**An Analysis of Virtually Structured Energy Efficient Data Aggregation Protocols in Wireless Sensor Networks**

**Abstract:** Wireless Sensor Network (WSN) is a collection geographically distributed self-governing sensors. It consists of sensor nodes constrained by its limited processing power and battery life. The major challenge in WSN is of developing an energy efficient approach for its long-term operations. Data aggregation is one of the appreciated and accepted approach which efficiently reduces energy consumption in the network. There is high demand for research and development of efficient mechanism to adapt data aggregation with the routing approaches of WSNs. Many arrgegation approachs have been proposed which claims to be efficient in their respective application scenario. This paper provides an exhaustive analysis of different structureless data aggregation algorithms in WSNs. Algorithms were analyzed for their networking features and performance criterias to get an hypothetical view of recent research work done in the specified area. This work helps in understanding different approaches and involved features which has there impacts on WSNs performance. Individual algorithms were traced with the environment for which it was developed.

**Keywords:** Virtual Structure, multi-sink.

**1. INTRODUCTION**

In recent days, WSNs have become an integral part of our daily lives because of its varied applications ranging from healthcare to military surveillance[1]. WSN contains sensors which can act as data generating and forwarding nodes. A sensor node consists of a sensing equipment along with a transceiver to transfer the sensed data. These sensor nodes are small, with limited battery and computational capacity. The sensor nodes are generally deployed in difficult-to-access locations, hence, it requires wireless transfer of data to the sink[2].

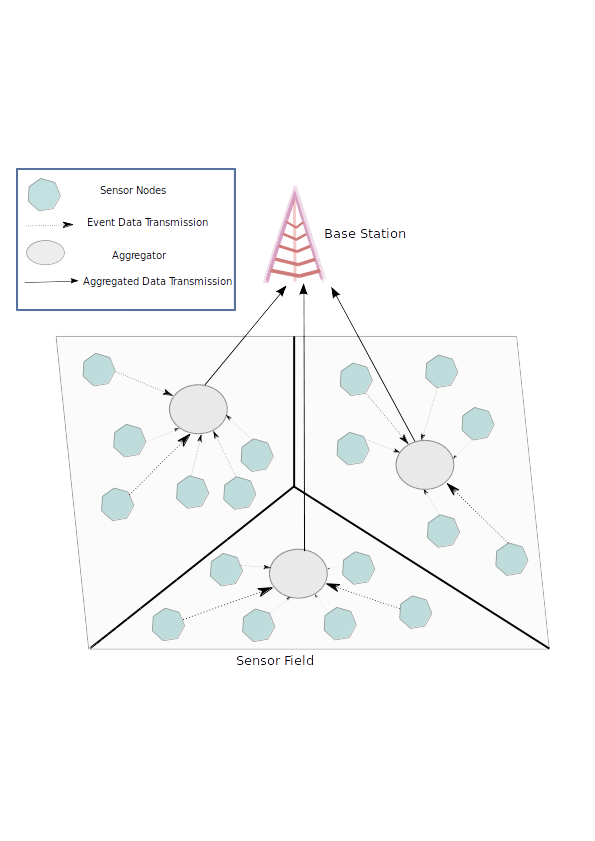
WSNs can be applied in various areas of military surveillance[3,4] , disaster relief, health care[5] and dangerous environment exploration. Its applications can be broadly classified into two categories of monitoring and tracking. Environmental and industrial environment monitoring are some of the monitoring based applications of WSNs[6]. Tracking applications include tracking movements of animals, human and vehicles like military aircraft.

With development in microelectromechanical systems(MEMS) technology, the design and expansion of battery powered multifunctional nodes is possible. The sophisticated and efficient communication protocols have made us realize the real potential of sensor networks. Generally, a WSN consists of a large number of sensor nodes, which are distributed randomly or in a predefined structure. A nodes in a WSN is battery powered device with limited processing and memory capabilities. These constraints are the key issues in establishing an efficient WSN with the longer lifespan.

Current state-of-the-art energy efficient routing protocols provide a quick fix for designing and developing WSNs in different environments[7]. With these existing protocols, nodes can be deployed in a structured pattern or in a self-organizing manner. These self-organizing protocols are also referred to as virtually structured protocols for sensor networks. Table 1 lists out some of the routing protocols used for self-organizing sensor networks.

As sensor is constrained by limited power, energy utilization is a critical concern in WSNs. Due to these reasons, considerable significance has been given for research in the field of minimizing energy usage in WSNs. Efficient energy usage plays an important role in increasing network lifetime. The energy depletion of a particular node significantly affects the functioning of the entire network, hence energy should be economically used to increase network life span.

Increase in network lifespan can be achieved through efficient and conscientious routing policies[10,11], smart sensor placement methodologies[8,9] and protocols that provide lofty data aggregation ratios[12]. Reduction in number of transmissions assists in increasing sensor lifetime and lowered bandwidth usage. It is possible by aggregating data from different sensor nodes and then forwarding it to sink as shown in Fig 1. Data aggregation assists in collecting and fusing only useful information from the multiple sensors. These fused data is then efficiently routed to the base station. Increased aggregation ratios play an integral part in energy conservation.

