

A Survey on Data Aggregation And Clustering Schemes in Underwater Sensor Networks

Rakesh Kumar¹ and NavdeepSingh²

¹Department of IT, M. M. University, Mullana, Ambala, India

²System Engineer, GMUADM, Infosys Ltd., Pune, India

Email: {¹raakeshdhiman, ²navi.rise}@gmail.com

Abstract

Energy consumption is one of the most challenging constraints of the design and implementation of the sensor network. Underwater sensor networking is the technology that enables the applications like environment monitoring, underwater exploration, seismic monitoring and other surveillance applications. In underwater sensor network, a sensor node senses the data and transmits it to the sink. Many routing algorithms have been proposed in order to make the network phase of UWSNs more efficient. In this report, we present a review and comparison of various data collection algorithms and clustering schemes, proposed recently in order to execute the demands of the ongoing researches. The main goal of data aggregation technique is to accumulate data in an energy efficient manner for a long-term network monitoring. The main purpose of this study is to present algorithms addressing issues like deployment and localization in UWSNs under different conditions.

Keywords: Data Aggregation, Clustering, Energy Analysis, Effective Routing, Underwater WSN, Architecture, Applications

IMAGES USED

- a. The basic principal for Data Aggregation [1].
- b. 2-D Underwater sensor network architecture [4].
- c. 3-D Underwater sensor network architecture [4].
- d. Mobile sensor network architecture [5].
- e. A high level view of Vector Based Forwarding [9].
- f. Topology Structure of clustered UWSNs in LEACH [13].
- g. An example scenario of LAA [16].
- h. Sphere Energy Depletion Model in REBAR [30].
- i. Example Network with DUCS [22].
- j. Topological Structure of CHSA.
- k. Underwater Network Topology for Multipath Virtual Sink architecture[30].

1. Introduction

In recent years, the growing interest in the wireless sensor network has resulted in thousands of peer-reviewed publications. We know very little about the Earth's water bodies. Almost 97% of the planet earth is water and surprisingly, not more than 10% of it has been discovered yet. An underwater network is generally made up of many autonomous and individual sensor nodes that perform data agglomeration operations as well as a store and forwarding operations to route the data that has been collected to a central node. The main challenges of deploying such a network are the computational power, the cost, the memory, the communication range and most of all the lifetime of any individual sensor in the UWSNs is limited, the number of sensor nodes that stop working due to power loss increases with lengthened deployment time, therefore the coverage area of wireless sensor network will shrink. Another crucial issue is the limited battery resources and it is a challenge for researchers to obtain long operating time without sacrificing system performance. Therefore, many top researchers have worked on proposing data aggregation techniques in order to develop energy efficient protocols for UWSN node functions. In UWSN data aggregation, the master node (aggregator node) aggregates the sensed data from surrounding nodes, process, and transmit the data to the sink. So, the major challenge of data aggregation in UWSNs is to minimize data redundancy while ensuring high data accuracy. Data aggregation is one of the major communication approaches in which multiple sources are sending data to the single sink as shown in Figure 1. We need to transfer data from multiple sources to the single sink. But as such kind of communication is performed individually between each source and the sink. There will be a heavy traffic or communication over the network that can result in energy and data loss. Because of this there is a requirement of data aggregation. Data aggregation has been examined as an essential technique for reducing energy consumption in wireless sensor network by minimizing superfluity from the data sensed by the sensor nodes. The process of data aggregation helps not only to enhance the accuracy of information that is obtained by the entire networks, but also reduces the redundant information and prolongs the network lifetime.

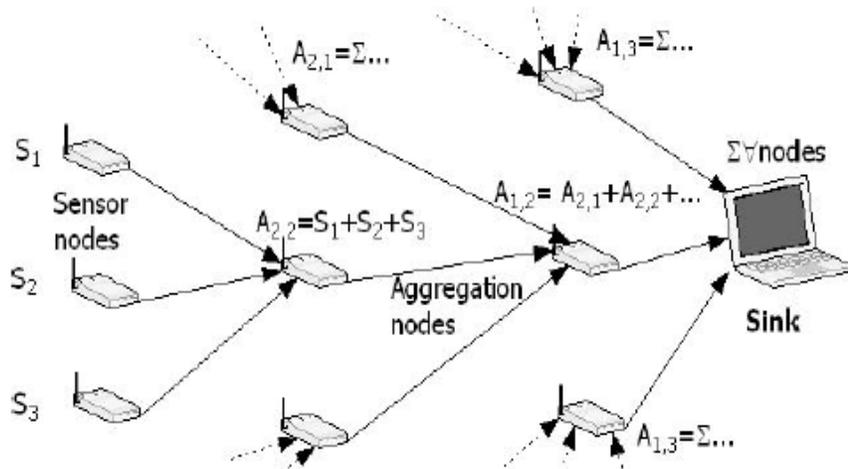


Figure 1. Basic Process of Data Aggregation [1]

The basic idea behind the aggregation is to avoid one to one communication to the sink node and build an aggregate path over the network. The main objective of data aggregation is

to reduce the network traffic and to save the energy loss over communication. The path selection in an aggregation depends on multiple factors such as type of network, network topology, etc. Two main security challenges in secure data aggregation are confidentiality and integrity of data. The main goal of data aggregation technique is to accumulate data in energy efficient manner for a long-term network monitoring. Data aggregation is one of the key process of aggregator in order to decrease the network consumption by eliminating the superfluous data as well as reducing the packet size being transmitted to the sink. It however complicates the already existing security challenges for underwater sensor networks and requires new security tailored.

