

# データ収集型無線センサネットワークにおけるデータ圧縮の分散化の検討

## A Consideration of Decentralized Data Compression for Data Collection WSN

山崎 慧<sup>1</sup>      勝間 亮<sup>1</sup>  
Kei Yamazaki   Ryo Katsuma

### 概要

センサネットワークにおけるデータ圧縮では、すべてのセンサに対してひとつのルールにのっとって圧縮すると、ある部分では非効率な圧縮になっている場合がある。そこで、複数の領域に分割してそれぞれの領域で適切なルールで圧縮することで、よりデータ圧縮の効率化を図る。

本稿では、可逆圧縮で、取得データに対して動的に効率的な圧縮ルールを決めることが可能なハフマン符号化に着目し、シミュレーション実験により圧縮の効率化が可能であることを示す。

### 1. はじめに

現在、センサネットワークの分野においては様々な研究が行われ、その技術は環境情報を取得することでセキュリティ、農業、気象観測などへの適用が期待されている。センサネットワークの典型的な例としては、まずセンサを空間上に散在させ、各センサで温度や湿度といった環境データを計測する。その後、センサ全体をいくつかのクラスタに分け、各クラスタで決められたクラスタヘッドがクラスタ内のデータを集約し、代表してそれをシンクに送信してデータを処理する。

このセンサネットワークにおける課題の一つとして挙げられるのが、各センサが送信するデータをどのように圧縮するかということである。センサネットワークで用いられるセンサは一般的にバッテリー駆動であり、頻繁にバッテリーを交換することが難しい。また、センサでの消費電力は送信するデータ量に比例するため [1]、省電力化という点でデータ圧縮は必要不可欠な技術といえる。さらに、データ圧縮は、記憶領域の節約や低遅延といった点においても効果的である。

センサネットワークにおけるデータ圧縮では、センサデータが刻々と変化するため、事前に決定した圧縮方法を静的に使い続けるよりは、センサデータに合わせた動的な圧縮方法を用いる方が効率的である。柳沢らの研究 [2] では、符号化するデータ系列を二分割した後、一方のデータからシンボルの出現頻度を算出し、他方のデータ

は算出された出現頻度を用いて低密度パリティ検査符号で符号化する手法を提案している。このような動的圧縮アルゴリズムの代表例としてハフマン符号化があり、適用する時点でのセンサデータに合わせた効率的な圧縮が可能である。しかし、この方法では一部で非効率なデータ圧縮になることがある。そこで、センサをクラスタリング後、各クラスタに最適な圧縮ルールを割り当てることで、圧縮率を高める方法についてシミュレーション実験をベースに検討していく。

### 2. 事前実験

ここでは、センサ全体を一つのルールで圧縮した場合より、各クラスタで別々のルールを用いて圧縮した場合の方がデータサイズを抑えられるのか確かめるため、それぞれの圧縮率を比べる実験を行う。

#### 2.1 環境設定

本シミュレーション実験では、センサを格子状に縦横 30 個の計 900 個並べ、それらすべてを使って温度を 5 回計測し、そのデータをまとめて符号化する場合をシミュレーションした。温度データは、図 1 のようにセンサ空間を三等分し、各グループの平均値を 20, 21, 22 度とした正規分布を用いて得られた数値を整数化したものを用いた。



図 1: センサ空間を三等分した図

#### 2.2 実験方法

センサデータの圧縮アルゴリズムとして、ハフマン符号化を使って実験を行った。ハフマン符号化は、出現頻度の高い記号には短い符号を、低い記号には長い符号を割り当てる符号化であり、センサデータの時間変化に応

<sup>1</sup> 大阪府立大学, Osaka Prefecture University, Sakai, Osaka 599-8531, Japan

じた効率的な圧縮が可能である。そこで、近くにあるセンサノード同士では同じデータになる可能性が高い温度観測を仮定した本実験において、各クラスタでハフマン符号化をすれば高い圧縮率を得られるのではないかと考えた。

