これにより、データの一貫性を維持し、問合せ時刻に依存する問合せに対応する。

3.4 予測方法

前節の検討結果にもとづき、データソースポットが固定の場合と、データソースポットが動的に変化する場合の問合せそれぞれにおける予測方法について述べる。

3.4.1 特定の P2P データポットをデータソースポットとする場合

複数か単数にかかわらずデータソースポットが固定である場合、個別のデータソースポットの履歴から今後発行される問合せの予測が可能である。例えば、同じデータソースポットに対し、繰り返し問合せが発行される場合、そのデータポットには表 3 のような履歴が蓄積される。ここでは、Query ID Q1-1, Q1-3, Q1-5 はいずれも同一内容の問合せが、DP1 から発行されていることが分かる。この問合せは繰り返し行われ、今後も発行されると予測することが可能である。

表 3 データソースポットの履歴例

Table 3 Example of query history records of a data source pot.

Query ID	Gateway ID	Query	Application ID	Date
Q1-1	DP1	現在の地中水分量	AP1-1	2010/12/20/12:00:10
Q1-2	DP2	現在の湿度	AP2-3	2010/12/20/12:10:15
Q1-3	DP1	現在の地中水分量	AP1-1	2010/12/20/12:30:13
Q1-4	DP3	過去 30 分間の温度データ	AP3-4	2010/12/20/12:46:13
Q1-5	DP1	現在の地中水分量	AP1-1	2010/12/20/13:01:13

繰り返し行われる問合せの発見は、過去一定期間に同一もしくは類似性の高い問合せが、何回発行されたかをカウントすることで行う。過去一定期間に発行された問合せのみに限定することで、以前は頻繁に発行されたが、最近では発行されなくなった問合せや、発行間隔が極端に長い問合せ等を除外する。

3.4.2 データソースポットとなる P2P データポットが動的に変化する場合の予測

データソースポットとなる P2P データポットが動的に変化する場合、問合せ履歴を共有することで、繰り返し発行される問合せを発見することが可能となる。ここでは、ある種類のアプリケーションにおける問合せ対象であるデータソースポットの動的な変化のパターンを発見し、今後同一種類のアプリケーションが出現した場合に問合せ対象がどのように変化するかを予測することを目標とする。

例えば、2.4 節のアプリケーション例で述べた登山者が天候を確認するようなアプリケーションでは、複数の登山者が個別に移動しながら、問合せを繰り返し発行すると考えられる。登山者 A は、山道の入口付近および入口付近から 3km ほど進んだ地点それぞれにおいて、現在地から移動方向に沿って 3km 先の地点の周辺気温情報を取得する問合せを発行するものとする。他の登山者も登山者 A と同一のアプリケーションを用いて行動するとすれば、事前にデータソースポットとなりうるデータポットから 3km 手前の地点に存在するデータポットに気温データを転送しておくことで、応答性を向上させることができる。これを実現するため、同一種類のアプリケーションによる問合せを発見し、データソースポットの移動パターンを予測する必要がある。

しかし、山道の入り口付近や、3km 先で各登山者が発行する気温取得の問合せを受けるゲートウェイデータポットやデータソースポットが必ず同一の P2P データポットであるとは限らないため、登山者が利用するゲートウェイデータポットやデータソースポットを P2P データポット単位で予測することは難しい。ある程度問合せ発行位置が異なった場合でも予測結果を適用できるようにするため、必ずしも同一のデータポットでない場合でも、一定の

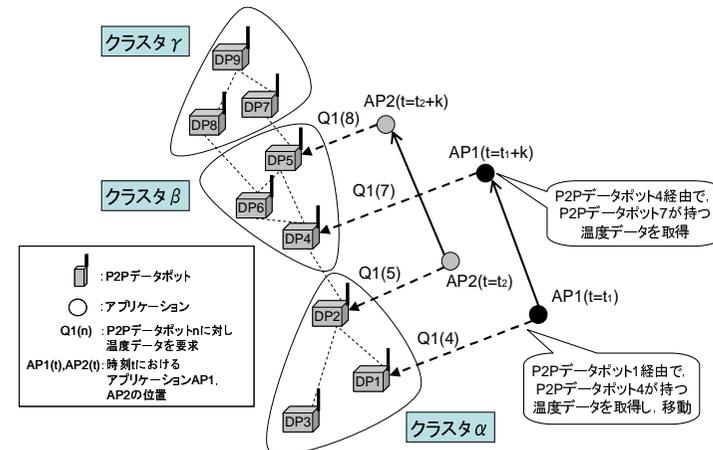


図 4 クラス単位での移動例

Fig. 4 The example of the movement by the class unit.

