International Journal of Engineering Innovation & Research
Volume 2, Issue 4, ISSN: 2277 – 5668

Query Routing and Data Aggregation in Wireless Sensor Networks: A Review from Database Perspective

Srinivasa.K
Department of Computer Science and Engineering,
Siddaganga Institute of Technology,Tumkur
Email: srinivas.karur@gmail.com

***Abstract* - In contrast to the traditional applications of sensor nodes, their computation, storage and wireless communication abilities changed the era from past few years, from just sensing applications to the large number of monitoring and tracking applications. From database perspective, a sensor network can be viewed as a large distributed database system. Hence each sensor node behaves like a tiny database management system which accepts, processes and answers queries submitted by the users. Most data queries over sensor networks are spatial in their nature due to the deployment of sensor nodes over large geographical areas. Due to the limited energy constraints and unavailability of the global view of the sensor nodes, the methods used for processing spatial queries in centralized database systems can't be applicable directly to wireless sensor networks. In this paper, a review on spatial query processing in wireless sensor networks is considered with the database perspective through a taxonomy and important research challenges are also highlighted for the future work.**

***Keywords* - Base Station, Correlation, Data Aggregation, Sink and Spatial Queries.**

I. Introduction

Today's sensors are no longer just simple sensing devices wired to a central monitoring site, but they have computation, storage and wireless communication capabilities. This has opened new opportunities of Wireless Sensor Networks (WSNs) for a variety of monitoring and tracking applications. Sensor nodes in WSNs are battery operated, which highly constrains their life span and it is often not possible to replace the power source of thousands of sensors deployed over a large geographical area. This necessitates the need for methods to increase the lifetime of the WSNs by conserving the nodes energy.

From a database perspective, a sensor network can be seen as a distributed database system (DBS) with sensor nodes that are tiny DBS themselves. In such case, WSNs are query driven and thus, a user can interact with the sensor network as a whole, and send queries to it. The sensor nodes run DBS locally, and participate in global query execution. This reduces a lot of unnecessary transmissions by limiting the data to the user interest and at the same time needs an energy efficient query processing method. The primary goal in query processing in sensor networks is to effectively answer queries while transmitting as little data as possible.

WSNs can be queried for values of certain attributes without specifying the location or the region as a query constraint. But, most data queries over sensor networks are spatial in their nature i.e. query the network based on location due to the deployment of sensor nodes in the physical world. The spatial query processing techniques such as indexing techniques used for centralized database systems can't be applied directly to WSNs due to its decentralized nature, energy limitation and unavailability of the global view of the network and hence this opens a new challenge in WSNs.

Many routing protocols have been proposed to route a packet to a specific location in the sensor network. Direct Diffusion [14] forwards the request based on the sender's interest such as location. Geographic based routing [15], [16] use geographic coordinates to route queries to a specific sensor. Unlike the general routing protocols, this paper focuses on running spatial queries that query the sensor network for sensors in a specific area. Although collecting spatial query results in a spatially enabled sensor network is the same as collecting normal or attribute query (query without spatial constraint) results, we can significantly reduce the energy consumption in processing spatial queries by exploiting the fact that, the sensors with the query results are usually located in the same geographical area and exploit high degree of correlation. This leads to a new approach for collecting the data which is called as correlation aware data aggregation, which considers correlation among the data from the sensors which are placed closely.

In this paper we address the issues and techniques involved in processing spatial queries in WSNs from database perspective through a taxonomy in detail.

The paper is organized as follows: Section 2 gives the motivation for the work. Section 3 describes the review of the survey work on the considered topic along with the necessary classifications in each case. Section 4 discusses the new paradigm of data aggregation used when the data reported by the sensor nodes are correlated and finally Section 5 concludes the topic by highlighting the open challenges in the field as a future work.

II. Motivation

WSNs gained importance for their sensing, processing and data dissemination capabilities, with the potential to serve many monitoring and tracking applications. Due to the energy constraints of the sensor nodes, extending the life time of these networks plays a significant role among other challenges like nodes deployment, topology management, area coverage and so on.

International Journal of Engineering Innovation & Research
Volume 2, Issue 4, ISSN: 2277 – 5668

Compare to sensing and computation, communication in particular is significantly more expensive operation for sensor nodes and it's often proved that, the energy consumed for 1 bit of data transmission can be used for processing 1000 instructions [13]. Hence most part of reducing the energy consumption lies in reducing the amount of data transmission and hence the cost of communication in WSNs.

This opens a new challenge for the database community by viewing the WSN as a distributed database and queries the sensor network for the data of interest to minimize the transmission of the unwanted data.