具体的な実験方法としては、センサデータ全体でハフマン符号化を行った場合と、センサデータを100個のセンサが属するクラスタAからIの9個に分け(図2)、各クラスタ内でハフマン符号化を行った場合の圧縮率を比べ、どちらがより高い圧縮率を得られるか調べた。

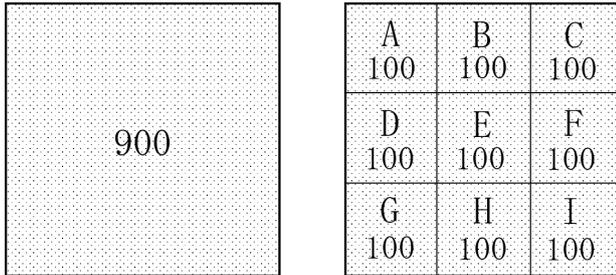


図2: クラスタリングのイメージ図

### 2.3 実験結果

圧縮率は、もとのデータサイズを  $D$ 、圧縮後のデータサイズ(圧縮ルールの定義部分も含む)を  $E$  とすると式(1)で表される。

$$\frac{D - E}{D} \times 100 \quad (1)$$

つまり、圧縮率が高い方がよりデータサイズを小さくすることができたということになる。

表1にセンサ全体のデータをハフマン符号化した場合とクラスタごとにハフマン符号化した場合の圧縮率をまとめた。

表1: 圧縮する領域を変化させた場合のそれぞれの圧縮率

	圧縮率 (%)
全体を一律に符号化した場合	49.896
クラスタごとに符号化した場合	55.071

### 2.4 考察

実験結果より、全体で一つのルールを用いて符号化するよりクラスタごとに別々の適切なルールを使って符号化の方が圧縮率が高くなる場合があることが確かめられた。表2は各クラスタAからIにおいて、全体での圧縮ルールで符号化した場合と、クラスタ独自の圧縮ルールで符号化した場合の圧縮率をまとめたものである。

全体での圧縮ルールで符号化した場合の圧縮率は、設定した温度データの差によってばらつきが出ているのに対し、クラスタ独自の圧縮ルールを適用した場合の圧縮

率は安定していることが分かった。これは、各クラスタにおける温度データの出現率をもとに、そのクラスタの最適なハフマン符号化を行えたためであり、気温の差が大きいセンサ空間においても有効だと考えられる。

表2: 各クラスタにおけるそれぞれの圧縮ルールでの圧縮率

	全体での圧縮ルール (%)	クラスタでの圧縮ルール (%)
クラスタA	42.16	54.66
クラスタB	42.80	55.30
クラスタC	42.52	53.20
クラスタD	55.42	55.42
クラスタE	55.38	55.52
クラスタF	55.56	56.48
クラスタG	51.18	54.98
クラスタH	52.32	55.96
クラスタI	51.72	54.12

### 3. まとめ

本稿では、センサネットワークにおいて、各クラスタで別々のルールを用いて圧縮する手法について、シミュレーション実験により効率化が可能であることを確認した。今後の実験では、ハフマン符号化以外の様々な符号化を各クラスタで実行できるようにしておき、センサデータによって最適な符号化を実行した場合の圧縮率が事前実験と比べてどのように変化するかを調査する予定である。

### 参考文献

- [1] 森下宏樹, 重井徳貴, 宮島廣美: センサネットワークのための近傍ノード数に基づく省電力クラスタリング通信. 情報処理学会研究報告 Vol. 2009-MBL-50, No. 9(2009).
- [2] 柳沢豊, 村松純, 岸野泰恵, 須山敬之, 納谷太: 無線センサネットワークを用いた環境モニタリングのためのデータ圧縮手法. マルチメディア, 分散, 協調とモバイル (DICOMO2014) シンポジウム (2014).