In this present work, we present a review of various data aggregation algorithms, proposed recently in order to fulfil the requirements. The main purpose of this study is to address issues like deployment and localization in UWSNs under different conditions. The rest of the paper is organized as follows. Section 2 provides the features of UWSNs, followed by Section 3 providing the overview of the applications of underwater sensor networks. Different architectures are presented in Section 4. TWSNs and UWSNs are compared in Section 5. In Section 6, we present several underwater data aggregation and clustering schemes and highlight the related issues. Open research issues are discussed in Section 7. Section 8 concludes the paper.

2. Features of UWSNs

UWSNs are significantly different from any terrestrial sensor networks in terms of the following aspects.

There are many aspects in which underwater sensor network is different from the terrestrial sensor network. Some of the important aspects are discussed below. Later, both the UWSNs and TWSNs are compared comprehensively in Section V.

- *Dynamic* – In underwater sensor network, the environment is such that the ocean current can repeatedly transport the nodes randomly to different areas as the nodes can move with the ocean currents underwater. According to the statistics and observation results as in [26], it is assumed that the nodes may move at a speed of 4-6 kilometers per hour approximately. Hence, node mobility results in making the underwater sensor network highly dynamic.
- *Error Prone* – Underwater is an environment where many parameters are to be taken into concern before designing a network. Factors such as noise, dopplerspread, water temperature, node mobility, battery exhaustion, multipath fading, signal attenuation results in degradation of the sensor network. The underwater sensor network is much more sensitive than the terrestrial network. Thus, the underwater sensor network comes out to be much more prone to errors as compare to the terrestrial network. In UWSNs, the rate of failure of the sensor nodes is much more because of the harsh environment conditions underwater.
- *Three Dimensional* – Underwater Sensor Network is deployed randomly by keeping in concern the node localization resulting in a 3-dimensional architecture underwater, whereas in most of the terrestrial sensor networks, a 2-dimensional orchestra is formed.

3. Applications

The underwater sensor network finds its application in a number of areas, including water characteristics, counting and imaging animal life. We have discussed its different applications below.

Ocean sampling networks – Sensor networks can perform adaptive sampling. This includes strategies to command the AuVs to places where data will be most useful. This approach has been proposed in pioneer monitoring missions such as [2] in order to improve the ability to predict the characteristics of the oceanic environment.

Pollution monitoring – UWSNs can perform monitoring of contamination such as chemical, biological and nuclear contamination. Monitoring ocean currents and climate change detection also count in the applications.

Underwater explorations – UWSNs help in keeping track of undersea oil fields and also determining paths and routes for undersea cables, and assist in the exploration for minerals.

Disaster prediction and prevention – Underwater sensor network helps in predicting the disasters and their remote location with the help of seismic monitoring.

Navigation assistant – Sensor network finds their application in assisting navigation with locating dangerous rocks, submerged wrecks *etc.*

4. Architecture

Network topology is in general an essential factor in determining the energy consumption, the capacity and the reliability of a network. While deploying sensor network, it is important to form a sensor network which is highly reliable. In order to avoid failure of nodes, the topological deployment scheme should be carefully engineered. As the capacity is limited in sensor nodes, it is important to organize the network topology in such a way that the energy remains conserved and is used efficiently. Network topology is an open research issue that needs further investigation. Underwater sensor network architecture can be categorized according to the sensor's mobility.

Static UWSN is one in which sensors are assumed to be stationary or with negligible movement. This architecture is further categorized into two-dimensional or three-dimensional.

Mobile UWSN is one where sensors are assumed to be floating freely with the ocean currents.

4.1. Static UWSN

As discussed earlier static UWSN can be further classified into two categories: -

4.1.1. Two-dimensional Static Underwater Sensor Network.

In static 2-D underwater sensor network, each sensor is interconnected to one or more underwater sinks (gateway) sensor by using wireless acoustic links as shown in Figure 2.

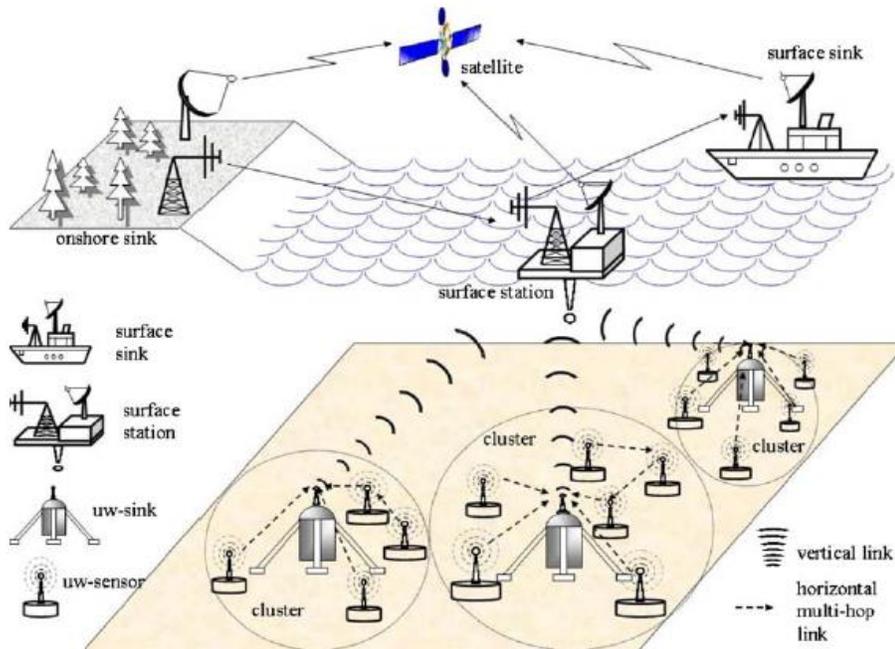


Figure 2. 2-D Underwater Sensor Network Architecture [4]

These gateways are network devices which are responsible for relaying data from the ocean bottom network to a surface station. In May 2006, in order to increase the robustness and energy efficiency, a two virtual dimensional, multi-tier topology was sunk together form a mesh proposed by W.K.G. Seah [3] where network and sensors transmit their data to the local sink via virtual sink within their own cluster. Underwater sensor network sinks are equipped with two acoustic transceivers namely a vertical and a horizontal transceiver. The horizontal transceiver is used by the sink to communicate with the sensor nodes and the vertical link is used by the sink to relay data to the surface station.