Queries used can be either attribute based or location based, based on the awareness of nodes about their location in the network. If nodes in the WNS are not aware of its geographical location in which they deployed, the WSN will be queried by attributes. In contrast, if the nodes know their location, they will be queried through some spatial constraints which represent the location of the nodes or a part of the region in the network.

Most of the queries in WSNs are location based due to the deployment of the nodes on physical world. Due to the unavailability of the global knowledge of the sensor nodes in the network and their limited energy constraints, necessitates the need for an energy efficient method for processing of such queries in a distributed manner due to the high distributed nature of sensor nodes.

III. Spatial Query Processing in WSNs

The following sections review the literature for spatial query processing issues in WSNs. Section 3.1 briefly discusses the various data delivery models used in WSNs along with highlighting the query driven model which is used here from database perspective. Section 3.2 defines spatial queries formally and lists different variations of the spatial queries used in different works. Section 3.3 describes the common steps involved in processing spatial queries. Finally Section 3.4 defines a taxonomy for spatial query processing techniques.

A. Data Delivery Model

From the application perspective, different data delivery or dissemination models are used in WSNs, each gives a way of data extraction from the WSNs and are as shown in the Figure 1 [29,28,2,4,10,17].

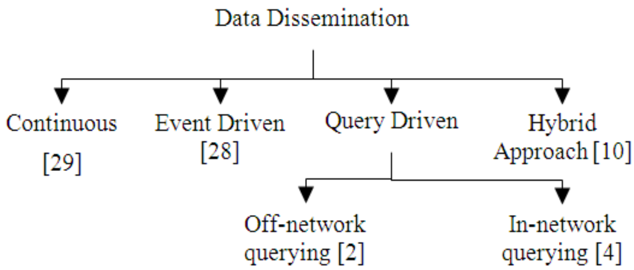

Fig.1. Data dissemination models

In traditional WSN applications, each sensor node transfers the sensed data continuously at a pre-specified rate to a central location which is normally termed as base station (BS). This model is considered as continuous.

Certain applications in which the sensor nodes communicate data among themselves or to the BS whenever they detect an event are called as event-driven WSNs. Such event-driven WSNs comprise the methods for detecting the events by gathering raw data from one or several sensors, recognizing a previously learned pattern in the raw data and mapping this pattern to an event that is semantically relevant to the application of WSN.

Third model which is observer initiated model makes use of storing and processing capabilities along with communication and sensing capabilities of sensor nodes. Unlike the previous two models, this model reduces the amount of unwanted data transfers from the sensors by transferring the data based on the user generated requests in terms of queries and hence is also called as query driven model. Based on the application requirements, the queries are generated either from BS or a sink node[1] which is placed either inside or outside the network or from any node within the network. In the former case it is considered as off-network querying and in the later as in-network querying. From the database perspective, our further discussion is based on this model.

Finally, the query driven and event driven approaches can coexist in the same network. This is referred as hybrid model.

B. Spatial Queries

Spatial queries are similar to the normal attribute based queries with the additional one or more spatial constraints which represents the area of interest. These are used to answer questions such as "*what is the average temperature in the region r?*". The area of interest 'r', is a subset of a Region 'R' in which the sensor nodes are deployed. R can be considered as a 'rectangle', if the area of nodes is assumed to be in Cartesian co-ordinate system or 'circular', if the area is assumed to be in polar co-ordinate system. Sensors located at position 'p', satisfies the spatial query constraint if p is inside r. Such queries in which the database or the sensor network is queried by location of the nodes or region of the network are called as spatial queries [1]. Few of the variations of the spatial queries used in different works are listed below:

1. Nearest Neighbor (NN) queries [4]: Nearest Neighbor (NN) query is defined as, finding the data object which is closest to the querying object given a set of objects. For ex: "What is the location of the nearest data object to coordinates (x, y)?".
2. Aggregated (A) queries: Using these queries, aggregated value of some attribute/s using aggregation functions such as min, max, avg, sum etc. from the sensors fall into a query region is asked. For ex: "SELECT AVG(*humidity*) FROM *sensors* WHERE *location* INSIDE *Room*-101 AND *temperature* < *threshold* ".

1. Query originating nodes and nodes which have query results are normally termed as sink and source nodes respectively. Sink nodes also responsible for collecting the results.

International Journal of Engineering Innovation & Research
Volume 2, Issue 4, ISSN: 2277 – 5668

a. With time window (TW)[2]: This specifies a time interval for which the data values need to be fetched from sensors with an optional field of total duration for which query should run. For ex: "get average temperature of section 'X' every minute for next 600 minutes".
b. Without time window (WTW): This is a just Aggregate query without the time window.

