

# 「構成物」構成に基づく高速センサデータ集積処理エンジンとそのサービス化に向けた考察

蔡 程緯<sup>†1</sup> 菊地 伸治<sup>†2</sup>

**概要:** 複数センサから収集したストリームデータ(RawData)を,そのセンサ群を組み込んでいる「構成物」構成に基づき加工,「構成物」の特性データに変換する専用アグリゲータ処理系を開発して来た.本稿では,当該専用アグリゲータ処理系を紹介すると共にサービス化する上でのデータ管理上のポイントについて概略説明する.

**キーワード:** センサデータ, アグリゲーション, アルゴリズム

## Implementing the fast aggregation algorithm of sensor data based on configuration of implementing structures and consideration for making it as a service

JOSEPH C. TSAI<sup>†1</sup> SHINJI KIKUCHI<sup>†2</sup>

**Abstract:** We developed a specialized software processor that can translate a set of gathered stream data (raw data) generated by multiple sensors into the characteristic behavioral data of the structure embedding these sensors after aggregating them according to the configuration of the structures. In this paper, we present the important points to make it as a utility as a service.

**Keywords:** Sensor Data, Aggregation, Algorithm

### 1. はじめに

著者らは,複数センサから収集したストリームデータ(RawData)を,そのセンサ群を組み込んでいる「構成物」構成に基づき加工,「構成物」の特性データに変換する専用アグリゲータ処理系を開発して来た[1]. ここではRawDataを特定するためデータモデル上を検索する処理と実際のRawDataを検索・集計する処理の2フェーズに分離することで,処理の高速化を図っている.本稿では,当該専用アグリゲータ処理系を紹介すると共にサービス化する上でのデータ管理上のポイントについて概略説明する.

### 2. 高速センサデータ集計処理エンジンの概要

図.1は,ストリームデータ(RawData)と「構成物」の特性データの視点の違いを記す.自動車では運転士が法定速度以下で走行しているか否かを確認するため,速度計が設置されている.これは運転士を含めた人間の操作に適用する様に,毎時の km 単位で記されるのが一般的である.しかし実際には車軸等に取り付けた回転計を元に計算式を用いて計算している.通常は個々センサからの Raw Data は,利用するアプリケーションで集計処理を施す必要がある.しかし,センサ群を組み込んでいる「構成物」構成に基づき,Raw Dataを加工,「構成物」の特性データに変換する処理を,それを利用するアプリケーションとは独立に,共通的な処理系とし

て定義出来れば,効率化・付加価値の点で大きな効果を得ることが出来る.しかし,この為には図.2 に記すセンサ,「構成物」を汎化した上で,多重拘束の関係を元に集計処理を実施する必要がある[2]. ナイーブな実装アルゴリズムを採用して実装した場合,図.3 の様に 3 重のループ構造が出現し,計算量を増大させる.これは Raw Data が巨大化する程,性能劣化は極端に悪化する.