Fig 1 Data Aggregation in Wireless Sensor Networks

Two necessary conditions for efficient data aggregation ratios are spatial and temporal convergence. Spatial convergence focuses on establishing efficient routes from discrete sensor nodes to the sink whereas temporal convergence is responsible for efficient timing and control. Data aggregation protocols can be broadly classified into two categories (A) Structured (B) Structureless.

**A. Structured data aggregation :**

Clustering and tree-based are the two approaches used in structured data aggregation. In cluster based data aggregation, the nodes of WSN is organized into different clusters based on performance and geographical features. Each cluster has a cluster head who is responsible to aggregate data from remaining nodes in the same cluster. It then forwards the aggregated data to the base station. Some of the proposed algorithms in this category are LEACH[13], DACP[14] and HEED[15]. Compared to LEACH, DACP has an additional feature of data prediction which increases data aggregation ratios. HEED provides a solution to avoid two sensor nodes in the transmission range to become cluster head at the same time by considering its relative position[15].

Tree based approach works by constructing a minimum spanning tree structure that connects all the nodes and then transferring the aggregated data upstream until it reaches the sink. Mingxin Yang proposed a tree based approach by constructing energy efficient tree which reduced energy consumption by constructing minimum spanning tree and applying information entropy to find relationship among the sensed data[16]. Another tree-based data aggregation protocol was proposed by Fen Zhou et al., which dealt with maximizing the lifetime of data-aggregation tree. They adapted three different aggregation modes: full aggregation, non-aggregation and hybrid partial aggregation. They also, proposed a solution to find optimal data gathering tree based on Mixed Integer linear Programming(MIP)[17].

**B. Structureless data aggregation protocols:**

A major drawback in structured data aggregation protocols is the cost involved in creation and maintenance network structure. Structureless data aggregation is an approach that overcomes these overheads. Each node sends an anycast RTS along with event data to figure out next hop to the base station. Next hop is selected based on the similarity of the event data or closeness to the sink. The next hop sends back a CTS packet which begins the. Some protocols also involve a randomized waiting scheme at each node to reduce the number of transmissions which reduces power consumption.

Further, we present a review on some of the virtually structured data aggregation protocols with their advantages and overheads. Furthermore, we analyse each of them based on different performance measures. Finally, the paper is concluded.

**2. LITERATURE SURVEY**

Data aggregation is an effective approach to save a certain amount of resources in a wireless sensor network. This theory helps to reduce the number of transmissions which in turn saves energy. Data-Aware Anycast and Randomized Waiting(DAA+RW) proposed by Kai-Wei Fan et al., showed an efficient data aggregation protocol without explicit maintenance of the structure[18]. The aggregation protocol developed by them achieved the advantages of early aggregation as well as fault tolerance. To achieve greater aggregation levels , they used the concept of “gather before transmit” where a node closer to the sink is made to wait for a longer time before forwarding its data. This scheme also covered both spatial and temporal data aggregation which led to higher aggregation ratios. The data gathering process begins by each node sending an anycast RTS(Request To Send) to determine the next hop to the sink. The node that receives this RTS serves as a candidate for the next hop. At this stage, it is assumed that all the nodes are aware of the geographic location of the one-hop neighbours as well as the sink. To achieve higher aggregation ratios , the nodes which had the same data to forward or its closeness to the sink was given a higher priority to reply with a CTS(Clear To Send). The RW(Randomized Waiting) scheme helped them to reduce the number of transmissions. All the nodes that had a data to report had a random waiting time to start their transmission. DAA follows an MAC layer anycasting, where a node with one hop distance from the sending node is selected. A major advantage of DAA+RW is that it does not require reconstruction of network structure whereas in the case of structured approaches a node failure may involve the overhead of network reconstruction.

A grid-based routing protocol was proposed by Nen-Chung Wang et al., termed as Energy-Aware Data Aggregation (EADA)[19]. It proved that EADA algorithm efficiently reduces energy consumption as well as data collision in a sensor network. In the first phase, an M x N grid is formed whose cells are identified using XY-coordinates. The next phase selects a gateway node. A gateway node is selected on the basis of the remaining residual energy. Whenever the gateway node detects an event, it informs to all the gateway nodes in the grid. Now, the mobile sink sends a request packet to the gateway node of the cell of the same grid where the sink is located. This cell's gateway node sends a query message to collect data from the interested gateway node where an event has occurred.

Another structureless aggregation protocol called RDAG(Real-time Data Aggregation) was proposed by Mohammad Hossein Yeganeh et al., which used Real-time Data-aware Routing and a Judiciously Waiting policies[20]. This protocol made assumptions that a node is already aware of its own as well as sinks geographic location. A FIFO(First In First Out) queue was used to hold the packets at the aggregator nodes. For the purpose of Real-Time Data Aware routing policy, they used packet’s TTL(Time To Live) field as well as the Euclidean distance between a node and the sink. This protocol also measures EHD(Estimated Hop Delay) for efficient data aggregation. It makes dynamic forwarding decision by sending only to those nodes which have similar kind of packets waiting in the queue. A Judicial Waiting Policy was proposed to achieve temporal convergence. This policy was applied to all the packets in the network by allocating each of them with a slack time in proportion to the remaining hop count to sink node. Thus, their waiting time policy not only improved data aggregation but also helped during the periods of high contention without any synchronization within the nodes.

