A SURVEY OF SPATIOTEMPORAL DATA COMPRESSION IN WIRELESS SENSOR NETWORKS

N.Boopal, Dr. S. Gunasekaran, V.Alamelu mangai

***Abstract*— The spatiotemporal compression is used to improve the efficiency, to maximize the utilization of the network and to reduce energy consumption. The various techniques are used for compression such as buffering method, suppression method, ordering method and clustering method. Among these methods, clustering method in the adaptive and composite compression improves compression ratio compared to the other methods. The comparative analysis is performed on the above methods.**

***Index Terms*— Aggregation, Order compression, Spatiotemporal compression, suppression.**

I. INTRODUCTION

The wireless Sensor network is the collection of sensor nodes that sends the data to the base station. Sensor nodes sense the information based on physical phenomena such as temperature, pressure, humidity, etc. Sensor nodes are referred as motes. It can be used in many application areas such as surveillance, volcano monitoring, environmental Monitoring, etc.. The key challenges in sensor networks are energy efficiency, storage, security and privacy. Among these challenges, the storage creates a lot of problems in wireless sensor networks, to eliminate this; data compression is an effective technique. The data compression is the method of reducing the number of bits needed to transmit or store data. It is mainly used to reduce the size of the data, save time and to reduce the energy consumption [3]. The spatiotemporal compression is one of the types of data compression method. The spatial compression is to find out the sensor nodes with similar readings. The temporal compression is carried out to eliminate the redundancy [1]. The spatial compression takes place on individual data to reduce space, whereas temporal compression occurs in a series of data and it removes the redundancy of the data. There are several methods developed for spatiotemporal data compression methods.

The four techniques used in spatiotemporal compression are Query sharing-reusing, redundancy removal, Compression based on spatial correlation, Compression based on temporal correlation, Data suppression on coding and decoding. Query sharing, aggregates and reduces the redundancy of the query [3]. The compression based on the spatial correlation combines the neighbor sensor nodes to collect the reading of the sensor nodes. Compression based on temporal correlation eliminates the redundancy, by comparing the data [1]. The Data Suppression on coding & decoding reduces the unnecessary data by encoding and decoding. This paper involves the spatiotemporal compression based on clustering, aggregation and ordering. The clustering is to satisfy scalability technique by grouping of sensor nodes into clusters, in which each cluster has a leader is referred as Cluster head, it performs the special function such as fusion and aggregation. The clustering reduces the number of nodes taking part in the transmission. The aggregation processes the query to obtain aggregate results from data collected by sensors. The aggregate queries executed functions such as min, max, sum, count, avg, median and the histogram [7]. The ordering is mainly used to encode the specific values, to locate the data history table and to optimize the data history table [3].

Manuscript received April, 2015.
N.Boopal *Assistant Professor,CIET,Coimbatore*
Dr. S. Gunasekaran*, Professor,CIET,Coimbatore*
V.Alamelu mangai*,PG Scholar,CIET,Coimbatore*

.,

II. SPATIOTEMPORAL COMPRESSION TECHNIQUES:

A. NETWORK COMPRESSION [10]

The Pipelined In-Network Compression (PINCO) eliminates the redundancy of the data obtained from sensors [10]. The data buffered in the network and combined groups of data through pipeline compression scheme. In the PINCO compression data method, data can be recompressed to eliminate the redundancy in the network. In this method, it reduces the latency by buffering. Thus, it eliminates the redundancy in spatiotemporal compression by compressing them into groups of data.

The data collection in wireless sensor networks is carried out by three methods is the plain data collection, network data aggregation and network compression. The plain data collection is a method where each sensor node sends its measurement to a base node at which data is stored and processed.

The network aggregation, aggregates the data collected from the different source nodes in order to reduce the communication overhead. The operations involved in aggregation function such as MIN, MAX, AVG and COUNT. The pipelined and interval based aggregation are two kinds of methods used in the network aggregation. The parent immediately receives data from the child node in the pipelined aggregation. Thus a parent node aggregates the result. In interval based aggregation, that node maintains the same interval time. Each child node waits for some time up to, which is assigned by the parent node.

In this network compression method, buffering takes place first. Before buffering it first converts into groups of data. The group of items consists of four attributes are defined as shared prefix, suffix list, node id list and timestamp list. The shared prefix length specifies are required to form a group. The suffix list specifies the least significant bits. Node identifier and timestamps match the order of sensor measurements. It can merge groups of data from children with newly produced groups of data with same prefix values. If a new data item is added, its suffix value is not inserted, but its node identifier and timestamp value are inserted. It again recompresses to reduce redundancy. PINCO involves some loss of data. PINCO data consider only for single valued data, it does not consider for complex data, such as audios and videos [10].