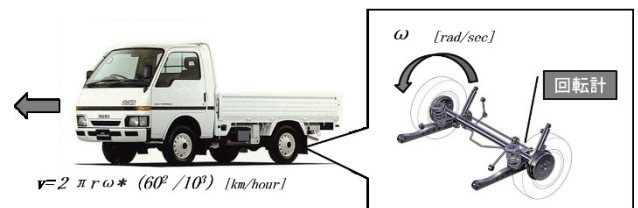


図 1 センサからの Raw Data と「構成物」の特性データの視点の違い

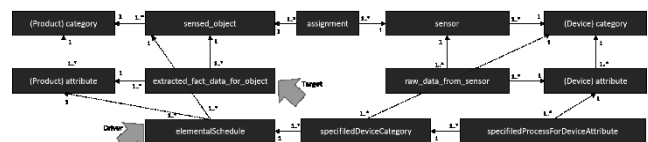


図 2 データモデルから見た関係の複雑さ

<sup>†1</sup> HTC Corporation

<sup>†2</sup> 公立大学法人 会津大学 先端情報科学研究センター

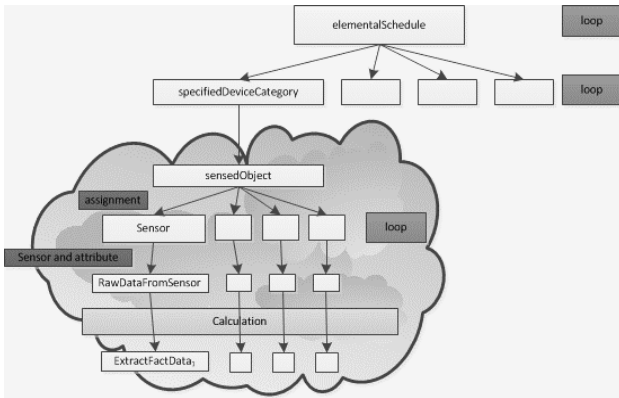


図 3 ナイブに実装した場合のアルゴリズム概要

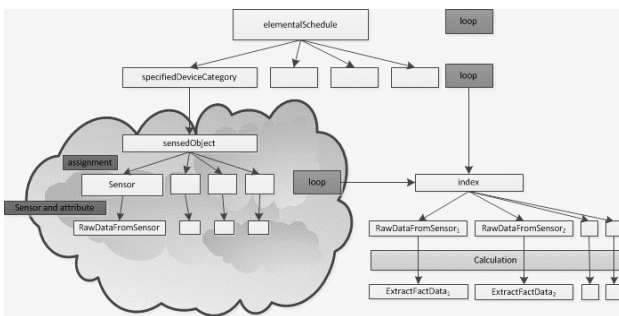


図 4 高速化を考慮したアルゴリズム概要

専用アグリゲータ処理系の設計・実装に当たっては、アルゴリズム的な工夫を施している。具体的には図 4 で記される様な構成管理に関するデータモデル、並びにオントロジ管理に関するデータモデルからの多重拘束の関係をトラバースに検索して中間索引表を生成する処理フェーズ、並びにこの中間索引表を元に、大量の Raw Data を直接検索する処理フェーズへの分離である。図.4 の提案方式を実装し、ナイブな方式と比較した測定結果を図.5 に記す。ここでは単一センサの場合に限定している。縦軸は処理応答時間である。パラメータとしては Raw Data の母数行数、並びに検索対象行数である。図左側は、ナイブに実装した場合の応答性能であるのに対して、図右側は、高速化した方式である。その効果は明白である。これは複数センサとした場合でも同様の効果を得ている。

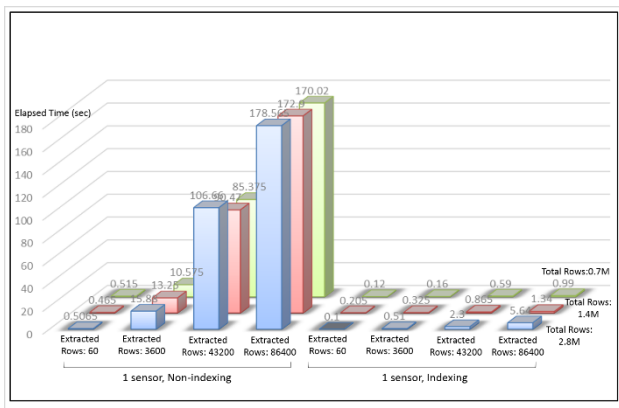


図 5 計測結果(単一センサの場合)

### 3. サービス化に向けた展開

図.6 は、データ管理の観点から見た当該機能の配置構成案である。ここで Storage(ForExtracedData)に専用アグリゲータ処理系が配置され、個々センサからの Raw Data は Storage(ForRawData)に保持される。サービス化をする上では、これらの参照関係を維持するために、等価的かつレイテンシーの低い通信方式でデプロイメントする必要がある。更に Storage 上の DBMS のタイプについても複雑な Query を処理し得る RDBMS と透過性の高いエンジンが求められる。この為には DBMS にネイティブなレプリカを持つ実装を採用する必要がある。これらについては検討途上であり、今後の報告事項である。

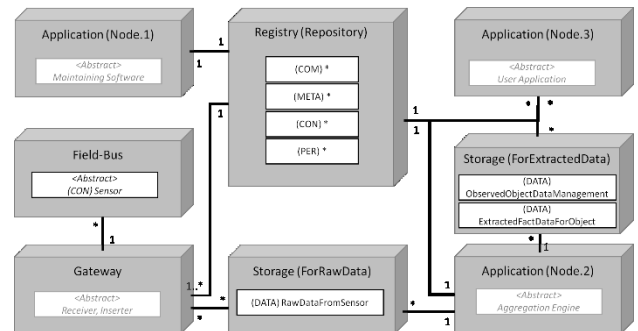


図 6 データ管理の観点から見た機能配置構成

### 参考文献

- [1] Joseph C. Tsai, Shinji Kikuchi, "Implementing the fast aggregation algorithm of sensor data based on configuration of implementing structures" Note of BASE 2016.
- [2] S. Kikuchi, A. Nakamura, D. Yoshino, "Evaluation on Information Model about Sensors featured by Relationships to Measured Structural Objects", Advances in Internet of Things, 2016, 6, pp.31-53, 2016.