Dr.B.Vinayaga Sundaram et al., proposed an efficient structureless approach for data aggregation based on inverse-square law and survival analysis[21]. They were successful reducing the number of transmissions along with an increase in packet delivery ratio. They found a problem with DAA[18] that it uses CTS and RTS packet transmission for every data transmission and it can lead to an increase in load which in turn reduces sensor battery life. They derived a two-step framework for data aggregation and forwarding. In the first step, they used Inverse-Square law and Survival analysis to detect their nearby neighbour nodes. Along with these two laws, they used Kaplan Meier Estimator(KPE) which denotes the probability of a living node to survive. For Delay calculation, they proposed an Enhanced Random Delay where maximum delay value depends on the distance between a node and the sink. Further, they increased network efficiency by providing a Fault Tolerance Mechanism (FTM) in which each node sends a beacon signal and starts the timer. Now , if the node does not respond in the required time , it is considered to be failed. This failure message is also sent to other nearby nodes.

Chih-Min Chao et al. came up with a Structure-Free and Energy-Balanced(SFEB) data aggregation protocol[22]. Their main aim was to reduce energy consumption by cutting down the number of transmissions. They assumed that SFEB works in a multihop network where each node is aware of their location as well as the sink. This protocol consists of two phases ,in the first phase, when an event is generated the aggregators are dynamically chosen to collect data from other nodes and in the second phase the aggregators forward the collected data to the sink. To accomplish the task in phase one entire network is divided in virtual parallelograms. After successful division , an aggregator is selected based on their nearness to the parallelogram corners , this aggregator is known as a primary aggregator(PA). Also ,within the range of sensor nodes, there is a secondary aggregator(SA). They achieve early aggregation by selecting primary and secondary aggregator and then forward it to the next node which in turn pass the data to a sink. In this protocol, the sensors are placed in a hexagonal pattern which reduces the number of aggregators further reducing the number of transmissions. The phase two deals with forwarding the collected data back to sink. This is attained by making the rear aggregators(RA) forward their data prior to the nodes which are closer to the sink.

Yung-Kuei Chiang et al. proposed a Cycle Based Data Aggregation Scheme (CBDAS) for grid-based wireless sensor networks[23]. This scheme consists of four primary stages, they are grid construction , cell head selection , cycle formation, and data transmission. In the grid construction stage , the entire sensor field is partitioned into MxN cells , where each cell has a cell ID associated with it. Each cell in CBDAS consists of a head node (aggregator) and other ordinary nodes. In the next stage, a cell head is selected based on the residual energy of each node. The next stage deals with the construction of a cyclic chain to send data to uplink as well as downlink node. This links up all the cell heads which result in an hamiltonian cycle. The last stage deals with data aggregation and transmission to the base station. Here the base station selects a cycle head node from the set of head nodes based on the leftover residual energy. The base station sends two packets to the cycle leader, which is recursively sent to the neighbouring nodes. Finally, the head nodes transfer data to the cycle leader which forwards it to the base station. The major disadvantage of this scheme is an overhead of maintaining a table for each sensor node to store its geographic location , cell ID, and current cell head. Also , extra packets need to pass to form a cycle of head nodes which serves as the drawback to this scheme.

A. V. Sutagundar came up with Wheel based Event Triggered data aggregation and routing (WETdar) scheme for wireless sensor networks [24]. This scheme consists of 7 phases. In the first phase, a dynamic wheel structure is constructed and a path is discovered to the sink node. The second phase deals with the identification of Secondary Aggregators(SA) along the spokes considering the factors like Euclidean distance , residual energy, spoke angle and connectivity. In the third phase, SA on the spoke aggregates the data from the neighbouring nodes. The fourth phase deals with invoking event detection and classification process. The fifth phase the boundary nodes aggregates data from the SAs along the spoke of the wheel. The sixth phase deals with maintaining the wheel structure. Finally, the aggregated data is forwarded to the sink along the selected path.

Abdul Waheed Khan et al. Designed a Virtual Grid based Dynamic Routes Adjustment Scheme(VGDRA) for wireless sensor networks[25]. Their main aim was to minimize route reconstruction cost to optimally route data to the sink new locations. Instead of using the grid construction approaches used in CBDAS[23] in DPBDAS[26] they used heuristics used in LEACH[29] , TEEN[30] , APTEEN[31]. In the next stage , for each cell, a node closest to the cell midpoint is selected as a cell head. This approach of head selection consumes lesser energy as GPS locations are directly used for its selection instead of node voting. After the node head election, notifications are sent to neighbouring nodes in the same cell as well as in the nearby nodes of the neighbouring cell. This enables them to identify cell adjacencies which form a virtual backbone structure. The next stage is to set up optimal routes to the sink. The major advantage of this scheme is routing information is stored only in the head node of each cell. Each time the base station changes its location the routes to the sink are dynamically changed which results in the optimal routes.