3. Spatial Pattern queries [7]: These are the kind of spatial join queries such as: "what is the high temperature reading in the vicinity of at least four high pressure readings".
4. Range Based queries [9][10]: It's an obvious extension to the NN queries where the data objects that fall into a query region is asked. Here queries may specify a temporal range and a spatial area such as "What was the lowest humidity recorded yesterday morning in the Lake Annete area?" Following are some of the range query examples in WSNs: What are the nodes whose temperature is great than T1 within a certain area? Reporting or counting the sensor nodes within the temperature range [T1,T2] and pressure range [P1,P2] ? In general cases, the queries are associated with geographic constrains, e.g, report the nodes with temperature greater than T1 within the rectangle area enclosed by [x1,x2] and [y1,y2]. If there is no explicit range constrains in the query, e.g, reporting the nodes with the temperature in [T1,T2] and pressure in [P1,P2], by default the query range is the entire coverage area. This is considered as a special case in complex range query.
5. Selective Aggregate Queries [6]: Queries that aggregate data only from a subset of all network nodes.
6. Spatial queries about the location of the data in the network (LD) [8]: Example query: "What are the location of all critical events within a range R from a point with coordinates (x,y)?"
7. Spatio-Temporal queries [11]: Queries include both the spatial and temporal constraints. For ex: "What was the temperature yesterday morning in area X?".

C. An Overview of Query Processing Methodology

Main issues with query processing in WSN's are as shown in the Figure 2.

As a first phase, two of the popular choices used in the design of routing protocols for query processing include:

- Construction of a routing tree or determining a routing path which will be used further for routing the query to and collecting the data from source nodes to sink node. Here the methods can use the same path to send the result as used by the query or use the different path as a further measure of energy efficiency.
- Using the location information of sensors and the sink to forward query and results.

SPIX [1], and GPSR [18] can be considered as representative examples of first and second choices, respectively. Unlike the second choice which is used to route the query to a sensor at specific location, here focus is given to the first choice to query the WSN for sensors in a specific area or region.

With respect to the spatial query processing this phase involves, developing methods which can send query to the user interested region rather than flooding the query to the whole network and minimizing the number of non-participating nodes in the collection path while collecting the result from source region to the sink. This can be summarized as:

- Bound the branches that do not lead to any result from reaching the query in the routing tree.
- Find a path to the sensors that might have a result.
- Few of the existing techniques which will be explained in the next section, uses the indexing methods to achieve the above two requirements. Based on these discussions, the unique characteristics of the WSNs that generate new challenges for processing spatial queries involve:
- Distributed query execution: Queries must run in a distributed manner, because sensor data are distributed in the network, and there is no global view of the data.
- Distributed and dynamic index maintenance: The high energy cost of communication requires deciding where to run the query to optimize the resource usage. In addition, sensors may fail at any time and thus the spatial index must reorganize itself, in case if indexing technique used.

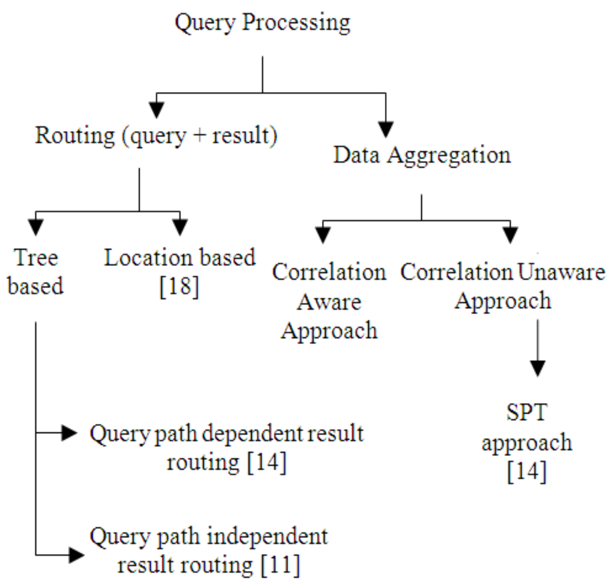

Fig.2. Query processing issues

Second phase involves processing the data as it moves from the source nodes towards the sink in the routing tree which reduces the number of packets transferred and hence saves the energy. This is termed as in-network aggregation or data aggregation. In achieving this, each sensor node uses a so called aggregation function to encode the data available at that node before forwarding it to the sink.

Due to the deployment of the sensors in the physical world, different sensor nodes partially monitor the same spatial region and data is often correlated. By exploiting

International Journal of Engineering Innovation & Research
Volume 2, Issue 4, ISSN: 2277 – 5668

the correlation among the nodes, more amounts of data transferred can be reduced compare to the methods which doesn't consider correlation among the nodes[ ].