範囲内に存在するものは同一視して扱うものとする。

このような予測を実現するため、近隣の複数の P2P データポットをまとめてクラスタとして扱うことで、クラスタ単位での移動履歴を作成する。クラスタ単位での移動履歴を用いることで、各登山者が異なるデータソースポットに要求を発行した場合でも、それらが同一クラスタに所属していれば同一種類の履歴であるとみなし、次に要求が発行されるデータソースポットが所属するクラスタを予測することが可能になる。図 4 のクラスタ単位での移動例を用いて予測の例を示す。

アプリケーション AP1 は時刻 $t = t_1$ においてデータポット DP1 に対し DP4 の気温データを要求する問合せ Q1(4) を発行する。その後 AP1 は移動し、時刻 $t = t_1 + k$ において DP7 の気温データを要求する問合せ Q1(7) を発行する。このとき、クラスタ単位でのデータソースポットの遷移履歴として、気温データを要求する問合せ Q1 がクラスタ から に遷移したという情報が記録される。アプリケーション AP1 と同一種類のアプリケーション AP2 も同様に、移動しながら $t = t_2$ および $t = t_2 + k$ において同内容の問合せ Q1(5) と Q1(8) を発行する。この場合も同様に、Q1 がクラスタ から に遷移して発行されたという情報が記録される。

AP1 と AP2 は問合せ内容 Q1 が同一であることから、同一種類のアプリケーションであ

表 4 地中水分量の問合せ一覧

Table 4 A view of queries on amounts of water under the soil.

Query ID	Gateway ID	Date
Q1-1	DP1	2010/12/20/12:00:10
Q1-3	DP1	2010/12/20/12:30:13
Q1-5	DP1	2010/12/20/13:01:13

ると推測することができる。したがって、上記の履歴情報より、Q1 を発行するアプリケーションがクラスタ内のデータソースポットに問合せを発行した場合、時間 k 経過後にクラスタ内のデータソースポットに問合せを発行することを予測できる。

4. アルゴリズム

データソースポットが固定されている場合の予測アルゴリズムを 4.1 節で述べ、動的に変化する場合のアルゴリズムを 4.2 節で述べる。

4.1 データソースポットが固定されている場合

まず、予測を行うため、単体の P2P データポットの履歴を下記の手順で整理する。

- (1) 同一のクエリ内容で問合せをまとめる。
- (2) (1) で抽出した問合せ毎に、クエリ ID、ゲートウェイデータポットの ID とクエリを受けた時刻をまとめる。

ここでは、表 3 を基に、上記の手順を Gateway ID が DP1 かつ、問合せ内容が現在の地中水分量である問合せに対して行う場合を考える。今回 DP1 はクラスタに所属しているものとする。表 4 のような結果が生成される。

表 4 を利用し、繰り返し行われる問合せか否かの判断を行う。判断方法を以下に示す。

- (1) 各クエリの発行間隔を調べる。
- (2) 発行間隔が、閾値 T を超えていないかを比較する。閾値 T は使用環境により異なる。
- (3) 超えていれば、カウンタを 1 にし、超えていなければカウントアップする。
- (4) (1)~(3) をすべてのクエリに対し行う。
- (5) カウントがカウントの上限 C_{max} 以上であれば今後も繰り返し行われる問合せであると予測する。カウントの上限 C_{max} は使用環境により異なる。

具体的には、表 4 の Query ID Q1-1 と Q1-3 の Date の差からクエリの発行間隔を求め

る。その後設定した閾値を超えているかをチェックする。超えている場合にカウンタを 1 にし、超えていなければカウントアップする。すべての問合せに対しこの処理を行うと、最終的に以下のようなリストが生成される。

条件	出現回数
Cluster = & Query = 現在の地中水分量	11
Cluster = & Query = 現在の気温	1
Cluster = & Query = 現在の湿度	8

このリストに含まれる問合せのうち出現回数が C_{max} 以上のものは今後も繰り返し発行されると予測する。この例では、 $C_{max} = 10$ とし、クラスタに属する P2P データポットがゲートウェイとなり、現在の地中水分量を要求する問合せが発行されると予測することになる。

4.2 データソースポットとなる P2P データポットが動的に変化する場合

すべてのデータポットの履歴から、アプリケーション毎に履歴をまとめることで、データソースポットが動的に変化するような問合せにも対応する。まず下記の手順ですべてのデータポットの履歴を整理する。

- (1) 同一のクエリ内容でかつ同じアプリケーション ID を持つ履歴を、問合せ毎にまとめる。
- (2) (1) で抽出した問合せ毎に、クエリ ID、データソースポットの ID とクエリを受けた時刻をまとめる。

整理の結果、例えば、表 5 のような結果を得ることができる。表 5 は、アプリケーション ID が AP3 でかつ問合せ内容が、過去 30 分間の温度データであるものをまとめた結果である。このようにまとめることで、時系列順に、データソースポット遷移状況の把握を容易にする。以下に、このようにまとめた履歴を用いて予測を行う手段を示す。

- (1) まとめた履歴のデータソースポットが所属しているクラスタの遷移のリストを作成する。
- (2) (1) で作成したリストから、アプリケーション ID は異なるが、問合せ内容が同じ履歴を比較し、クラスタの遷移に同様の規則がみられないかチェックを行う。
- (3) 規則がみられた場合、今後も同様のクラスタ遷移をするようなアプリケーションが問合せを発行すると予測する。

表 5 AP3 が発行した過去 30 分間の温度データの問合せ一覧

Table 5 A view of queries on temperature data of the past 30 min issued by AP3.