Neng-Chung Wang et al. proposed a Dual-Path-Based Data Aggregation Scheme (DPBDAS) for the grid-based network[26]. This work is an enhancement to Cycle Based Data Aggregation Scheme proposed (CBDAS) proposed by the same authors[23]. The only difference between both the schemes is that DPBDAS selects two cell heads within a grid (cell head A and cell head B) and then connects cell heads A and cell heads B resulting in a path for aggregated data transfer to the sink, all other procedure being same as CBDAS.

Kaushik Ghosh et al. came up with a Fermat Point Based Energy Efficient Data aggregating routing protocol for multi-sink wireless sensor networks[27]. Their main aim was to come up with a distance vector protocol for data aggregation WSNs. The major benefit of this protocol is that it does not require to form any virtual structure like VGDRA [25] and DPBDAS[26]. In this protocol, after nodes have been deployed, a fermat point is selected by forming a logical triangle/polygon from the source node to multiple sinks. They used Minima algorithm to find a fermat point between source and sinks[32]. Now, the transmission begins by first transferring data from source to Fermat Node(FN) and then from FN to sinks. The major drawback of this algorithm is that if the FN is much far away from the source node, this will inversely affect the network lifetime.

Shubhra Jain et al. Developed a Vertical and Horizontal Segregation based data dissemination protocol[28]. They wanted to reduce the amount of energy consumed in informing the nodes about the current location of the base station. This protocol consists of five phases starting from backbone creation , neighbour discovery, tree construction and finally sharing sinks location and data dissemination. In the backbone creation stage, the square sensor field will be divided into four equal parts by a vertical and horizontal stripes. Again each of these four areas is divided into two equal parts by 45-degree angle from the center. All the nodes falling inside the horizontal and vertical strips are known as spine nodes. In the next stage, all the regular nodes will send packets to spine nodes to share their locations. The next stage deals with the decision of nodes region discovery which helps to identify the belongingness of regular nodes to a part of horizontal and vertical strips. Then the sink sends its location to the nearest node to itself. This information is then sent to all the spine nodes that form the backbone. The regular nodes then communicate with nearest spine node to get the sink location information and then forwards the data in a multi-hop manner.

The following table provides an overview of the protocols reviwed in this section.

| Protocol | Author | Virtual Network Structure |
| --- | --- | --- |
| Data-Aware Anycast and Randomized waiting (DAA+RW)[18] | Kai-Wei Fan, Sha Liu et al. | No virtual structure |
| Energy-Aware Data Aggregation (EADA)[19] | Nen-Chung Wang, Yung-Fa Huang et al. | Virtual Grid |
| Real-time Data Aggregation (RDAG)[20] | Mohammad Hossein Yeganeh,Hamed Yousefi et al. | No structure |
| Data aggregation using Inverse Square and Survival analysis(SA-ERD)[21] | Dr.B.Vinayaga Sundaram, Rajesh G et al. | No virtual structure |
| Structure Free and Energy Balanced (SFEB) Data Aggregation[22] | Chih-Min Chao and Tzu-Ying Hsiao | Virtual Parallelogram and sensor placement in hexagonal pattern |
| Cycle Based Data Aggregation Scheme(CBDAS)[23] | Yung-Kuei Chiang, Neng-Chung Wang et al. | Virtual Grid |
| Wheel based Event Triggered data aggregation and routing (WETdar)[24] | A. V. Sutagundar and S. S. Manvi | Virtual Wheel Structure |
| Virtual Grid based Dynamic Routes Adjustment Scheme(VGDRA)[25] | Abdul Waheed Khan, Abdul Hanan Abdullah et al. | Virtual Grid |
| Dual-Path Based Data Aggregation Scheme(DPBDAS)[26] | Neng-Chung Wang, Yung-Kuei Chiang et al. | Virtual Grid |
| Fermat Point Based Energy Efficient Data Aggregation routing protocol(KPS)[27] | Kaushik Ghosh, Pradip K. Das et al. | Polygon/Triangle |
| Vertical and Horizontal Segregation  Based Data Dissemination Protocol[28] | Shubhra Jain, Suraj Sharma et al. | Square division into Octants |