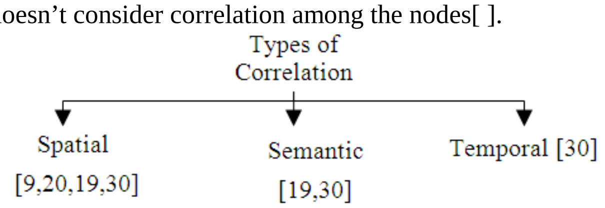

Fig.3. Types of Correlation

Accordingly, if any data aggregation approach takes correlation among the nodes into consideration, it's called as correlation aware data aggregation and conversely other as correlation unaware data aggregation approach. Here the focus is given to correlation aware data aggregation scheme. Correlation unaware data aggregation schemes involve methods that use Shortest Path Tree (SPT) or their approximations as their data aggregation backbone.

Correlations exist can be of three types as shown in the Figure 3[9,20,19,30].

- *Spatial Correlation:* Given the typical dense deployments of sensors in WSNs, it is likely that the sensing regions of two different sensors within the rectangular region overlap. Consequently, the data reported by these sensors are spatially correlated. If the two sensors are very close to each other, the data reported by both sensors is practically the same, which implies that the sensors are perfectly correlated.
- *Semantic Correlation:* Responses to statistical queries such as min, avg, max, usually fall under this category. If the query imposed is about certain statistics about an event such as: "is the total number of cars in the rectangular region (x1, y1, x2, y2) greater than K?", then even if the sensors are reporting data about different cars, the information reported may be correlated based on semantic of the query, as it is only required to find the total number of cars and consequently determine if it is greater than K.
- *Temporal Correlation:* When the sensor readings change slowly over time, the values of sensors can be considered as temporally correlated.

  Hence the second phase involves building an energy efficient data aggregation backbone which performs routing and compressing the data together while routing the result from source nodes to the sink node. The details about the correlation aware data aggregation are described in section 5.

D. Taxonomy

Figure 4 depicts the classification criteria for Spatial Query Processing methods in WSNs. The existing methods classified based on:

- methods which uses indexing
- network topology used &
- query originator

Few of the existing techniques found from the survey which comes in each category are also mentioned and explained in brief below. Summary of these techniques with some of the common properties of each is listed in Table 1.

SPIX [1]: SPatial IndeX (SPIX) method creates a hierarchical structure in the network, which is an index structure built on top of the sensor network and forms a routing tree to route the query and get back the result. The spatial query will be disseminated into the routing tree and the result will be sent back to the root (BS). When a sensor receives the query, it must decide if the query applies locally and/or needs to be submitted to one of its children in the routing tree. A query applies locally if there is a non-zero probability that the sensor produces a result for the query. This is achieved by maintaining minimum bounded area (MBA) as an index at each node which covers area information of itself and its children.

STI [2]: This method is advancement over the previous one and uses the same indexing technique as [1]. Here they proposed a Sectioned Tree Indexing, which divides the network area into several squares and each square has a local index subtree organized within that square. Local trees are interconnected to form one big tree in the network. In this method they showed that, MBA indexing of [1] itself is not very energy efficient when a query region includes two or more branches from a node that has a common nearest ancestor.

PT [4]: This method proposes a peer-to-peer version of the centralized R-tree index structure and is called as Peer Tree structure which addresses the problem of nearest-neighbor queries in sensor networks. In this model a query can be posed in any node of the network without the need of a central server.

SAP [5]: Here methods to collect the aggregated value from spatially distributed nodes were proposed which take into consideration the distribution of the values generated by the sensor nodes. They proposed the use of spatial interpolation methods derived from the fields of spatial statistics and computational geometry to answer spatial aggregations and developed prototypes for distributed Spatial Aggregate Processing.

PDT [6]: In this work the focus was on selective aggregate queries and presents an aggregation algorithm called Pocket Driven Trajectories (PDT) that minimizes the number of nodes used in data aggregation i.e. that minimizes the number of non-participating nodes in the path of data aggregation.

ESPQ [7]: Here protocols were devised for 'in-network' Evaluation of Spatial Pattern Queries. Here they provide the cost models which can be used by a query optimizer to determine a suitable evaluation method based on query parameters and data statistics.

ADI [8]: Here an energy efficient hierarchical indexing approach called Adaptive Distributed Indexing for spatial data in sensor networks was proposed. This indexing technique allows roaming users to navigate through sensor networks distributed over large geographical areas and to pose spatial queries about the location of the data in the network. The major challenge in designing such indexes is the minimization of the total amount of traffic needed to create and maintain the indexes, which is a function of region activity and the actual query rates.