4.1.2. Three-dimensional Static Underwater Sensor Network

3-D underwater networks are used to detect and observe phenomena that cannot be adequately observed by means of ocean bottom sensor nodes. In 3-D UWSNs, sensor nodes float at different depths in order to observe a given phenomenon. For 3-D case, I.F. Akyildiz *et al.*, [4] proposed a very innovative approach where sensors are anchored to the ocean bottom, equipped with floating buoy that can be inflated by a pump. The buoy pulls the sensor towards the ocean surface as shown in Figure 3. The depth of the sensor can then be regulated by adjusting the length of the wire that connects the sensor to the anchor. Three deployment strategies have been proposed: the 3D random, the bottom random and the bottom grid. In the first two, sensors are randomly deployed on the bottom where they are anchored, however in the bottom grid strategy, sensors need to be assisted by one or multiple authors, which deploy the underwater sensors to obtain grid deployment on the bottom of the ocean. Simulation results show that the coverage ratio of obtained with the bottom-random strategy and the 3D-random strategy.

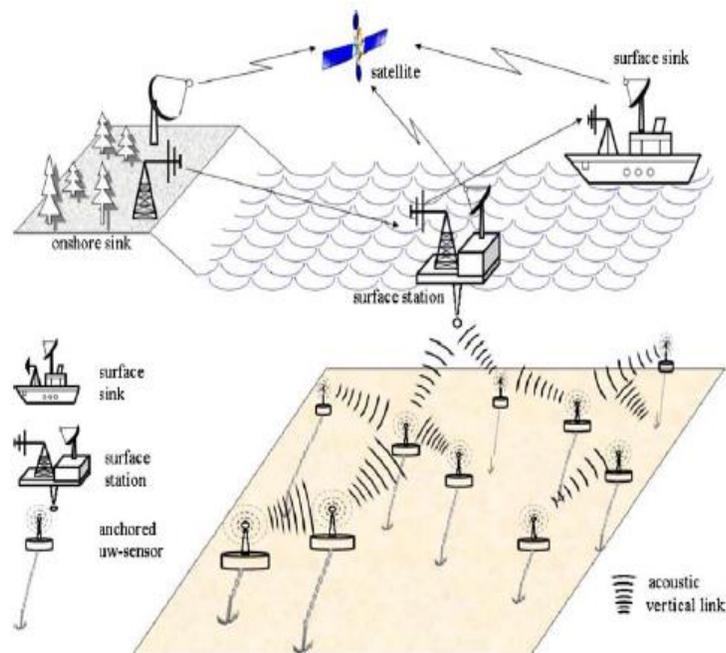


Figure 3. 3-D Underwater Sensor Network Architecture [4]

4.2. MobileUWSN

As shown in Figure 4, the other type of underwater sensor network is mobile sensor network in which the nodes are free floating and are not fixed or stationary, that is the nodes are assumed to be moving with the ocean currents. A mobile UWSN is a self-organizing network. Underwater sensor nodes may be relocated and moved by the process of advection and dispersion. After transport, the sensor must maintain and reorganize as a network in order to maintain communication. One of the design objectives of such a network is to make them rely on local intelligence and less dependent on communications from online shores. In [6], I.F. Akyildiz, introduce two classifications for mobile underwater sensor network:

- Mobile UWSNs for long-term non-time-critical aquatic monitoring – finds its applications in oceanography, Marine biology, deep-sea archaeology, seismic predictions, pollution detection and oil/gas field monitoring.
- Mobile uwsns for short-term non-time-critical aquatic monitoring – finds its applications in natural resource discovery, hurricane disaster recovery, loss treasure discovery.

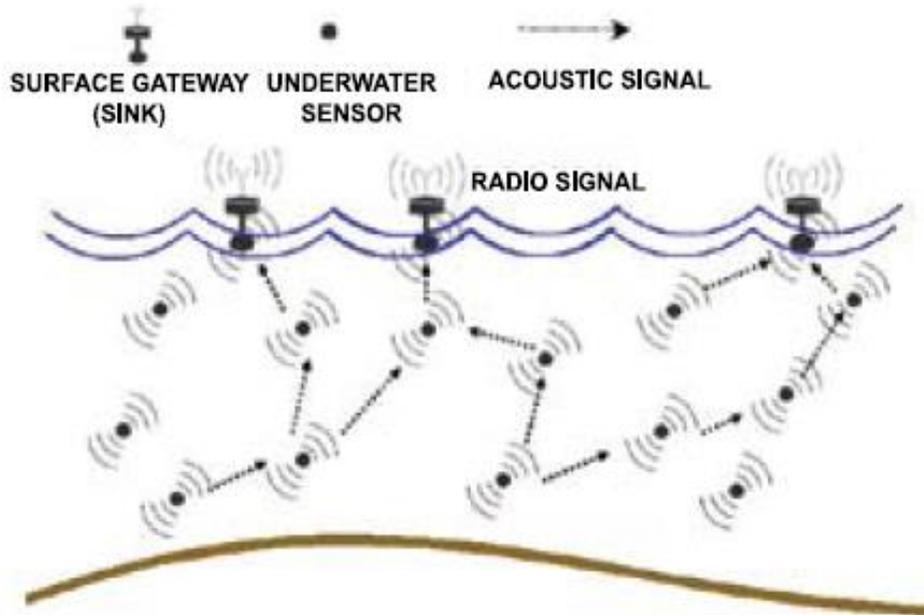


Figure 4. A Mobile Underwater Sensor Network [5]