| Phase description | Types of packets |
| --- | --- |
| Sending an anycast request to determine next hop to sink | RTS (Request to send)  - Class A (Same AID closer to sink)  - Class B (Same AID farther away from sink)  - Class C (Different AID closer to sink)  CTS (Clear to send) |
| Selecting a node for the next hop from the candidates who have received RTS. |
| Randomized waiting |
| Grid construction | - Event registration packet  - Query message packet  - tree\_construct packets |
| Gateway node selection |
| Event registration |
| Data collection |
| Real-time data aware routing | -Holding packets using FIFO queue  -Time to Live (TTL)  - Estimated One hop Delay (EHD) |
| Judiciously waiting |
| Applying Inverse Square to identify distance | -Beacon signal |
| Survival Analysis to for Event density function |
| Kaplan Meier Estimator to calculate the lifetime probability |
| Dynamically choosing the aggregators | -PA request (Primary Aggregator)  -SA request (Secondary Aggregator)  -RTS  -CTS |
| Collecting and Forwarding it to sink node |
| Grid Construction | - Head\_electing  - Head\_confirming  - Path\_forming  - Path\_reply  - Token to Cycle Leader |
| Cell Head selection |
| Cycle Formation |
| Data transmission |
| Wheel Structure Construction | - Event triggering packet  - Spoke generation packet |
| Identification of Secondary Aggregators(SA) |
| Data aggregation by SA's |
| Event Detection and classification |
| Data aggregation from boundary nodes |
| Maintaining Wheel structure |
| Data dissemination to sink |
| Grid construction | - Sink location update packet  - Route update packet |
| Cell head selection |
| Identification of optimal route to sink |
| Data dissemination |
| Grid Construction | - Head\_electing  - Head\_confirming  - PathA\_forming  - PathA\_reply  - PathB\_forming  - PathB\_reply  - Token to Cycle Leader of A group  - Token to Cycle Leader of B group |
| Two cell head selection |
| Two cycle formation |
| Data transmission |
| Node Deployment | - neighbour\_becon control packet |
| Fermat point selection |
| Data transfer from source to FN |
| Aggregated data forwarding to sinks from FN |
| Backbone creation | -NBR\_MSG  -NBR\_REPLY  -T\_MSG  -HELLO\_MSG  -HELLO\_REPLY  -S\_MSG  -SLQ  -SLR |
| Neighbour discovery |
| Tree construction |
| Sink location information |
| Data dissemination |

| Aggregator node selection | Overheads | Assumptions |
| --- | --- | --- |
| Aggregation ID generated for each packet. Any node near to sink can serve as an aggregator | -Increase in load because of CTS and RTS transmission which can reduce battery performance  -Tree construction overhead | - Node are aware about the geographic location of its one hop neighbours.  - Interference range is twice as transmission range  - Nodes are time-synchronized |
| - Cell head selection based on residual energy | - Event registration overhead.  - Increased rerouting overhead | - Static sensor nodes  - Multihop communication  - Aware of its geographic location  - Location table of all the gateway nodes at the sink  -Bidirectional Channels |
| Dynamic forwarding of packets to the nodes with similar packets. Any node with similar packets can serve as aggregator | -Artificial Waiting policy  -Estimation of end to end delay  -Calculating the required velocity based on packet's TTL | - Nodes are aware about its own geographic location and of sink  - No resource limitation on the sink  - Sensors are battery operated with limited capacity  -Non-adjacent nodes communicate hop by hop. |
| Dynamic forwarding of packets to the nodes with similar packets. Any node with greater survival probability can serve as agregator | Probability Value Calculation for each node | - Failed node is a data forwarding node.  - Failed node is a sensor node as well as aggregator node |
| -Primary Aggregator(PA) is selected based in the nearness of the parallelogram corners  -Secondary Aggregator is selected within the range of PA | -High Computation overhead  -Longer Waiting in phase one may be in vain if aggregation levels are low | -Each node is aware about its own as well as sinks geographic location.  -SFEB operates in a multihop network |
| -Cell head based on remaining residual energy  -Cycle leader selection based on distance from Base station | - Passing of extra packets to form a hamiltonian cycle  - Maintaining the cycle | -Static sensor nodes  -Base station located far away  -Location aware  -Nodes are aware about  residual energy.  -Bidirectional channels |
| -Spoke Aggregators(SA) are along the spokes of the wheel  -Selected on the basis of residual energy and minimum deviation from sensor nodes | - Dynamic wheel structure maintenance  -Calculation of euclidean distance and spoke angle | -Multihop communication with sink on event detection  -Static nodes  -Each sensor node has security features to perform inter-agent communication |
| - Cell head selection based on nearness to the midpoint of the cell.  - Re-election considers midpoint closeness as well as residual energy | - Route updation at cell heads because of sink movements  - Optimal route calculation based on sink location | - Random node deployment  - Static sensor nodes and moving sink  - Mobile sink does not have any resource constraints |
| -Cell head A and cell head B selection based on residual energy  -Cycle leader A and Cycle leader B selection based in distance from base station | -Passage of extra packets to form cycles of head A and head B  -Token passing from base station for collection of aggregated data | -Static sensor nodes  -Base station located far away  -Location aware  -Nodes are aware about  residual energy.  -Bidirectional channels |
| - It's a Fermat Node located on the fermat point calculated by Minima algorithm[14] | -Fermat point calculation  -Multihop transfer to far away fermat nodes  -Confirming the residual energy above threshold of a fermat node after each transmission | - Fermat point are near to the source node  - Multiple sink are located at the edges of the network  - Source node are aware about the position of fermat node |
| - Spine nodes in the stripes forms the backbone. | - Octants creation by dividing square into 8 parts  - Large number of communication packets transferred between spine and regular nodes | -Location aware sensor nodes  -Static nodes  -Square shaped sensor field |