SWIP [9]: Spatial Window Processing is a framework developed for processing spatial range queries. This framework has two phases. In the first phase, a path from the query originator node to a suitable sensor node located in the query region will be constructed. Due to its role in the second phase of SWIP, located node termed as the *query coordinator*. For the second phase, the coordinator node disseminates the query to the sensor nodes located within the query region, collects the query answers from them, and returns the answers to the originator node on the path discovered during the first phase. Within this framework, they proposed the spatial window flooding (SWIF), a processing strategy formed by two algorithms, one for each phase.

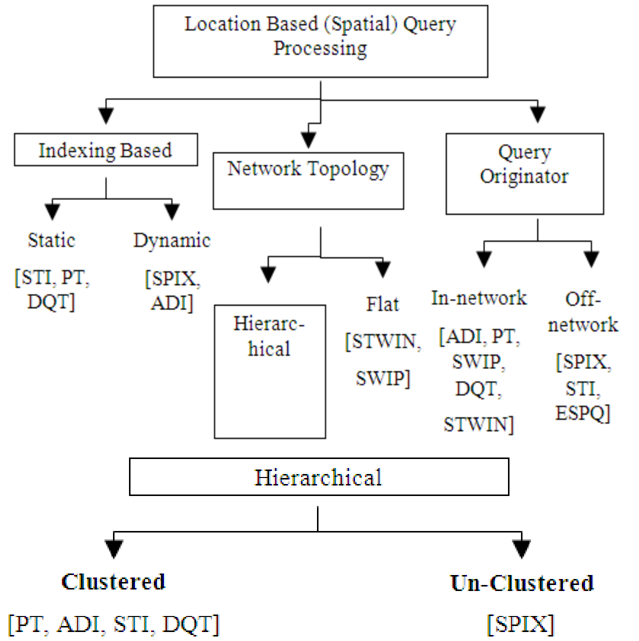

Fig.4. Taxonomy of Spatial Query Processing Techniques

DQT [10]: By exploiting the location information and geometry of the network an in-network querying infrastructure was proposed and named as Distributed Quad Tree structure. DQT can handle range queries and it satisfies efficient in-network information storage as well as distance-sensitive querying i.e. the cost of answering a query for an event is at most a constant factor of the distance 'd' to the nearest event in the network.

STWIN [11]: Here three algorithms proposed, one for the first phase which is used to search for a path from the query originator node to a sensor located within the query's spatial window and two algorithms for the second phase where located sensor assumes query coordinator role and gathers the answers from all query relevant sensors and ships them back to the query originator. Here a greedy routing algorithm used in the first phase, while for the second phase a parallel flooding, the other using a depth first strategy proposed.

IV. Correlation Aware Data Aggregation

For data gathering applications in which data originates at multiple correlated sources and is routed to a single sink, aggregation would primarily involve in-network compression of the data along with finding an optimal routing strategy for routing the data towards sink.

Here an effort is made in highlighting correlation aware data aggregation schemes and the issues related with them to the best of our knowledge from the survey.

Three schemes that helps in understanding the interaction between routing and compression is as shown in the Figure 5 [20].

(1) Distributed Source Coding (DSC): If the sensor nodes have perfect knowledge about their correlations, they can encode/compress data so as to avoid transmitting redundant information. In this case, each source can send its data to the sink along the shortest path possible without the need for intermediate aggregation.

(2) Routing Driven Compression (RDC): In this scheme, the sensor nodes do not have any knowledge about their correlations and send data along the shortest paths to the sink while allowing for opportunistic aggregation wherever the paths overlap.

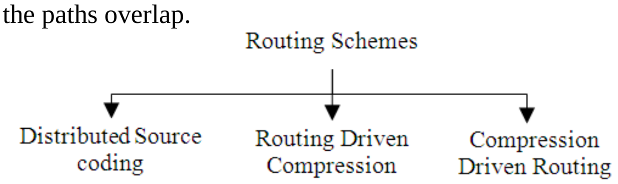

Fig.5. Interaction of routing with compression

(3) Compression Driven Routing (CDR): As in RDC, nodes have no knowledge of the correlations but the data is aggregated close to the sources and initially routed so as to allow for maximum possible aggregation at each hop. Eventually, this leads to the collection of data removed of all redundancy at a central source from where it is sent to the sink along the shortest possible path.

During the aggregation, a sensor node uses a so called aggregation function to encode the data available at that node before forwarding it to the sink. Several coding strategies were proposed in recent research that can be classified as shown in the Figure 6 [21]:

On the one hand there are the so called multi-input coding strategies where aggregation is performed at a node only if all input information from multiple nodes is available in order to exploit correlation among several nodes. On the other hand, there also exist single-input coding strategies where the encoding of a node's information only depends on the information of one other node.