5. Comparison Between TWSNs and UWSNs [30]

TERRESTRIAL WSNs	UNDERWATER WSNs
Different topologies can be applied, assuming that the nodes are stationary.	Nodes continue to move with water current so network cannot be viewed as a fixed topology.
Nodes are considered moving in 2D space.	Nodes can move in 3D volume without following any pattern.
The destination is fixed and seldom changes its location.	Sinks are placed on the water surface and can move with water current.
Deployment is deterministic as nodes are deployed manually, so routing path is pre-determined.	More self-configuring and self-organizing routing protocols are required because of non-uniform deployment.
Nodes are assumed to be homogeneous throughout the network.	Nodes are heterogeneous resulting in multiple technical issues related to data routing.
Radio waves are available, nodes can communicate with low propagation delays [27].	Acoustic waves replace radio waves (at a speed of 1.5×10^3 m/s) result in high propagation delays [28].
High data rate, normally in the order of MHz.	Low data rate, normally in the order of KHz.

Larger batteries can be used and can be replaced or recharged with ease.	Battery power is limited and usually cannot be recharged or replaced.
Nodes are less error prone.	Nodes are more error prone and can die soon because of the harsh environment conditions. Also, nodes may leave the area range of communication.
Schemes such as ToA and TdoA are used.	TdoA and ToA are not feasible.
Memory- Very limited storage capacity.	May need some more memory to be able to do some data caching as channels are intermittent.
More stable, especially in terms of communication links.	Link stability is an important issue in underwater sensor networks.
Cost- Expected to become increasingly inexpensive.	They are more expensive, especially due to the more complex protection.
Spatial Readings- readings are often correlated.	Readings are mostly correlated due to the higher distance between sensors.
Geographic routing is supported.	Geographic routing is not supported because of the low frequency bands in UWNSs [29].
Forward error correction schemes are used in order to increase and make the network more robust.	Due to the smaller bandwidth in underwater sensor network, Forward error correction schemes are not easily practiced.
This network is considered to be more stable as compared to UWSNs in terms of the communication links.	This Network is less stable because of the aspect of node mobility with ocean currents.

6. Data Aggregation Protocols ForUWSNs

Data aggregation is a fundamental issue for any network, and data aggregation protocols are considered to be in charge of discovering and maintaining the efficiency of a sensor network. Underwater sensor networks have been studied for decades, underwater networking and data aggregation schemes are still in a chaotic age of research. In this section, different aggregation schemes used or proposed by researchers are shown and reviewed for any kind of sensor network.

6.1. Parametric Chain Based Routing Approach (PCRA)

Feng Sen *et al.* [25], proposed Energy Efficient PEGASIS Based protocol (EEPB) which has a network architecture that resulted in serious performance issues such as non efficient election of the leader node. Additionally, it also failed to take certain parameters into concern, such as the distance between nodes and the leader nodes, energy, residual energy and the congestion are the major issues involved in underwater sensor networking. So in order to handle these environmental and other issues, Aarti *et al.* [7], 2013 proposed a parametric chain based routing approach (PCRA) in which the algorithmic approach for next neighbor selection is defined. The work is about to define an optimized path that will be started from the farthest node from the base station and after covering all the network nodes, transfer the aggregate data to the base station. This chain based aggregate routing assumes that all the nodes are in floating motion as the sensor network type in underwater. The work is about to generate a chain based path over a less congested nodes so that network life will be improved. The nodes will be organized to form chains which can either be accomplished by sensor nodes themselves using a greedy algorithm starting from some node. When a node dies, the chain is reconstructed in the same manner to bypass the dead node. The cluster head initiates a token passing procedure to start the transmission of data from the chain ends. Each node aggregates the data from its neighbor with its own data and generate a single packet of the same length. An efficient and reliable algorithm is defined for next neighbor selection. However, it has eliminated the issues based on parameters such as distance, residual energy, congestion associated with the nodes, and node localization is done at the maximum level keeping the limitations of underwater environment and is inspected and maintained by the base station in the absence of GPS which further helps in increasing making the connectivity between the nodes more efficient and reliable at the maximum level as no node is lost using this approach but it has a serious performance issues as well, such that all the sensor nodes are assumed to be homogenous but in underwater sensor network heterogeneous network is common. Inclusion of a heterogeneous set of sensor nodes raises multiple technical issues related to data routing (Shin *et al.* [8], 2007).

6.2. VBF

In underwater sensor network, When it comes to the 3-D architecture, the mobility causes the maintenance as well as the recovery to turn into an issue. In order to handle this issue, Vector based forwarding (VBF) has been proposed by Xie *et al.* [9], in 2006. In this, all the data packets are transmitted from sink to the source using a redundant path which handles the issue regarding the loss of packets and the failure of the sensor nodes. The idea of a virtual routing pipe which is a vector like pipe from source to sink as shown in Figure 5, is proposed and all the packets are aggregated at the head node and the forwarded through this pipe from the source to the destination resulting in less network traffic, and also the network topology is efficiently managed.

Here, the information status of node is required, as only few sensor nodes take part in the transmission of data. But it has serious problems as well, such as it has routing efficiency that is not up to the mark as the use of virtual pipe can be an issue. Additionally, node mobility is an issue as it could happen that no node is responsible for the data packet transmission because it does not lie in the path of the virtual pipe which further results in less transmissions in large areas. Also, link quality and data packet efficiency is lost in case there exist a path other than the virtual pipe as it could be an issue because of the ocean currents. Additionally, Battery power is exhausted when a node along a routing path are used repeatedly to transmit the packets of data from source to the sink. Its three-way handshake nature also lessens the

link quality. They use the same idea of routing pipe, but rather than using a single pipe from source to destination, it defines per hop virtual pipe for each forwarder where every node checks its own current location and then chooses the path which helps in finding a routing path till the time a single node is available within the communication range.

