An Effective Data Compression Technique Exploiting Spatial-Temporal Correlations in WSN

Venkatesan S[1*], Xavier arockiaraj S[2], Karuppasamy P[3]

[1*] Department of Computer Science and Engineering,
Adhiyamaan College of Engineering, Hosur-635109,
TamilNadu, India.

[2] Department of Electronics and Communication,
Adhiyamaan College of Engineering, Hosur-635109,
TamilNadu, India.

[3] Department of Electronics and Communication,
Adhiyamaan College of Engineering, Hosur-635109,
TamilNadu, India.

***Abstract*— This paper presents a novel data compression and dimensionality reduction scheme for data fusion and aggregation in wireless sensor networks (WSNs), aimed at preventing data congestion and reducing energy consumption. The approach utilizes an unsupervised neural network to analyze temporal and spatial data, enabling efficient compression without losing significant data accuracy. A new distributed algorithm for data compression is introduced, based on a hierarchical network structure. The algorithm employs two distinct compression techniques: Delta Encoding and Run-Length Encoding (RLE). Delta Encoding is used to compress consecutive data by transmitting differences between values, while RLE groups similar data at the cluster head for further compression. The proposed scheme effectively reduces data size and minimizes communication overhead, optimizing energy usage in the network. Through these techniques, the method ensures high compression efficiency and energy conservation without compromising the accuracy of the transmitted data.**

I. INTRODUCTION

Wireless sensor network is an emerging technology. In which it has lot of applications it can be used to measure the temperature, humidity etc. In wsn the Data compression is used to reduce the number of data send without any error or data lost. the basic idea of the past paper is compressive sensing it consists of one sink node and N node for collecting the data in which hybrid network is used which is efficient but its cluster size is lager means it required to collect data within the cluster head between the cluster must be very high. Instead of using hybrid network the hierarchical network is used in which the data can be collected by using energy efficient data gathering method.

In this we proposed that the data can be compressed by spatial-temporal correlation. Compression will be done at cluster head and it will then transmit to the base station. Cluster head can be elected through the high efficient energy in the node and then it will group the node in certain distance. the compression can be done by two ways first is delta encoding and then run length encoding. The delta encoding is used to compress the data by grouping difference between consecutive data and similar data are grouped together by using the run length encoding. By using this spatial-temporal correlation we can attain the energy efficient.

II. METHODOLOGY

The methodology of this paper revolves around utilizing spatial-temporal correlation for efficient data compression in a hierarchical wireless sensor network (WSN). The core idea is to reduce the amount of data transmitted from the sensor nodes to the base station, thus improving energy efficiency and reducing the communication overhead. Spatial and temporal correlations are exploited in two primary ways: **Run Length Encoding (RLE)** for spatial correlation and **Delta Encoding** for temporal correlation. These two techniques work synergistically to compress the data at the cluster head, which is the energy-rich node in the network, achieving energy balancing and minimizing energy consumption for sensor nodes. This methodology is highly effective in hierarchical sensor networks where the data collected by individual sensor nodes often exhibit both spatial and temporal dependencies.

1. Spatial Correlation Using Run Length Encoding (RLE)

Spatial correlation occurs when sensor nodes that are located close to each other in a wireless sensor network measure similar physical phenomena. For example, in a temperature monitoring network, neighboring sensors often report similar temperature values because they are located in the same environmental conditions. The data from these neighboring nodes tend to have long runs of the same value or closely related values, which makes it a prime candidate for **Run Length Encoding (RLE)**.

Run Length Encoding (RLE) is a simple and efficient data compression technique that replaces sequences of identical data points (runs) with a single data value and the length of the run. For instance, if several consecutive temperature readings from neighboring sensor nodes are the same (e.g., 20°C, 20°C, 20°C), RLE compresses this sequence into a single value (20°C) and stores the count of the consecutive identical values (3). This drastically reduces the data size and minimizes the amount of information that needs to be transmitted.

In the context of the hierarchical network, **compression at the cluster head** is crucial. The cluster head is a node that aggregates data from the sensor nodes within its cluster. As it is typically more powerful and energy-rich than the sensor nodes, it is well-suited to perform the computationally intensive task of applying RLE. Since the data transmitted from the sensor nodes to the cluster head often contains spatial redundancy (due to the proximity of the nodes), RLE is applied to exploit this redundancy and compress the data.