| Advantages |
| --- |
| -No overhead of constructing the structure.  -Reduced number of transmissions. |
| - Reduced overhead of discovering new paths to sink  -Increased network lifetime |
| -No explicit maintenance of structure.  -Implementing artificial delays resulting in increase in temporal convergence |
| -Node failures are considered  -High packet delivery ratio because of node failure avoidance mechanisms |
| -Increase in event size has a little effect on aggregation ratio |
| -Easy cyclic chain formation  -Extended node lifetime |
| -Two level aggregation increases aggregation ratio |
| - Optimal routing path is determined  - Reduced reconstruction cost |
| -Evenly distributed energy depletion  -Dual path transmission |
| - No structure maintenance overhead  - Reduced disadvantage of greedy packet forwarding algorithm |
| - Reduced number of transmissions to send sinks location |

**3. ANALYSIS**

The efficiency of a data aggregation protocol depends on many factors. Some of the performance-based aspects are structure maintenance cost, sinks and nodes mobility, routing table maintenance and allowance of sleeping nodes.

[A] Node Mobility:

A mobile node causes the frequent update of location tables, to attain maximum route efficiency. Node movement thus, consumes energy to inform neighbouring nodes about their current location by passing numerous control packets.

[B] Transmission:

Two data transmission approaches are generally adopted in sensor networks. It can be a single hop transmission directly to the sink or a multihop transfer where multiple sensor nodes are involved in forwarding the data to the sink.

[C] Structure Maintenance Cost:

Movement of nodes or a node failure can cause major changes in the network structure. This causes the reconstruction of network structure which consumes considerable amount battery power. A data aggregation protocol must consider these factors to attain appreciable aggregation ratios.

[D] Location Awareness:

Nodes in a wireless sensor network are randomly deployed in the sensor field. Each node needs to be aware of its GPS location. Later, the location details are used for distance calculations to form a virtual structure.

[E] Base Station Mobility:

Like nodes, even base stations mobility affects the performance of a data aggregation protocol. The movement of base station causes a change in the network structure at some or all the nodes. This incurs a considerable usage of node energy as they transfer control packets to retrieve the sinks current location.

[F] Node Failure:

Factors such as energy depletion, link breakage, physical damage and malicious attacks can cause a node to fail. A node failure might result in the breakdown of a part or an entire network. To overcome this, parts network where node failure occurred needs to be restructured which adds to the maintenance cost.

[G] Routing Table:

Nodes in a wireless sensor network maintain a routing table to identify nodes to forward their data. The amount of routing information stored at each node should be limited because of the memory restrictions in sensor nodes.

[H] Energy Tracking:

Rapid node energy depletion is one of the crucial factor affecting the lifetime of a sensor network. As compared to other nodes, the energy of the nodes involved in aggregation and forwarding gets depleted at a faster rate as it involves extra processing. As a result, energy level tracking becomes a compelling feature of data aggregation protocols.

[I] Sleeping Nodes:

In some sensor networks, event data is required after certain time intervals. A predefined downtime can be set for the nodes in such network. This approach aids in increasing energy conservation as sleeping nodes do not consume any battery power.

[J] Interference:

A sensor network consists of nodes with limited transmission range and bandwidth. Random node deployment increases the chances of intersection between transmission range of different nodes. This interference may cause packet collision that can result in invalid event data or complete loss of packets.

[K] Multisink Support:

There are networks wherein event data needs to be sent to multiple base stations. Here, the geographical location of multiple sinks needs to be maintained at aggregator nodes for efficient forwarding.

Based on the above disscused performance aspects, a study was conducted to analyse various data aggregation protocols(table 2).

| **Protocol** | Node Movement | Transmission | Location Aware |
| --- | --- | --- | --- |
| DAA+RW[18] | Dynamic | Multihop | Yes |
| EADA[19] | Static | Multihop | Yes |
| RDAG[20] | - | Multihop | Yes |
| SA-ERD[21] | - | Multihop | No |
| SFEB[22] | Dynamic | Multihop | Yes |
| CBDAS[23] | Static | Multihop | Yes |
| WETdar[24] | Static | Multihop | Yes |
| VGDRA[25] | Static | Multihop | Yes |
| DPBDAS[26] | Static | Multihop | Yes |
| KPS[27] | Static | Multihop | Yes |
| [28] | Static | Multihop | Yes |

| **Protocol** | Mobile Station Movement | Node Failure  Consideration | Routing Table Maintenance |
| --- | --- | --- | --- |
| DAA+RW[18] | No | Yes | At each node |
| EADA[19] | Yes | Yes | Gateways |
| RDAG[20] | - | Yes | At each node |
| SA-ERD[21] | - | Yes | At each node |
| SFEB[22] | - | Yes | PA and SA |
| CBDAS[23] | - | Yes | Cell heads |
| WETdar[24] | No | Yes | Spoke Aggregators |
| VGDRA[25] | Mobile | Yes | Cell heads |
| DPBDAS[26] | Static | Yes | Cell heads |
| KPS[27] | Static | Yes | Fermat Point |
| [28] | Mobile | - | Spine nodes |