Although, according to the simulation results in hop by hop vector based forwarding, it has proved out to be an efficient algorithm. Hence, produces better results for packet delivery ratio, mostly in the sparse deployment of nodes in large areas as compared to VBF. But, because of the hop-by-hop nature of this algorithm, HH-VBF succeeded in producing more signalling overhead as compared to the VBF.

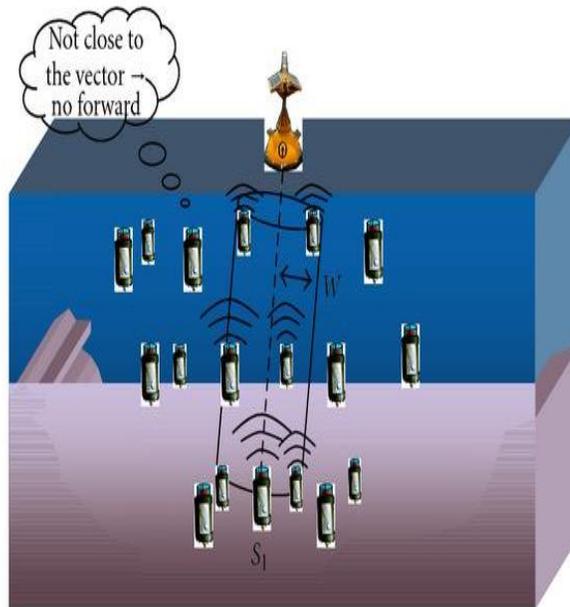


Figure 5. A High Level View of Vector Based Forwarding [9]

6.3.SFEB

Energy constraint is the most important aspect of underwater sensor network. Many proposals have been suggested in order to reduce the energy consumption. Chao and Hsiao *et. al.* [15], (2009) proposed a Structure Free and Energy Balanced Data Aggregation mechanism which reduced energy consumption by reducing the number of transmission and receiving. In SFEB, the data aggregation protocols are classified into two classes, structured class in order to achieve efficient data gathering by using a cluster based structure constructed at the time of network initialization and the other one is structure-free where it is assumed that the nodes may fail in such an environment in order to achieve reduced maintenance overhead and robustness as it does not spend extra energy to build a structure. This works in two phases: (i) data aggregators are dynamically selected to collect sensing data from the neighbor; (ii) the aggregators send the collected information to the sink. Although, simulation result shows that significant improvement in aggregation efficiency can be found if compared to other data aggregation algorithms such as DAA+RW, also an energy constraint is kept in mind and achieve a high data packet ratio, but if is to be set-up in an environment like underwater, node mobility, deployment of the nodes will result into performance issues as it could be possible for the nodes to travel with the ocean currents and go out of range.

6.4. LEACH

Heinzelman *et al.* [11], 2002, proposed a protocol for terrestrial sensor network named Low-Energy Adaptive Clustering Hierarchy (LEACH). This algorithm divides the sensor network into clusters and each cluster is composed of cluster head and numerous cluster nodes. The cluster head is responsible for the data communication between the clusters. Each cluster head can be chosen by numerous election algorithm. Clare *et al.* [12], also developed a time division multiple access protocol for LEACH where the sensor nodes are allowed to stay in a sleep state for a long time period to save power, whereas the cluster head must always stay awake to receive all the data from its cluster nodes and transmit it to other heads or sinks shown in Figure 6. However, it is assumed that all the nodes are within the communication range of each other and the sink, but in case of the underwater sensor network where node mobility is the prime concern, deployment of nodes become non-uniform, that makes this algorithm unsuitable for such an environment.

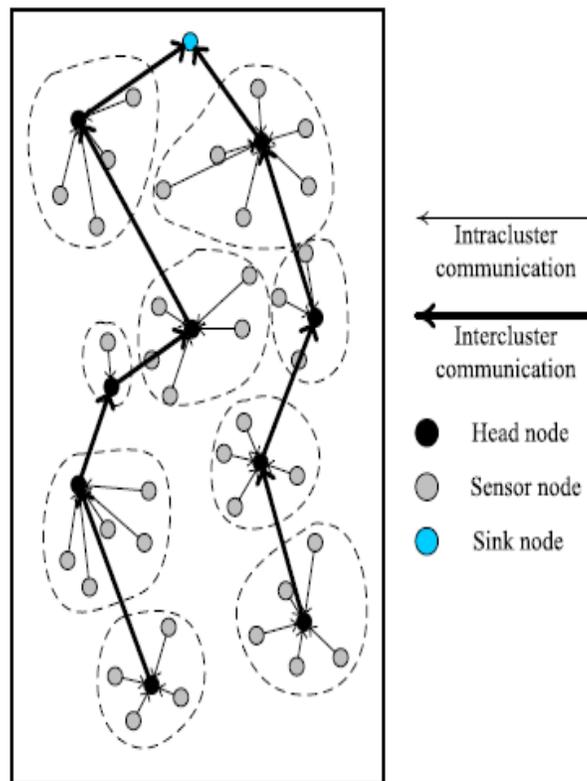


Figure 6. Topology Structure of Clustered UWSNs in LEACH [13]

6.5.LAA

UWSNs usually consist of a single sink node and a number of sensor nodes which send data packets to the sink node. Data aggregation is considered the best way to minimize the energy consumption and to achieve high data-packet ratio. Based on the certain constraints such as error rate, bandwidth, propagation delay etc, and keeping in mind the reliability of packet transmission, Dongseung *et al.* [16], (2011) proposed a novel location-based *k*ACK aggregation method for underwater sensor networks known as the LAA. The work is proposed in order to reduce the number of ACK (Acknowledgement at Transport layer)

transmissions in the network. In LAA, the sink node look for k ACK packets which are scheduled to be sent to k different sources and aggregates them into one aggregated-ACK packet. So instead of transferring one packet at a time, it sends the aggregated packet along with a minimum cost path. The assumption made by them was that the nodes know the information about the location of all the source nodes. An algorithm for aggregation cost, queue management also resulted in a more efficient and reliable protocol. LAA uses a cumulativeACK technique and operated on a scoped flooding based routing protocols in contrast to other such protocols proposed till date [34, 35].