B. CODING BY ORDERING [5]

Aggregation and compression is done with coding by ordering method. In this method compressing can be done by ordering [5]. The ordering can be done by suppressing the data by encoding. Every sensor node compresses its own data based on its relative ordering. In this method the order of the data sent is not important. The border node and controller node are used in this approach. The border node receives the packet from other sensor nodes. It converts node packet into large super packet containing data of all nodes. The border node should include the id of each node because the entire packet reaches the controller at the same time. The border node has an option to any ordering of packets within super packets. This makes the border node to suppress the packets. Consider the nodes A, B, C, D, and E. The ordering can be done by three nodes at 3! =6 .The possible ordering from which the permutations are shown in table 2.1

Table 2.1: Permutation ordering

Integer	*Permutation*
0	*ABC*
1	*BAC*
2	*CAB*
3	*ACB*
4	*CBA*
5	*BCA*

To suppress l out of n packets, the relationship must be satisfied.

$$\frac{n!}{l!} \geq \binom{d-n+l}{l} k^{l} \qquad (1)$$

Here 'n' be the number of packets in the encoder 'k' be range of possible values generated by each sensor and 'd' range of node ids of the sensor[7], l be the largest number of packets that can be suppressed.

Striling's equation can reduce inequality to equivalent one

$$(n+0.5)\ln n + (d-n+0.5)\ln(d-n) + 0.5\ln(d-n+l) + (d-n+l)[\ln(d-n+l)-1] + l\ln k \geq 0 \qquad (2)$$

The equation (1) and (2) is necessary for suppressing the packets. At first payload can be encoded and packet suppression can be done. There is more freedom to choose the packets in which it is to be suppressed.

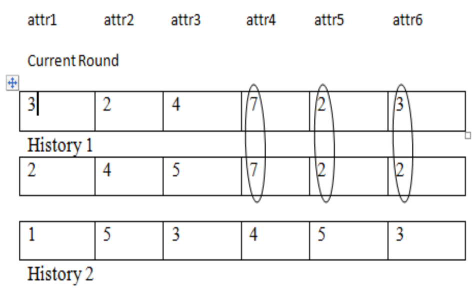
There is no efficient algorithm for mapping permutation to a data value, since permutation values obtained during time of compression are very large.

C. ORDER COMPRESSION METHOD[3]

The data suppression is a technique in which source node sends newly collected data to aggregate node in which suppression is carried out [3]. The data loss is one of the problems with suppression technique. In order to avoid this problem, suppressed data can be carried out without loss. It can compress the data size using network prediction. The data record collected by sensor node is carried out by two methods in which first method leads to transmission back to the parent and the second method is to suppress the data transmission. The steps involved in data compression are data aggregate query, prediction with weighted data trend and order compression. Data aggregate query combines the high level user queries. SQL query language is used for data aggregation. Prediction with weighted data trend is used to evaluate data aggregation queries. The sensor report consists of m different attributes based on which data aggregation operations are performed. Consider, for n different attributes, it performs ordering as m!

The order compression algorithm first checks whether there is a data aggregate query. If it is not a data aggregate query, normal data processing technique is used to process the query evaluation. Based on network prediction model coherency c is checked. In compression algorithm, it checks for reported attributes. If the attj has a similar value to atti, the order of previous attribute will be used to report the location.

Fig.1: Ordered values in sensor network



The current collected record consists of six attributes in the sensor nodes. Consider attributes in which 3rd and 4th attributes has no change in the value of record. It immediately cancelled the attributes based on the network suppression. Next it compares the value with previously collected value in the data history table. Every sensor node maintains the data history table. If it remains the same value it cancels the transmission, but it maintains the order. The sensor node has m different attributes. It performs the order operations as m!. The order compression reduces the processing time for some data attributes to optimize data history tables.

D. ADAPTIVE AND COMPRESSIVE METHOD [1]

Adaptive and composite method uses the possible data delivery latency tolerance to minimize the amount of data transported to the sink [1]. It maintains the accuracy of the data. In this method, it deals with the continuous data collection. The continuous data collection involves the real time decision making and delay tolerant process. The adaptive modeling algorithm uses the delay tolerance of the data collection to maximize the data compression. For some applications like volcano monitoring and ecosystem it involves high spatiotemporal resolution. This needs to send large volume of data so that it leads to high energy consumption.

The temporal correlation and redundant deployment reduces the communication overhead. The redundant deployment reduces the redundancy in the data. The compression of sampled attribute temporally is referred as temporal correlation. It considers the precedence about the quality and quantity of the data. This latency tolerance improves the efficiency in wireless sensor network. The adaptive and composite method involves three stages are 1 hop cluster formation, temporal modeling of cluster and merging 1 hop cluster.

In the first stage, one hop cluster formation uses the redundant deployment, it is compared with the historical values reduces the redundancy based on its data is transmitted to the sensor node. In stage two it uses the temporal correlation by constructing the models and produces a small number of clusters is referred as master cluster. In stage three, master cluster sends the model to the neighboring cluster. It accepts or rejects based on their sampled model. If it is accepted, it forms a large cluster. The stage one is executed only once, Stage two and three are continuously executed in order to reduce space and time. The main drawback of this method leads to node crashes and it does not extend to multivariate compression.