Query ID	DataSourcePot ID	Date
Q1-21	DP3	2010/12/25/9:00:10
Q1-35	DP6	2010/12/25/11:05:15
Q1-45	DP9	2010/12/25/13:10:34

具体的には、表 5 の場合、DP3 がクラスタ に、DP6 がクラスタ に、DP7 がクラスタ に所属していた場合、手順 (1) より各アプリケーションと問合せ内容毎に以下のようリストが作成される。

条件	クラスタの遷移状況
Application ID = AP3-4 & Query = 過去 30 分間の温度データ	
Application ID = AP1-6 & Query = 現在の湿度	
CApplication ID = AP2-6 & Query = 過去 30 分間の温度データ	

このリストを元に手順 (2) を実行すると、AP3-4 と AP2-6 のものが同じ問合せ内容で、同様のクラスタ移動パターンがあることが判明する。このような問合せが複数回行われている場合、クラスタ , , に所属する P2P データポットに対し、過去 30 分間の温度データが繰り返し発行されると予測する。

5. 関連研究

文献 4) では、無線マルチホップネットワーク環境下において、データロスや遅延の緩和を行うため、コンテンツやサービスのより効率的なレプリケーション手法を提案している。効率的なレプリケーションを実現するために、コンテンツやサービスのオブジェクトの平均アクセス時間の最適化や、各ノードにおけるレプリカデータの置換アルゴリズムの最適化を行っている。しかし、対象としているコンテンツやサービスは、更新頻度が低いことが想定されており、センシングデータのような更新頻度が高いデータのレプリケーションには適切でないと考えられる。また、対象がコンテンツやサービスであり、本研究で扱うリレーショナルデータのようなデータモデルは想定されていないため、この手法をそのまま利用することは難しい。

文献 5) は、モバイルアドホックネットワークデータベースのレプリケーションの研究に

関するサーベイである。アドホックネットワーク環境におけるデータの複製、配置を行う既存の研究を、省電力性、データの可用性、リアルタイム性などさまざまな観点からまとめている。アドホックネットワーク環境下のレプリケーションの研究には様々なものが存在するが、目的の多くはノードの移動や電力枯渇によりデータが利用できなくなる場合に備え、データを複製、配置し、データの可用性を向上させることである。一方、アドホックネットワーク環境下で発生する応答性の低下については、十分考慮されていない。また、これらの研究で対象としているデータは、ファイルなど更新頻度が極めて低いものを想定しているため、本研究が対象とするリレーショナルデータモデルのように頻繁に更新される可能性があるデータに対する十分な考慮がなされていない。

6. おわりに

本稿では、センシングデータのプロアクティブデータ転送・検索機構の概要について述べ、問合せ履歴を用いたプロアクティブデータの予測手法を提案した。今後は、プロアクティブデータの配置手法と本提案手法の評価方法の検討を行う予定である。

参考文献

- 1) 藤崎友樹, 鈴木和久, 横田裕介, 大久保英嗣: “無線アドホック通信を利用したセンサネットワーク向け協調ストレージシステム”, 情報処理学会研究報告, 2008-MBL-44/2008-UBI-17, pp. 149-156, 2008 年 3 月.
- 2) 井邊研吾, 陶山優一, 西原雄太, 藤原秋司, 横田裕介, 大久保英嗣: “センサデータの分散管理システムにおけるプロアクティブ転送・検索機構”, 情報処理学会第 72 回全国大会, 第三分冊, 3Z-3, 2010 年 3 月.
- 3) Yokota, Y., Ibe, K., and Okubo, E.: “A Sensor Data Management System with a Proactive Distribution Mechanism”, 2010 International Workshop on Ubiquitous Service Platforms (IWUSP 2010), IEEE Computer Society, Kyoto, Oct. 2010.
- 4) Shudong Jin, Limin Wang: “Content and Service Replication Strategies in Multi-hop Wireless Mesh Networks”, International Workshop on Modeling Analysis and Simulation of Wireless and Mobile Systems, pp.79-86, 2005.
- 5) Padmanabhan, P., Gruenwald, L., Vallur, A., and Atiquzzaman, M: “A survey of data replication techniques for mobile ad hoc network databases”, The VLDB Journal Volume 17, Issue 5, pp. 1143-1164, 2008.