

センサネットワークにおける検索クエリの効率化 Proposal on Efficient Query Processing in Sensor Network

齋藤 祐里† 藤田 悟‡
Yuhri Saito Satoru Fujita

1. はじめに

現在、気温センサや人流センサなどの多様なセンサが各地に設置されつつある。今後、これらの大量のセンサから検出されるデータを、ユーザが必要な時に欲しい情報として検索できるサービスの需要が高まると予想される。そこで本稿では、主にサーバ負荷軽減の観点から、ユーザから発信された検索クエリを、より迅速に処理できるシステムモデルの構築について検討する。まず、検索クエリをいくつかの型に分類し、それぞれの型の特徴に合わせたシステムモデルを構築する。そしてそれぞれのシステム負荷を分析し、クラウドを用いることによって検索クエリを効率的に処理できるシステムを提案する。

関連研究には、TomuDB[1]がある。TomuDBとは、異種センサネットワークからの多様なデータを効率よく処理し、一般ユーザがセンサデータを容易に統合利用することを目的としたオーバーレイネットワークの概念に基づく分散センサデータベースである。本稿では、TomuDBが複数のセンササーバ群をオーバーレイネットワークとしている部分を参考にしている。

以下、第2節では、検索クエリ型によるセンサデータの分類、第3節ではそれらの検索クエリを処理するにあたってのクラウドを利用した検索処理の効率化、第4節では考察、そして第5節ではまとめについて述べる。

2. 検索クエリ型によるセンサデータの分類

センサネットワークに対する検索は、気象センサによる避難勧告システム[2]のようなリアルタイムでセンサ情報が要求されるもの、大規模火災における温度や湿度の変化[3]のように過去のセンサ情報の抽出と分析が要求されるもの、そして人流計測によるなど避難誘導システム[4]のように避難時の各出口の人流密度を計測することで効率的な避難誘導を実現できるものなど、検索対象に個々の特徴がある。本稿では、これらをストリーミング検索型、リアルタイム検索型、履歴検索型の3つに分類する。

(1)ストリーミング検索型

連続的に流れるデータからクエリの検索条件に合ったデータを見つけた時に、通知するモデルである。気温センサを例にとると、「東京で気温が30℃に達した場合通知してほしい」という条件を設定した時に、東京に対応した気温センサが低い温度から30℃を計測した時点でユーザ側に通知を送ることになる。

(2)リアルタイム検索型

このモデルは、一定期間のセンサ情報をまとめて検索するものである。気温センサの例では、ある地点での一定期間の気温を要求するクエリが送られ、その地点に対応したセンササーバに保管されている過去のセンサデータから指定された期間のデータ群をグラフ化し、ユーザに送信する。

(3)履歴検索型

履歴検索型とはある地点においてクエリ送信時からのセンサデータが統計的条件を満たした時に、ユーザに通知を送るモデルである。基本的にはストリーミング検索型に近いが、一定時間ごとに送られるセンサデータを統計処理する事が異なる。人流センサを例にとると、「新宿における通行人数が100人に達した場合通知してほしい」というクエリを送るとすると、新宿のセンサから一定時間ごとに送られて来る人数を加算し、100人に達した時に通知を送る。ここでは加算を例にとったが、「クエリ送信時点から1時間ごとの平均が〇人以上」や「クエリ送信時点から1時間ごとの合計人数が〇人」といったより複雑な演算を要求されるクエリも想定されるため、これらのクエリの処理量を想定したシステムモデルの構築が要求される。

3. クラウドを用いた検索の効率化

本節では、前記の3つの検索クエリ型について、それぞれを効率的に処理するシステムモデルを提案する。必要に応じて、クラウドを用いたスケールアウトの構成を活用することを検討した。

3.1. ストリーミング検索型

ストリーミング検索型のクエリを処理するにあたってボトルネックになる部分はクエリの管理である。仮に全てのクエリをセンササーバに登録して処理するモデルを採用すると、内容が重複するクエリが多数存在する場合に冗長な処理が発生してセンササーバのパフォーマンスが低下する。これに対して提案するモデルは、クエリと送信元のクライアントのアドレスを登録するデータベースをクラウド内に設け、同一のクエリを1つにまとめて、センササーバに登録する。メインサーバはユーザからのクエリを適切なセンササーバに登録する機能を担う(図1)。

この方式のメリットは、ユーザ数が多くなって、重複するクエリが増加した場合でも、内容が重複する機会が増え、それによるセンササーバへのトラフィックや負荷が増加することを抑制できる点である。

† 法政大学大学院情報科学研究科 Yuhri Saito, Hosei Univ. Computer and Information Sciences.

‡ 法政大学情報科学部 Satoru Fujita, Hosei Univ. Computer and Information Sciences.