This works in three phases: (i) *Queuing*, where data packets are received from the source nodes and the sink node generates ACK packets corresponding to the data packets as shown in Figure 7. These ACK packets are stored in an ACK queue and goes into the waiting state for the time AWT (Aggregation Waiting Time). As soon as the waiting time expires, the sink chooses the packet from the queue (ChACK- k) and determines which packet in the ACK queue is eligible for the aggregation. The aggregation cost is determined by calculating the physical distance and accordingly the aggregation priorities are set, and the process is repeated for k -ACK packets.

(ii) *Aggregation*, The lowest cost ACK packet is chosen and the process keeps repeating itself until all the eligible ACK packets are aggregated and no eligible packet exists. The sink aggregates the k -ACK packets chosen into one aggregated-ACK packet (AggrACK- k) where, the respective destination addresses are included in the AggrACK- k packet;

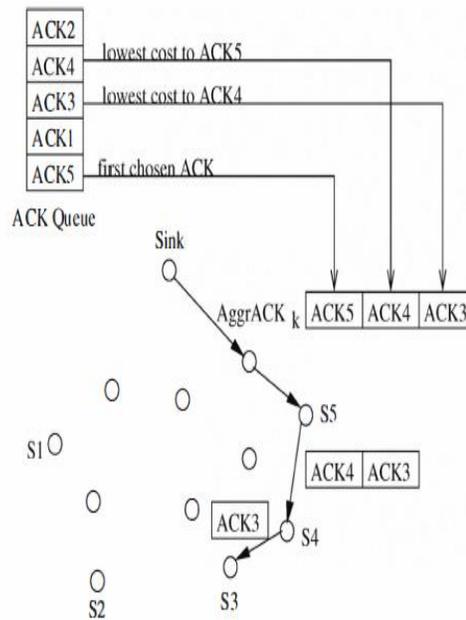


Figure 7. An Example Scenario of LAA [16]

(iii) *Network Phase*, The sink node sends the AggrACK- k to the destination node of ChACK-1, where the destination node takes ChACK-1 and deletes it from the AggrACK- k with its address information and retransmit the AggrACK- k to the next destination. This process keeps repeating itself until all the ACK packets are received by the destination nodes. Aggregation cost is determined and is then compared. The source node and the sink

node with the largest distance is set to have the longest AWT value because in case the ACK packet travels a long distance, the number of ACK transmission is reduced when it is aggregated with other ACK packets. The basic use of queue in the algorithm was: (i) Aggregation of multiple ACKs transmitted to different destinations; (ii) Removal of redundancy and duplicate data in the same data packets; (iii) Finally to merge the number of ACKs which are destined to the same destination but different numbers into one ACK packet. Sink does not reply to the source node with the corresponding ACKs immediately because it could result in ACK explosion. So, it holds the ACKs for a given AWT and awaits the rest of the ACKs eligible for the aggregation. However, unlike other algorithms based on end-to-end acknowledgement mechanism [32] leads to wasting of the network bandwidth.

LAA has numerous advantages over other data aggregation schemes and it has resulted in reduced data traffics as the ACK transmissions in the networks are reduced, increased throughput and high-data packet ratio with lower number of ACK transmission. End-to-end Optimization is one of the problem that needs to be resolved.

6.6. Reliable and Energy Balanced Routing Algorithm (REBAR)

The underwater sensor network has a drawback that the underwater environment is more dynamic and node mobility of the sensor nodes is one of the most important issues when it comes to the underwater sensor network. Jinming *et al.* [17], (2008) considered node mobility and took it as a positive factor, and proposed a reliable and energy balanced routing protocol (REBAR) in contrast with Jornet *et al.*, 2008 [35], which can be helpful in balancing energy depletion in the network. They solved the problem of sensor node depletion by frequent alterations of the locations of the nodes. Nodes near the sink are prone to die much earlier than other nodes so they tried to solve this problem by taking node mobility into concern and altering the position of the nodes. Therefore, nodes closer to the sink are set to smaller values in order to reduce the chances of being involved in the routing process as discussed. In this, all the sensor nodes are assigned a unique ID and possess a fixed range as shown in Figure 8. The sink is stationary on the water surface and the sensor nodes are randomly deployed. It is assumed that each node knows about its location as well as about the location of the sink.

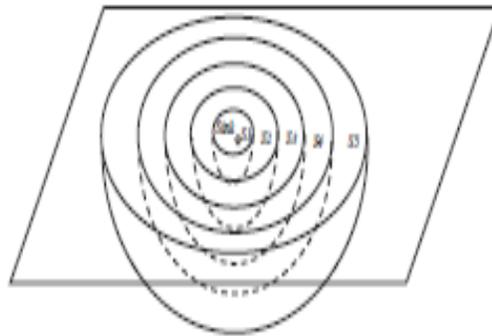


Figure 8. Sphere Energy Depletion Model in REBAR [30]

However, it has a serious problem. Large movements could also affect network performance since nodes would be required to update their location more frequently. Additionally, it is also assumed that the movements are completely dynamic in terms of directions, both vertically and horizontally and in such cases, a bottom node will move to the surface and then move back to the bottom as it is not possible to navigate vertically as only

small fluctuations are shown [35]. Moreover, the simulation results have not provided any information about the end-to-end delays, which can vary with the different node movement type. Delivery ratios are smaller and increases with an increase in the node movements. Energy consumption is also one of the most important issues as the energy is consumed critically.