There exist two classes of single-input coding, namely self coding and foreign coding. Using self-coding, data is only allowed to be encoded at the producing node and only in the presence of side information from at least another node. With foreign coding in contrast, a node is only able to encode raw data originating at another node as it is routed towards the sink using its own data.

The available methods which deal with correlation aware aggregation are categorized under three schemes as shown in the Figure 7.

Clustered in-network aggregation approaches exploit the spatial correlation of sensor readings to preserve node energy. The Clustered Aggregation (CAG) and Geographic Adaptive Fidelity (GAF) approaches, use the fact for spatially correlated nodes form clusters so that only one node needs to respond an aggregation query [22] [23]. Similarly, in static clustering [20] the network is statically partitioned into grid cells. For each grid cell one node is appointed as a cluster head node that acts as a gateway node. In each cell data is routed via a local shortest path tree, aggregated at the local gateway and then communicated directly to the sink.

In [24], the authors introduced techniques to exploit spatial correlations by selecting a small subset of sensors called correlation dominating set, whose data values are sufficient to interpolate value at all points with sufficient accuracy. A hyper graph is used to represent the correlation correlation structure. As computing the minimum dominating set problem is NP-hard, they proposed some approximation algorithms.

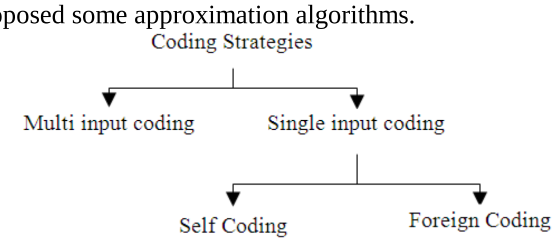

Fig.6. Coding schemes used for Data Aggregation

In [25], the authors proposed the distributed compression algorithm in which encoding is achieved in a completely blind manner without any side information of the correlated sensors. The main advantage of this scheme is that it is flexible to capture spatio-temporal correlation structure without requiring any state to be maintained at the individual nodes.

In [26], the authors investigated the interaction between the correlation structure and the routing tree to transport the data. They considered two approaches. The first considers both compression and routing independently, where each node has the global correlation structure to perform compression using Slepian-Wolf coding. In the latter, source coding is performed with some side information obtained through explicit communications as part of the transport toward the sink.

In [21], the authors focused on single-input coding strategies as they feature several advantages against multi-input coding strategies. Since the problem of finding a minimum-energy data gathering tree is NP-complete for optimal self-coding, they proposed an approximation algorithm. This algorithm, termed LEGA, is based on shallow light tree that combines the properties of the spanning tree and the shortest path tree. For foreign coding, they introduced the algorithm termed the minimum-energy gathering algorithm (MEGA) that gives an optimal solution.

Few of the schemes proposed correlation aware data aggregation tree structures and are categorized under two schemes as shown in the Figure 8.

When the full knowledge about source location is known, the Steiner tree over all sources, sink and non source nodes gives the optimal message cost when the degree of correlation[2] is close to 1. However, the computation of Steiner tree is an NP-hard problem [27].

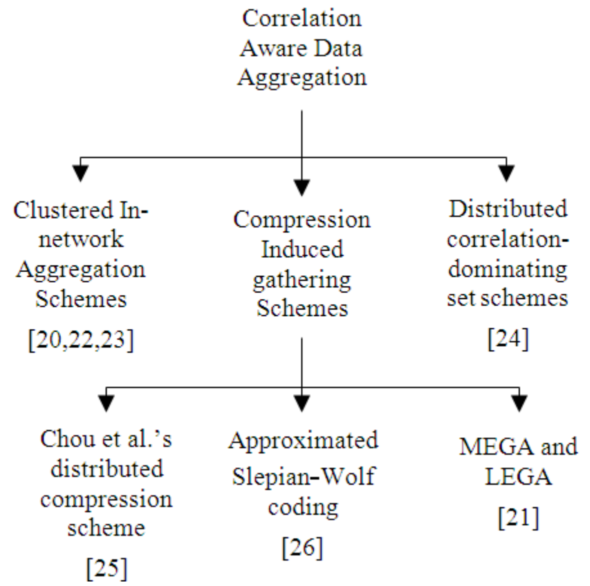

Fig.7. Correlation Aware Data Aggregation schemes

[3, 12] address the more general problem of building aggregation structure with optimal expected cost when the knowledge of sources is incomplete.