| **Protocol** | Sleeping Nodes | Interference  consideration | Collision Avoidance | Multi sink target |
| --- | --- | --- | --- | --- |
| DAA+RW[18] | No | Yes | Yes | No |
| EADA[19] | Yes | - | Yes | No |
| RDAG[20] | - | Yes | Yes | No |
| SA-ERD[21] | - | Yes |  | - |
| SFEB[22] | - | Yes | Yes | No |
| CBDAS[23] | Yes | - | - | No |
| WETdar[24] | Yes | Yes | - | No |
| VGDRA[25] | - | - | - | - |
| DPBDAS[26] | Yes | - | - | - |
| KPS[27] | - | - | - | Yes |
| [28] | - | - | - | No |

**4. CONCLUSION**

Energy efficiency is one of the crucial aspects in the development of data aggregation protocols for constrained wireless networks. We provided an overview of some of the data aggregation protocols which are based on virtual structure formation. The overview briefly discussed the various features of data aggregation protocols along with their merits and overheads. Later, we presented an analysis of the surveyed protocols considering various performance measures.

**REFERENCES**

[1] Ian F. Akyildiz, Mehmet Can Vuran, “Wireless Sensor Networks”, Wiley 2010.

[2] Jennifer Yick, Biswanath Mukherjee, Dipak Ghosal, “Wireless sensor network survey”,Computer Networks 52, Elsevier B.V, 2008.

[3]G. Simon, M. Maroti, A. Ledeczi, G. Balogh, B. Kusy, A. Nadas, G. Pap, J. Sallai, K. Frampton, “Sensor network-based countersniper system”, Proceedings of the Second International Conference on Embedded Networked Sensor Systems (Sensys), Baltimore, MD, 2004.

[4] J. Yick, B. Mukherjee, D. Ghosal, “Analysis of a Prediction-based Mobility Adaptive Tracking Algorithm”, Proceedings of the IEEE Second International Conference on Broadband Networks (BROADNETS), Boston, 2005.

[5] Hande Alemdar, Cem Ersoy, “Wireless sensor networks for healthcare: A survey”, Computer Networks 54 , Elsevier B.V., 2010.

[6] Kay Soon Low, Win Nu Nu Win, Meng Joo Er,“Wireless Sensor Networks for Industrial Environments”, Proceedings of the 2005 International Conference on Computational Intelligence for Modelling, Control and Automation, and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC’05).

[7] Gaddafi Abdul-Salaam, Abdul Hanan Abdullah, Mohammad Hossein Anisi, Abdullah Gani, Abdulhameed Alelaiwi, “A comparative analysis of energy conservation approaches in hybrid wireless sensor networks data collection protocols”, Telecommun Syst, Springer, 2015.

[8] Mustapha Reda Senouci, Mohamed El Yazin Boudaren, Mohamed Abdelkarim Senouci, Abdelhamid Mellouk, “A smart methodology for Deterministic Deployment of Wireless Sensor Networks”, [Smart Communications in Network Technologies (SaCoNeT), 2014 International Conference](http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=6857452), IEEE,2014.

[9] Carlos E. Otero, Ivica Kostanic, Luis D. Otero, “A Decision-Making Methodology for stochastic Deployment of Wireless Sensor Networks”, UKSim 2009: 11th International Conference on Computer Modelling and Simulation, IEEE, 2009.

[10] Yongrui Chen, Li Gao, Yulong Xing, Weidong Yi, “Cross-layer Design for Energy-Efficient Reliable Routing in Wireless Sensor Networks”, 2015 11th International Conference on Mobile Ad-hoc and Sensor Networks, IEEE, 2015.

[11] Ioannis Ch. Paschalidis, Wei Lai, David Starobinski, “Asymptotically Optimal Transmission Policies for Large-Scale Low-Power Wireless Sensor Networks”, IEEE/ACM TRANSACTIONS ON NETWORKING, VOL. 15, NO. 1, FEBRUARY 2007.

[12] Feng Wang, Jiangchuan Liu, “Networked Wireless Sensor Data Collection: Issues, Challenges, and Approaches”, IEEE COMMUNICATIONS SURVEYS & TUTORIALS, VOL. 13, NO. 4, FOURTH QUARTER 2011.

[13] Wendi B. Heinzelman, Anantha P. Chandrakasan,Hari Balakrishnan, “An Application-Specific Protocol Architecture for Wireless Microsensor Networks”, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, VOL. 1, NO. 4, OCTOBER 2002.

[14] Lingjun Meng, Huazhong Zhang, Yun Zou, “A Data Aggregation Transfer Protocol based on Clustering and Data Prediction in Wireless Sensor Networks”, [Wireless Communications, Networking and Mobile Computing (WiCOM), 2011 7th International Conference](http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=6036127) IEEE, 2011.