By using RLE, the cluster head can transmit much smaller packets of data to the sink or base station, significantly reducing the bandwidth requirements. This compression process ensures that fewer resources are used for transmission, extending the lifetime of the network.

2. Temporal Correlation Using Delta Encoding

Temporal correlation refers to the relationship between successive data points over time. In many sensor networks, sensor readings over time tend to be similar, with only small changes between consecutive measurements. For example, temperature readings from a sensor node might fluctuate by only a few degrees over time, leading to small differences between successive values. This kind of correlation is best handled using **Delta Encoding**.

Delta Encoding compresses data by recording the difference between consecutive data points rather than the actual values themselves. For instance, if a temperature sensor records the following sequence of values: 20°C, 20.5°C, 21°C, 21.5°C, delta encoding stores the differences between consecutive values: 0.5°C, 0.5°C, 0.5°C. This reduces the amount of data stored and transmitted, as the deltas (the differences) are often smaller in magnitude than the original values.

By applying **Delta Encoding** to the data collected over time, temporal redundancy is effectively exploited. This method is particularly useful in sensor networks where the data does not change drastically between measurements and only small variations are observed.

When this technique is applied at the **cluster head**, which collects and processes data from multiple sensor nodes over time, the temporal correlation between the successive readings from the nodes can be compressed into smaller, more efficient data packets. This process reduces the communication overhead and decreases the amount of energy required to transmit data to the base station.

3. Data Compression at the Cluster Head

The hierarchical structure of the network is key to the data compression process. In this network, sensor nodes are grouped into clusters, and each cluster has a **cluster head**. The cluster head is responsible for collecting data from all the sensor nodes within its cluster, performing data aggregation and compression, and then transmitting the compressed data to the base station or sink.

The process of compressing the data at the cluster head leverages both **spatial correlation** (via RLE) and **temporal correlation** (via Delta Encoding). By performing the compression at the cluster head, we achieve several benefits:

- **Reduced Data Transmission:** Both RLE and Delta Encoding reduce the size of the data, meaning that less data needs to be transmitted between the cluster head and the base station. This reduces the amount of bandwidth required, leading to more efficient use of network resources.
- **Energy Efficiency:** Since the cluster head is more powerful and energy-rich compared to the sensor nodes, it is capable of handling the computationally expensive task of compression. The sensor nodes only need to transmit raw data to the cluster head, which reduces their energy consumption since they do not need to perform any complex processing.
- **Energy Balancing:** Energy balancing is achieved because the compression process is shifted to the cluster head. This reduces the communication load on the sensor nodes, preventing them from depleting their energy reserves quickly. By transferring the computational burden to the cluster head, the network can achieve a more balanced energy distribution, enhancing the overall network lifetime.
- **Minimized Communication Overhead:** By reducing the amount of data that needs to be transmitted, the overall communication overhead is minimized. This is particularly important in wireless sensor networks where communication energy costs are high, and reducing the frequency and size of transmissions can significantly prolong the life of the network.

4. Energy Balancing in Hierarchical Networks

Energy balancing is one of the most critical concerns in wireless sensor networks, especially when nodes have limited battery life. In hierarchical networks, where sensor nodes are grouped into clusters with a central cluster head, it is vital to ensure that energy consumption is evenly distributed to avoid early depletion of energy in certain nodes, leading to premature network failure.

In this methodology, **compression at the cluster head** not only reduces the bandwidth but also ensures that the sensor nodes do not perform complex operations that could drain their energy. Since only the cluster head is responsible for compressing the data, the sensor nodes send raw data with minimal energy expenditure. This results in **energy-efficient data communication** and extends the overall lifetime of the network.

Furthermore, the cluster head can be chosen based on its available energy resources, ensuring that the most capable node handles the compression and transmission tasks. This allows the network to dynamically balance energy consumption, promoting the longevity of the network as a whole.

III. PROPOSED WORK

The proposed work focuses on improving the energy efficiency and data compression in a wireless sensor network (WSN) using spatial-temporal correlation. The methodology includes distance calculation for cluster formation, the use of **Run-Length Encoding (RLE)** and **Delta Encoding** for data compression, and the election of **Cluster Heads (CH)** in the network. The aim is to create a system that reduces communication overhead, balances energy consumption, and extends the overall network lifetime.