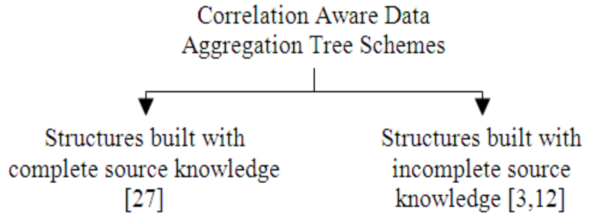

Fig.8. Correlation Aware Data Aggregation Tree Schemes

In [19] authors proposed a called Semantic/Spatial Correlation Tree (SCT). SCT is based on the identification of an aggregation backbone which is used to generate efficient aggregation trees, regardless of sources distribution and density. The aim is to efficiently build and maintain a network structure for data aggregation. To this end, the authors of [19] propose a ring-sector subdivision of the network. A subset of nodes is elected as aggregation nodes and they are organized in a spanning tree to form the data aggregation backbone. Each node belonging to the backbone aggregates messages coming from a certain sub-area.

2. Degree of correlation characterizes the correlation among the sensors data. For ex: If it is equal to 1, then sensors are perfectly correlated

V. Conclusion and Future Work

This paper discusses a number of issues related to spatial query processing in WSNs from database perspective like, data delivery models, spatial queries and the several variations of these queries used in different works, their importance in WSNs, steps involved in in-network query processing in WSNs, a literature review of the existing techniques with a taxonomy and finally correlation aware data aggregation which is a new paradigm for aggregation used in environments where the sensors data are correlated.

From the discussion, we note that most of the existing research on spatial query processing focuses on routing, often considering only very simple approaches to aggregate data. We also note that most of the methods are based on clustering. The major advantage with this approach is, it limits the aggregation at the cluster heads only by avoiding the overhead with more number of aggregations.

Even though, existing methods proposed various indexing schemes for locating the answers in an efficient manner for a given query, very few considered the issue of dynamic index management which is very crucial in WSNs environment if any sensor leaves or lost its energy. Sometimes managing the index dynamically may consume more energy than the underlying communication scheme.

Though much of the work is going on in data gathering from correlated nodes, a very few methods tried to combine the correlation aware data aggregation with spatial query processing, as correlation aware data aggregation places an important role in this context due to the spatil-temporal constraints and semantic nature of the spatial queries submitted to the WSNs. Even though few of the schemes deals with spatial and temporal correlation, very less work has been done in case of semantic correlation.

References

[1] Amir Soheili, Vana Kalogeraki and Dimitrios Gunopulos. Spatial Queries in Sensor Networks.GIS'05.ACM.November 4.2005.

[2] Kyungseo Park, Byoungyong Lee and Ramez Elmasri. Energy Efficient Spatial Query Processing in Wireless Sensor Networks. 21st International Conference on Advanced Information Networking and Applications Workshops (AINAW'07). IEEE.2007.

[3] D.R. Karger, M. Minkoff. Building Steiner trees with incomplete global knowledge.41st Annual IEEE Symposium on Foundations of Computer Science (FOCS'00).2000.pp 613–623.

[4] Hakan Ferhatosmanoglu, M. Demirbas. Peer-to-peer spatial queries in Sensor Networks.3rd IEEE International Conference on Peer-to-Peer Computing (P2P'03).2003.

[5] Mehdi Sharifzadeh,Cyrus Shahabi. Supporting Spatial Aggregation in Sensor Network Databases. GIS'04.November 12–13.2004.

[6] Muhammad Umer, Lars Kulik and Egemen Tanin. A Location Based Aggregation Algorithm for Selective Aggregate Queries in Sensor Networks.2nd International Conference on Geosensor Networks (GSN'06).2006

[7] Man Lung Yiu, Nikos Mamoulis and Spiridon Bakiras. Evaluation of Spatial Pattern Queries in Sensor Networks. HKU CS Tech Report TR-2007-02.

[8] Vladimir Dyo,Cecilia Mascolo. Adaptive Distributed Indexing for Spatial Queries in Sensor Networks.IEEE Proceedings of 8th International Workshop on Mobility in Databases and Distributed Systems (co-located with DEXA05). August 2005.IEEE Computer Society Press.

[9] Alexandru Coman,Joerg Sander,Mario A. Nascimento. Adaptive processing of historical spatial range queries in peer-to-peer sensor networks.Springer.Distrib Parallel Databases (2007) 22: 133–163.

[10] Murat Demirbas, Xuming Lu. Distributed Quad-Tree for Spatial Querying in Wireless Sensor Networks. Proc. IEEE Int'l Conf. Comm. (ICC '07), June 2007.

[11] Alexandru Coman, Mario A. Nascimento and JÄorg Sander. A Framework for Spatio-Temporal Query Processing Over Wireless Sensor Networks. 1st international workshop on Data management for sensor networks in conjunction with VLDB 2004. August 30-30. 2004.