[15] Ossama Younis, Sonia Fahmy, “HEED: A Hybrid, Energy-Efficient, Distributed Clustering Approach for Ad Hoc Sensor Networks”, IEEE TRANSACTIONS ON MOBILE COMPUTING, VOL. 3, NO. 4, OCTOBER-DECEMBER 2004.

[16] Mingxin Yang, “Constructing Energy Efficient Data Aggregation Tree Based on Information Entropy in Wireless Sensor Networks”, 2015 IEEE Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), IEEE, 2015.

[17] Fen Zhou, Zhenzhong Chen, Song Guo, Jie Li, “Maximizing Lifetime of Data-Gathering Trees with Different Aggregation Modes in WSNs”, 2015 IEEE Global Communications Conference (GLOBECOM), IEEE, 2015.

[18] Kai-Wei Fan, Sha Liu, and Prasun Sinha , “Structure-Free Data Aggregation in Sensor Networks”, IEEE Transactions on Mobile Computing,VOL. 6,NO. 8, AUGUST 2007, IEEE.

[19] Nen-Chung Wang, Yung-Fa Huang, Jong-Shin Chen, Po-Chi Yeh , ”Energy-Aware Data Aggregation for Grid-Based Wireless Sensor Networks with a Mobile Sink”, Wireless Pers Commun (2007) 43:1539–1551, Springer , 2007.

[20] Mohammad Hossein Yeganeh,Hamed Yousefi,Naser Alinaghipour,Ali Movaghar ,“RDAG: A Structure-free Real-time Data Aggregation Protocol for Wireless Sensor Networks”, 17th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications, IEEE, 2011.

[21] Dr.B.Vinayaga Sundaram, Rajesh G, Khaja Muhaiyadeen A, Hari Narayanan R, Shelton Paul Infant C,Sahiti G, Malathi R, Mary Priyanga S ,”An Efficient Approach for Data Aggregation Routing using Survival Analysis in Wireless Sensor Networks” ,International Journal of Wireless & Mobile Networks (IJWMN) Vol. 3, No. 2, April 2011

[22] Chih-Min Chao , Tzu-Ying Hsiao , “Design of structure-free and energy-balanced data aggregation in wireless sensor networks”, Journal of Network and Computer Applications 37 (2014) 229–239,Elsevier Ltd 2013.

[23] Yung-Kuei Chiang, Neng-Chung Wang and Chih-Hung Hsieh, “A Cycle- Based Data Aggregation Scheme for Grid-Based Wireless Sensor Networks”, [Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2013 Seventh International Conference,](http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=6603432) Sensors, [201](http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=6603432)4.

[24] A. V. Sutagundar, S. S. Manvi , “Wheel Based Event Triggered Data Aggregation and Routing in Wireless Sensor Networks: Agent Based Approach” ,Wireless Pers Commun (2013) 71:491–517, Springer, 2013

[25] Abdul Waheed Khan, Abdul Hanan Abdullah, Mohammad Abdur Razzaque, Javed Iqbal Bangash, “VGDRA: A Virtual Grid based Dynamic Routes Adjustment Scheme for Mobile Sink based Wireless Sensor Networks”, IEEE Sensors Journal, IEEE, 2013.

[26] Neng-Chung Wang, Yung-Kuei Chiang, Young-Long Chen, and Chih-Hung Hsieh, “A Dual-Path-Based Data Aggregation Scheme for Grid-Based Wireless Sensor Networks”, 2014 Eighth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing, IEEE, 2014.

[27] Kaushik Ghosh, Pradip K. Das and Sarmistha Neogy, “KPS: A Fermat Point Based Energy Efficient Data Aggregating Routing Protocol for Multi-sink Wireless Sensor Networks”,Advanced Computing and Systems for Security,Springer India 2016.

[28] Shubhra Jain, Suraj Sharma and Neeraj Bagga, “A Vertical and Horizontal Segregation Based Data Dissemination Protocol”, Emerging Research in Computing, Information Communication and Applications, Springer India 2016

[29] W. B. Heinzelman, A. P. Chandrakasan, S. Member, and H. Balakrishnan, “An Application-Specific Protocol Architecture for Wireless Microsensor Networks,” IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, VOL. 1, NO. 4, IEEE, OCTOBER 2002.

[30] A. Manjeshwar and D. P. Agrawal, “TEEN: a routing protocol for enhanced efficiency in wireless sensor networks”, Proceedings 15th International Parallel and Distributed Processing Symposium, IEEE, April 2000.

[31] A. Manjeshwar and D. P. Agrawal, “APTEEN: a hybrid protocol for efficient routing and comprehensive information retrieval in wireless sensor networks,” Proceedings 16th International Parallel and Distributed Processing Symposium, IEEE, April 2001.

[32] Ghosh, K., Roy, S., Das, P.K., “An alternative approach to find the fermat point of a polygonal geographic region for energy efficient geocast routing protocols: global minima scheme.” NETCOM’09. First International Conference on Networks and Communications, 2009, pp. 332–337, IEEE, 2009