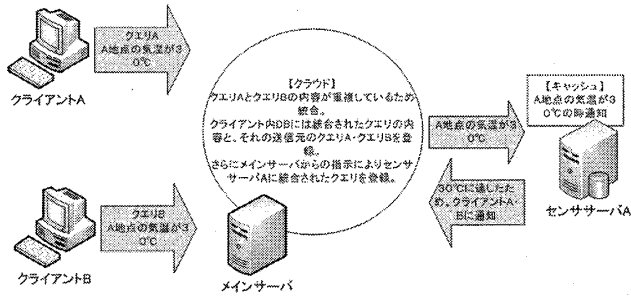


図1: ストリーミング検索型クエリ処理モデル

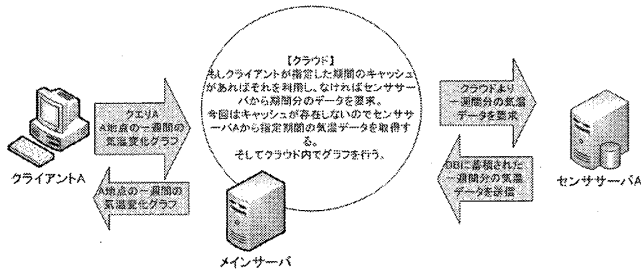


図2: リアルタイム検索型クエリ処理モデル

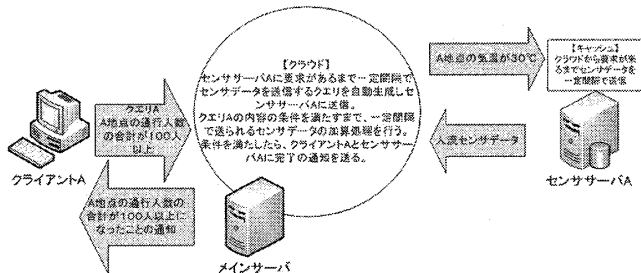


図3: 履歴検索型クエリ処理モデル

3.2. リアルタイム検索型

リアルタイム検索型クエリを処理するにあたって、ボトルネックとなるのは過去の気温情報の検索処理である。ユーザから情報を要求されるたびに過去の気温情報をセンササーバから検索するモデルであると、センササーバとメインサーバの間のトラフィックが増加し、さらにデータ処理の増加と共にサーバの負担も増大する。これに対して提案するモデルでは、クラウドがセンサデータベースからの一定期間の気温情報を受け取り、グラフ化を一括して行う。クエリはストリーミング検索型と同様にクラウド内のデータベースに登録される(図2)。

一度受け取った気温情報をクラウド内に一定期間キャッシングすることでセンササーバから同じ気温情報を検索することがなくなる。そのためセンササーバからのトラフィックを抑える効果が期待できる。

3.3. 履歴検索型

履歴検索型処理のプロセスにおいてストリーミング検索型に似通った特徴を持つが、最大の違いは一定時間内のセンサ情報の統計処理が加わる部分であり、ここが履歴検索型のボトルネックである。センササーバ上で統計処理を行うモデルでは、データベース登録処理と並行して統計処理も行うためセンササーバのパフォーマンスが落ちる。これを解決するため、提案するシステムモデルは、統計処理

機能とクエリ管理機能をクラウド内に設ける。メインサーバはストリーミング検索処理モデルと同様に、ユーザからのクエリを適切なセンササーバに登録する機能を担当する。センサデータサーバにはクエリ登録キャッシュに加え、一定時間ごとにセンサ情報をクラウドに送信する機能を持たせる(図3)。

4. 考察

クライアントの人数が10万人で、1人につきクエリを10個送信し、これらを10台のセンササーバで処理することを仮定し、システムの性能を検証する。

(1) ストリーミング検索型

クエリの総数は100万個となり、1台につき10万個のクエリを処理することになるため負荷は非常に大きくなる。しかし、ユーザ数が増加するほど、重複するクエリが増えるため、独立したクエリは高々1000~10000個程度と考えられる。そうするとセンササーバ1台につき100~1000個の処理数になるため、センササーバでも処理できる処理量となる。

(2) リアルタイム検索型

10万人のうち、同時に検索要求する確率が1%程度と考えると1000クエリの同時処理が必要である。ただし、1クエリごとのデータ量はストリーミング検索型よりも大きいいため、1000クエリ×データ量の検索処理量がかかる。ここにおいてクラウドでキャッシュを用いることで100分の1程度までに処理量を削減することが期待できる。

(3) 履歴検索型

基本的にはストリーミング検索型と同様だが、統計処理を行うためその分負荷は高くなる。ここでクラウドのスケールアウト機能を用いることでストリーミング検索型と同等のパフォーマンスを維持できると考えられる。

5. 結論

以上によりサーバの負荷を考慮するという点においてはクラウドを利用して検索クエリ処理を負荷分散することは有効と思われる。今回は机上シミュレーションで評価したが、今後実機上の評価を行い、有効性の確認を進めていく予定である。

参考文献

- [1] Yoh Shiraishi, Niwat Thepvilajanapong, Yosuke Tamura, Tatsuro Endo, Koichi Yamada, Nayuta Ishii, Hiroki Ishizuka, Keisuke Kanai, Yoshito Tobe, "TomuDB: Multi-Resolution Queries in Heterogeneous Sensor Networks through Overlay Network", Poster session, The Fifth ACM Conference on Embedded Networked Sensor Systems (SenSys 2007).
- [2] 松浦知史, 洞井晋一, 落合秀也, 石塚宏紀, 江崎浩, 砂原秀, "Live E!: センサストリーム制御機構および地理位置に基づくオーバーレイネットワークを利用したセンサ情報共有基盤の構築", 情報処理学会 マルチメディア, 分散, 協調とモバイル(DICOMO2007) シンポジウム論文集, pages 1161-1167. Jul. 2007.
- [3] 森戸貴, 猿渡俊介, 南正輝, 森川博之, "Smart Dust から10年: 無線センサネットワークの展開", 森川研究室 技術研究報告書 No.2008002, 2008-5-1.
- [4] 山下倫央, 副田俊介, 野田五十樹, "人流計測による避難誘導効果の実証的検証", ヒューマンインタフェース学会研究報告集 11(5), 131-138, 2009-11-12.