1. Distance Calculation for Cluster Formation

In the proposed system, distance calculation plays a crucial role in forming efficient clusters. Each sensor node in the network calculates the distance between itself and the base station using the following equation:

s.xy(:,3) = sort(power((50-s.xy(:,1)),2) +power ((50-s.xy (:,2)),2)) calculate the distance between nodes is given by sadist (i j) = calculate distance (s xy (i ,:), s.xy (j , :)) the clustering will form at radius 25cm is used and we explained the compression technique in this paper. The first technique is run length encoding and next is delta encoding.

Once the distances are calculated, clustering is done within a fixed radius of 25 cm. The nodes that are within this radius from each other form a cluster, and the **Cluster Head (CH)** is selected from among the nodes in each cluster. This localized clustering ensures that the communication overhead is minimized, as nodes will primarily communicate within their cluster before forwarding data to the base station.

2. Run-Length Encoding (RLE)

Run-Length Encoding (RLE) is a simple but effective lossless data compression technique. It works by replacing sequences of repeated data (called runs) with a single data element and a count representing the number of repetitions. In the context of sensor networks, this is particularly useful when sensor data from nearby nodes exhibit redundancy or repetition.

For example, consider a string:

aaaaaaaaaabbbaxxxxyyyzyx.

Without RLE, this string has a length of 24 characters. Using RLE, the string becomes:

a10b3a1x4y3z1y1x1.

This reduces the length of the string to 17, which is about 70% of the original size. The major advantage of RLE lies in its simplicity and efficiency when compressing sequences of identical data values.

Advantages of RLE in Sensor Networks

- **Simple to implement**: RLE is easy to implement and does not require significant computational resources.
- **Effective for repetitive data**: It is highly efficient when there are long sequences of repeating data. In sensor networks, many readings from neighboring nodes can be similar or identical, making RLE highly applicable.
- **Compression of sparse data**: In situations where data contains large areas of repeated values, such as when multiple nodes measure the same environmental condition, RLE achieves significant data reduction.

However, RLE is not optimal when the data contains little or no repetition. In such cases, the compressed data might even be larger than the original, but it is still highly effective in networks with spatially correlated sensor readings.

3. Delta Encoding

Delta Encoding is another data compression technique that exploits **temporal correlation** in sensor data. Instead of storing the absolute values of sensor readings, delta encoding stores the difference (or delta) between successive values. This method is particularly useful when changes in data are small over time.

For example, consider the sequence of sensor readings:

2, 4, 6, 9, 7

Using delta encoding, we would store:

2, 2, 2, 3, -2

The first value (2) is retained as the initial reading, and the subsequent values represent the differences between consecutive readings. By recording the deltas instead of the absolute values, the range of values is reduced, which in turn reduces the space required for storage and transmission.

Advantages of Delta Encoding

- **Efficient for small variations**: When sensor readings change little over time, delta encoding significantly reduces the amount of data needed to represent those changes.
- **Space-efficient**: In sensor networks where many measurements are similar or change only slightly, delta encoding can greatly reduce the amount of transmitted data.
- **Improves compression**: By reducing the range of values (especially for closely spaced readings), delta encoding helps to further compress the data.

For instance, in a temperature monitoring system, if the temperature readings vary by only a small margin over time, delta encoding will reduce the size of the data that needs to be transmitted, leading to lower energy consumption.

4. Cluster Head Election

In hierarchical sensor networks, **Cluster Heads (CH)** are critical for aggregating data from the sensor nodes within a cluster and transmitting it to the base station. The **Cluster Head Election** process is vital for ensuring that the most suitable nodes are chosen to perform the task of data aggregation and compression.

To select the Cluster Head, the sensor nodes evaluate the distance from the base station, the energy available in the node, and the proximity to other nodes. The nodes within a 25-meter radius of each other form clusters, and the Cluster Head is elected from these nodes based on certain criteria, such as energy levels and proximity to other nodes. A typical approach involves selecting the node with the highest energy as the Cluster Head.

When a Cluster Head fails or runs out of energy, neighboring nodes detect the failure and initiate a new election process. The new CH election follows the same method, ensuring that the network can adapt dynamically to changes in the cluster head's availability.

5. Cluster Head Failure and Recovery

Cluster Head failure is an inevitable occurrence in wireless sensor networks due to the limited energy resources of sensor nodes. When a Cluster Head fails or its energy depletes, the neighboring nodes detect this failure and initiate the election of a new Cluster Head. The failed Cluster Head's neighboring nodes broadcast a message to all other nodes within the cluster to start the new CH election process.

This ensures that the network can continue to function even when individual nodes fail. The election and recovery mechanism helps maintain the robustness and energy efficiency of the network.