III. COMPARATIVE STUDY

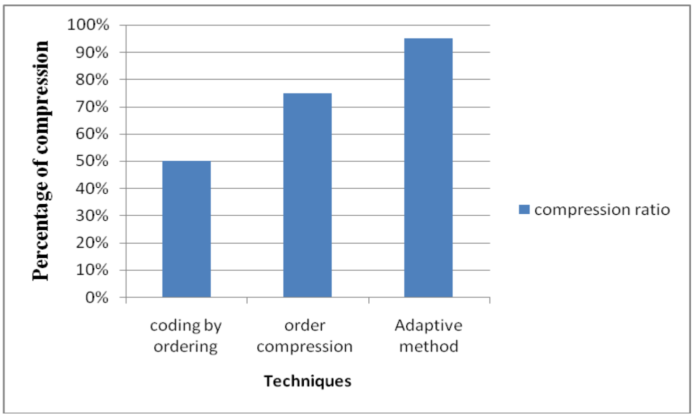
The comparative study of the network compression, coding by ordering, order compression, and adaptive and compressive method are given below

S.no	Name	Advantage	Disadvantage
1	Network Compression	Compression can be done again in order to reduce redundancy[10]	Only Single valued data[10]
2	Coding by ordering	Removes the similarities of the data to greater extend[5]	Need not consider the order of data[5]
3	Order compression	Reduces the processing time of data attributes[3]	Maintaining data history table is difficult[3]
4	Adaptive and compressive	Minimizes the amount of data transferred to the sink[1]	Leads to node crashes[1]

IV. COMPARISON CHART:

Following figure shows the comparison of spatiotemporal data compression of wireless sensor networks. The comparison is done based on compression ratio. The significance of the comparison is that Adaptive and Compressive method achieves higher performance when compared with other techniques.

Fig. 2: Analysis of Compression methods



V. CONCLUSION

In this survey the four different types of spatiotemporal data compression such as network compression, adaptive and composite compression, order compression and coding by ordering are discussed. Adaptive and composite method achieves a high compression ratio while compared to the other methods. The adaptive and composite method minimizes the amount of data transferred to the sink, reduces the processing time of data attributes and it uses the buffering method to remove the similarities of the data. It is necessary to consider for the complex data, such as audio and video. So the future works may be lying on multi-valued data.

REFERENCES

[1] A.Ail,A.Khelil, P.Szczytowski, N.Suri, "An adaptive and composite spatio-temporal data compression approach for wireless sensor networks" ,in: Proc.of ACM , pp.67–76,2011.

[2] Akyldiz IF, Su W, Sankarasubramanian Y, Cayirci E." A survey on sensor networks". IEEE Communications Magazine, vol 40, no 8,pp.102–114,2002.

[3] C.Yang, Z.Yang,K.Ren, C.Liu, "Transmission reduction based on order compression of compound aggregate data over wireless sensor networks", International Conference on Pervasive Computing and Applications(ICPCA'11) , pp.335–342,2011.

[4] C.Chong and S. Kumar,"Sensor Networks: Evolution, oppurtunities, and Challenges," Proceedings of IEEE,vol. 91,no 8, pp.1247-1256,2003.

[5] D. Petrovic, R. C. Shah, K. Ramchandran, and J. Rabaey, "Data Funneling: Routing with Aggregation and Compression for Wireless Sensor Networks," In Proceedings of First IEEE International Workshop on Sensor Network Protocols and Applications, May 2003.

[6] Jeong-Joon Kim, In-Su Shin, Yan-Sheng Zhang, Dong-Oh Kim, and Ki-Joon Han, "Aggregate Queries in Wireless Sensor Networks",International Journal of Distributed Sensor Networks, vol 2012 .

[7] Kimura and S. Latiti, "A Survey on Data Compression in Wireless Sensor Networks," Proc. of the International Conference on Information Technology: Coding and Computing (ITCC'05), pp. 1-6,2005.

[8] N.D.Phame et al. "SCCS Spatiotemporal clustering and compressing schemes for efficient data collection applications in WSNs". International Journals of Communication Systems ,vol 23, no 11,2010

[9] S.H.Yoon, C.Shahabi,"An experimental study of the effectiveness of clustered aggregation (CAG) leveraging spatial and temporal correlations in wireless sensor networks", ACM Transactions in Sensor Networks ,vol 3, no 1,pp. 1–36 March 2007.

[10] T. Arici, B. Gedik, Y. Altunbasak, and L. Liu, "PINCO: a Pipelined In-Network Compression Scheme for Data Collection in Wireless Sensor Networks," International Conference on Computer Communications and Networks, October 2003.