6.7. Location based Clustering Algorithm for Data Gathering (LCAD)

Data needs to be transmitted and for the sensor nodes, it turns out to be the main source of energy consumption and this depends upon the distance between the sender and receiver. Also, the large number of data packets also results in draining off the sensor node's energy. By looking at the nexus between the two issues and keeping them in concern, Anupama *et al.* [18], in 2008 proposed a cluster based architecture for a three dimensional underwater sensor network where the sensor nodes are deployed at fixed relative depth from each other. All the sensor nodes are organized at their respective positions with the help of respective cluster heads, where the cluster heads are interlinked with each other by horizontal acoustic links, with a length restricted to a maximum of 500m for an effective communication proved in [42]. For a range of more than 500m, the use of autonomous underwater vehicle is practiced.

The communication process completes itself in three phases: (i) Setting up Phase: The cluster head is selected by using an optimal algorithm. (ii) Data aggregation: The data sensed by the cluster nodes is forwarded to the cluster head. Cluster head receives the data and aggregates that data in a less number of data packets in order to ensure the highest data packet delivery ratio. (iii) Transmission Phase: Finally, the data aggregated by the cluster head is forwarded to the base station by autonomous underwater vehicles. Autonomous underwater vehicles collect data packets from the cluster heads rather than every single sensor node in the network. Cluster heads possess more memory and energy as compared to the member nodes as more memory and energy is consumed by the cluster heads as compared to the member nodes and retrospect, it makes the sensor network not only more reliable, but also balance the load in the network. It has formed a grid structure just like a cellular network, where the cluster head is located at the center, which helps it in communicating efficiently with the respective member nodes. Therefore, for the most optimal results, they supported a compact deployment of sensor nodes at the lower level of the structure and more dispersed at the higher level of the structure.

In accordance to the simulation results, the work has proved itself to be effective in terms of network lifetime. Though, location based clustering algorithm for data gathering has numerous advantages over other similar proposed network protocols, it has serious performance issues as well, especially that it depends completely on a grid structure, which is highly efficient when it comes to the terrestrial environment but is not feasible in underwater environments as node mobility is the prime concern as in case of the underwater sensor network. In underwater sensor networks, the nodes can move with the ocean current and can leave the network. Therefore, it could be possible that the nodes enter the neighbor grid or leave different grids frequently resulting in communication loss.

6.8. Minimum-Latency Aggregation Scheduling (MLAS)

High propagation delay and data aggregation latency are the major issue in developing an efficient data aggregation algorithm for underwater sensor network. In order to handle this issue, Zuodong *et al.* [19], (2011) proposed a minimum-latency aggregation scheduling (MLAS) addressing the problem of minimizing the data aggregation latency and minimizing the propagation delay through the network phase. They succeeded in developing a virtual slot based approximation algorithm which efficiently exploits multiplexing opportunities of time

domain and present its theoretical bound. Shortest Hop Tree (SHT) completes in three phases: (i) Tree construction, where each node sets its parent and constructs a tree based structure; (ii) A concept of virtual slot is introduced in order to efficiently exploit multiplexing opportunities of the time domain. (iii) Network phase, the data is aggregated and is sent to the sink.

The advantage of SHT is that it introduces a virtual slot with a key feature that they are allowed to overlap and be assigned with a different slot sizes resulting in minimizing the aggregation latency, and also the propagation delays. However, it can be enhanced in order to obtain more efficient time domain utilization. Additionally, node mobility and packet loss is one of the major concern and must be considered.

6.9. Distributed Minimum-Cost Clustering Protocol (MCCP)

Protocols proposed for cluster formation in the terrestrial sensor network do not hold the same reliability and efficiency in underwater sensor network as in such an environment, node mobility prediction is a challenging aspect in underwater sensor network arises due to the fleeting characteristic of nodes underwater. In order to improve energy efficiency and prolong the network life, Pu *et al.* [20], (2007) proposed a distributed minimum cost clustering protocol (MCCP) which uses a cluster based approach where energy, residual energy and relative location of cluster head and sink are the basic three parameters for cluster formation. In this approach, initially all the member nodes are candidates for cluster heads and cluster members. Every node constructs a neighbor set to form a cluster. The average cost is calculated and the one with the minimum average cost is selected as a cluster head. After that, the cluster head broadcasts an INVITE message to all the other cluster nodes to become its cluster's member, otherwise it sends a JOIN message to the specific cluster head.

Although, MCCP is an efficient centralized algorithm for data aggregation as it eschews the formation of hot spots near the underwater sink. Also, it has the ability to rebalance the traffic by clustering the sensor nodes time to time. But it has a few drawbacks as well, such as it does not support multi-hop routing. It has the ability to re-cluster the network, but re-clustering the network can take months, due to which it could be possible that the nodes can leave and enter different clusters as the underwater sensor network is assumed to be mobile.

6.10. Distributed Underwater Clustering Scheme (DUCS)

In underwater sensor network, sensor nodes have limited power and battery so it is not feasible to recharge or replace the nodes. Keeping this issue into concern, Domingo and Prior *et al.* [22], 2007 proposed a clustering scheme which supports the node mobility and energy degradation issue, and named it as distributed underwater clustering scheme (DUCS), an adaptive self-organizing protocol where the nodes organize themselves into clusters and a cluster head is selected from each cluster. The cluster head agglomerates the data sent by the respective cluster members as shown in Figure 9. Data transfer between the member nodes and the cluster head takes place via a single hop. After that, the cluster head performs data aggregation on the received data and forwards the data to the sink via multi-hop routing with the help of remaining cluster heads, the redundancy of data is minimized and the energy is saved. Cluster heads are responsible for both inter-cluster communication as well as intra-cluster communication.