IV. RESULTS

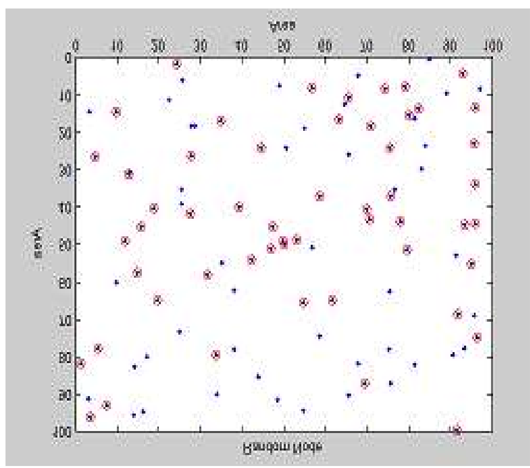

Fig.1 Random node

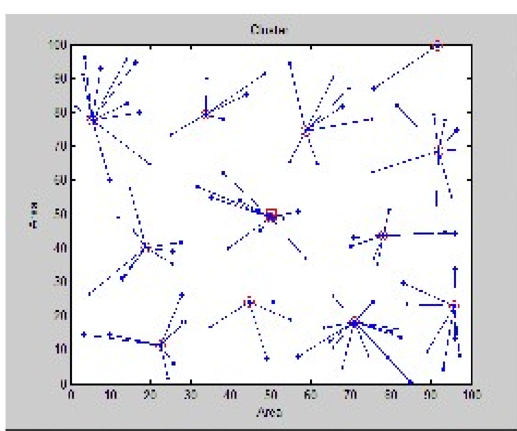

Fig.2 Forming cluster

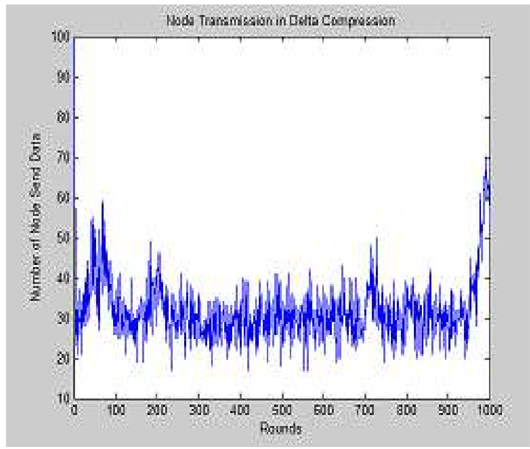

Fig.3 node transmission in Delta compression

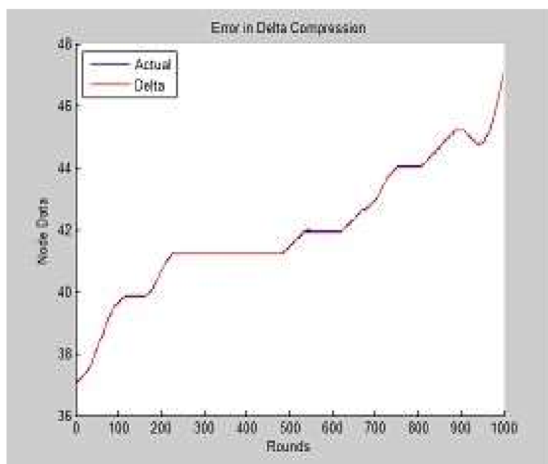

Fig.4 Error in delta compression

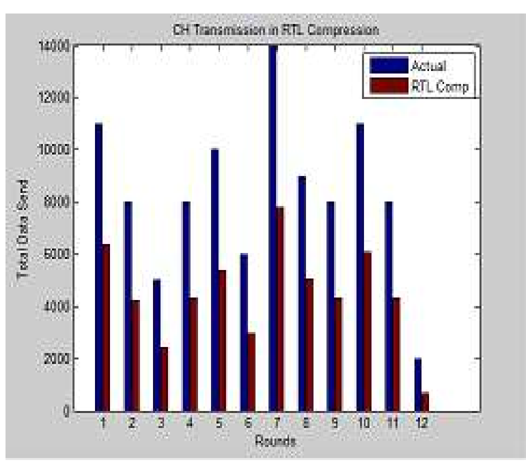

Fig.5 RTL compression

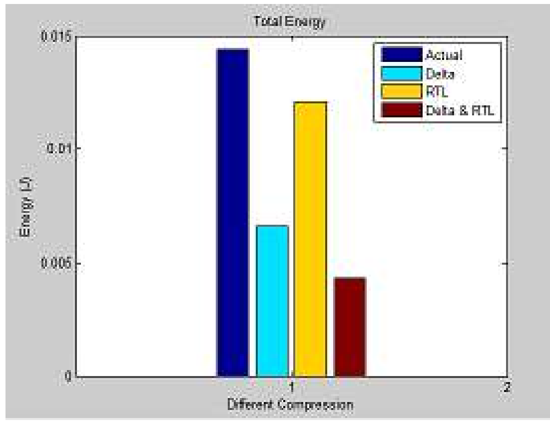

Fig.6 Total Energy

V. CONCLUSION

In this paper, we have proposed a data compression approach based on spatial and temporal correlation to reduce energy consumption in wireless sensor networks (WSNs). By grouping similar data and sending the differences between consecutive data points to the cluster head, the proposed method significantly reduces the bit rate required for data transmission, thereby extending the lifetime of sensor nodes. The use of Run-Length Encoding (RLE) for spatial correlation and Delta Encoding for temporal correlation has proven to be highly effective in compressing the sensor data before it is transmitted to the base station. The reduction in data size not only lowers the communication overhead but also helps conserve energy by minimizing the number of transmissions required. Simulation results indicate that the proposed algorithm outperforms conventional methods, achieving a 29% reduction in average energy consumption when compared to a network without clustering. This significant reduction demonstrates the efficiency of the clustering approach combined with data compression techniques in energy-constrained environments. The results show that the proposed algorithm is particularly effective when there is a high degree of spatial and temporal correlation in the data. This is common in many real-world applications, such as environmental monitoring and industrial sensor networks, where sensor readings are often correlated over time and space. Future work will explore the evaluation of different clustering algorithms and predictors to further improve the performance of the system. In particular, clustering algorithms that adapt dynamically to changes in the correlation characteristics of the data will be considered. Additionally, exploring the use of more sophisticated predictors and machine learning techniques could provide further improvements in data compression and energy efficiency.