[12] A. Gupta, M. Pal, R. Ravi and A. Sinha, Boosted sampling: approximation algorithms for stochastic optimization.34th Annual ACM Symposium on Theory of Computing (STOC'04).2004.pp 417–426.

[13] D. Culler, D. Estrin and M. Srivastava. Overview of Sensor Networks. IEEE Computer.Special Issue in Sensor Networks. Aug. 2004.

[14] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann and F. Silva. Directed Diffusion for Wireless Sensor Networking. IEEE/ACM Feb. 2003.

[15] D. A. Coffin, D. J. Van Hook, S. M. McGarry and S.R.Kolek. Declarative ad hoc sensor networking. SPIE Integrated Command Environments Conf.2000.

[16] Y.-B Ko,N. H. Vaidya. Location-aided routing (LAR) in mobile ad hoc networks. Mobicom'98.Oct 1998.

[17] Sameer Tilak,Nael B. Abu-Ghazaleh and Wendi Heinzelman. A Taxonomy of Wireless Micro-Sensor Network Models. Mobile Computing and Communications Review. Volume 6. Number 2.

[18] B. Karp, H. Kung. GPSR: Greedy perimeter stateless routing for wireless networks.6th Annual International Conference on Mobile Computing and Networking (MOBICOM'00).2000.pp 243–254.

[19] Y. Zhu, R. Vedantham, S. J. Park and R. Sivakumar. A Scalable Correlation Aware Aggregation Strategy for Wireless Sensor Networks. IEEE WICON 2005.Jul 2005.

[20] S. Pattem, B. Krishnamachari and R. Govindan. The impact of spatial correlation on routing with compression in wireless sensor networks. Third International Symposium on Information Processing in Sensor Networks (IPSN'04).ACM Press.2004. pages 28–35.

[21] Pascal von Rickenbach, Roger Wattenhofer. Gathering Correlated Data in Sensor Networks. DIALMPOMC'04. ACM. October 1.2004.

[22] Y. Xu, J. Heidemann and D. Estrin. Geography-informed energy conservation for ad hoc routing.7th Annual International Conference on Mobile Computing and Networking (MobiCom '01). ACM Press, 2001.pages 70–84..

[23] S. Yoon,C. Shahabi. Exploiting spatial correlation towards an energy efficient clustered aggregation technique (CAG).IEEE International Conference on Communications (ICC 05).pages 82–98, 2005.

[24] H. Gupta,V. Navda,S. Das and V. Chowdhary. Efficient gathering of correlated data in sensor networks. 6th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc '05).ACM Press, 2005. pages 402–413.

[25] J. Chou, D. Petrovic and K. Ramachandran. A distributed and adaptive signal processing approach to reducing energy consumption in sensor networks. 22nd Joint Conference of the IEEE Computer and Communications Societies (INFOCOM). volume 2.2003.pages 1054–1062,.

[26] R. Cristescu, B. Beferull-Lozano and M. Vetterli. On network correlated data gathering. 23rd Joint Conference of the IEEE Computer and Communications Societies (INFOCOM).2004. pages 2571–2582,.

[27] H. Takahashi, A. Matsuyama. An approximate solution for the Steiner tree problem in graphs. Math. Japonica 6 (24) (1980) 573–577.

[28] Linnyer Beatrys Ruiz, Isabela G. Siqueira and Leonardo B. e Oliveira. Fault Management in Event Driven Wireless Sensor Networks.MSWiM'04.October 4-6.ACM.2004

[29] Jia-Liang Lu and Fabrice Valois. On the Data Dissemination in WSNs. Third IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob 2007). 2007

[30] Elena Fasolo, Michele Rossi, J¨org Widmer and Michele Zorzi. In-network aggregation techniques for wireless sensor networks: a survey.Wireless Communications.IEEE.2007.

Table 1: Summary of Spatial Query Processing Methods

Method	**Type of Query Emphasized**	**Query Originator**	**Emphasis on Routing(R) or Data Aggregation(DA) or Both**	**Indexing Model**	**Is the index tree is adaptive ?**
SPIX [1]	-	BS	R	MBA	Yes
STI[2]	ATM	BS	R	MBA	No
ADI[8]	LD	Any Node	R	Quad tree	Yes
PT[4]	NN	Any Node	R	R-tree	No
SWIP[9]	Range queries	Any Node	Both	-	-
DQT[10]	Range queries	Any Node	R	Quad tree	No
ESPQ[7]	Spatial Pattern queries	BS	Discusses the evaluation of the query	-	-
SAP[5]	Spatial Aggregation queries	-	DA	-	-
PDT[6]	Selective Aggregate queries	-	DA	-	-
STWIN [11]	HSTQ	Any Node	Both	-	-