REFERENCES

[1] Ruitao Xie, and Xiaohua Jia, *Fellow, "*Transmission Efficient Clustering Method for Wireless Sensor Networks using Compressive Sensing" in the year of FEB 2014.

[2] . J. Wang, S. Tang, B. Yin, and X.-Y. Li, "Data gathering in wireless sensor networks through intelligent compressive sensing,"in INFOCOM 2012, Mar. 2012.

[3] C. Luo, F. Wu, J. Sun, and C. W. Chen, "Compressive data gathering for large-scale wireless sensor networks," in Proc.ACM MobiCom'09, Sep. 2009.

[4] C. Luo, F. Wu, J. Sun, and C. W. Chen, "Efficient measurement generation and pervasive sparsity for compressive data gathering," *IEEE Trans. Wireless Commun.*, vol. 9, no. 12, pp.3728–3738, Dec.

[5] L. Xiang, J. Luo, and A. Vasilakos, "Compressed data aggregation for energy efficient wireless sensor networks," in *Proc .IEEE Sensor, Mesh and Ad Hoc Communications and Networks(SECON'11)*, Jun. 2011, pp. 46–54.

[6] D. B. Johnson and D. A. Maltz, "Dynamic source routing in ad hoc wireless networks," in *Mobile Computing*.

Kluwer Academic Publishers, 1996, pp. 153–181.

[7] C. Perkins and E. Royer, “Ad-hoc on-demand distance vector routing,” in *Mobile Computing Systems and Applications (WMCSA’99). Second IEEE Workshop on*, Feb. 1999, pp. 90–100.

[8] L. Xiang, J. Luo, and A. Vasilakos, “Compressed data aggregation for energy efficient wireless sensor networks,” in *Proc .IEEE Sensor, Mesh and Ad Hoc Communications and Networks(SECON’11)*, Jun. 2011, pp. 46–54.

[9] F. Fazel, M. Fazel, and M. Stojanovic, “Random access compressed sensing for energy-efficient underwater sensor networks,” *IEEE J. Sel. Areas Commun.*, vol. 29, no. 8, pp. 1660–1670, Sep. 2011.

Q. Zhang, X. Cheng, N. Zhang, Y. Cui, Y. Li, and Liang, “Sparse target counting and localization in sensor networks based on compressive sensing,” in *INFOCOM 2011*, Apr. 2011, pp. 2255–2263.