Cluster heads are selected through a randomized rotation among different nodes in order to avoid draining of the battery from a particular node. DUCS works in two rounds: (i) Set-up Phase: A network is formed by dividing the network into a number of clusters and cluster heads are formed using the respective cluster formation algorithm; (ii) Network operation

phase: Transfer of the packets is completed in this phase. DUCS has turned out to be an efficient clustering and aggregation scheme as simulation results have shown increased throughput and also achieves a high packet delivery ratio.

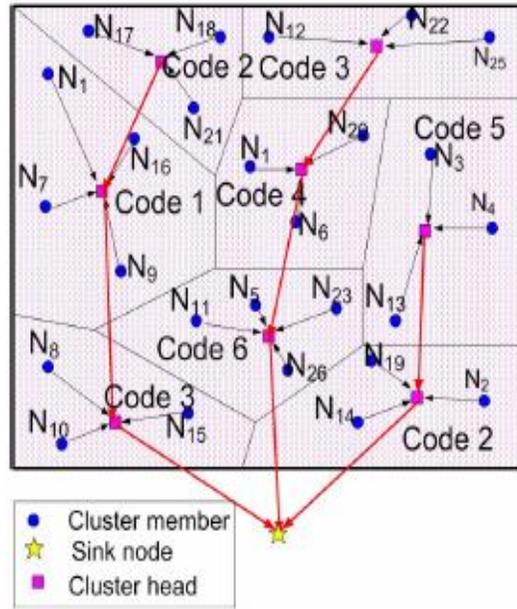


Figure 9. Example Network with DUCS [22]

Although this scheme is efficient, but it has some serious issues like node mobility is not considered so node movements due to ocean currents can affect the structure of clusters that it reduces the network overhead. Also, the cluster head is bound to send the data to another cluster head only. In that case, ocean currents can move two cluster head nodes far away such that, while data transfer between the two cluster head, there are a few non-cluster head nodes available between them.

6.11.A Self Healing Clustering Head Selection Algorithm (CHSA)

The unique characteristics of up/down transmission direction, good cluster head distribution, energy consumption for cluster heads are the basic parameters in underwater sensor network. Node clustering is a useful topology management approach to reduce the energy consumption and make use of and deriving benefit by data aggregation in underwater sensor networks. To tackle these challenges, Chenn *et. al.* [13], 2010 proposed a clustering head selection algorithm, assuming that each node is able to act both as a source and cluster head also rotating the server role among all nodes such that the energy consumption is uniformly balanced. Cluster head consume more energy as compared to member nodes so in order to balance the energy consumption, the cluster heads rotation is taken into concern where if a node acts as the head within a time period, another sensor node will be selected as the head at the next time period as shown in Figure 10.

The entire communication process is completed in three phases: (i) Candidates for the cluster head are located and a cluster head is selected on the energy efficiency basis and the member nodes are linked with the respected cluster heads, and the steps keep repeating until an efficient network is formed; (ii) The sink broadcasts message containing each cluster

head's ID and its member nodes' IDs, and a header to all the sensor nodes to confirm the formation of the clusters; (iii) Re-clustering is practiced after each fixed time period for the cluster head rotation for load balancing.

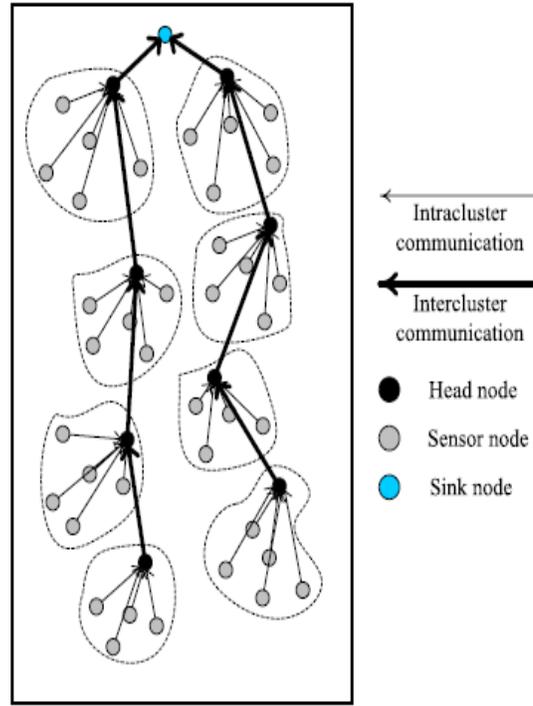


Figure 10. Topological Structure of CHSA[13]

Simulation results have proven that the work outperforms the working of LEACH and other such algorithms in terms of reliability. This algorithm has various advantages in terms of the number of nodes alive over a specific period of time, the total amount of data received at the sink over time and the total amount of data reviewed at the sink for a given amount of energy. Even though, a number of similar protocols have been proposed for the terrestrial environment but they do not hold good in underwater conditions. It also turns out to be a self-healing system where the system can automatically recover itself after a failure occurs. Although the work result in the efficient selection of the cluster heads but heavy energy consumption could result in re-clustering. Also, nodes degradation with underwater disasters have not been addressed.

6.12. Multi-Sink Opportunistic Routing Protocol

Node deployment is one of the open research issue which is to be handled in order to develop an efficient and reliable routing algorithm. In order to handle the issue, Tonghong *et al.* [23], (2008) proposed a multi-sink opportunistic routing protocol for underwater sensor network. In this approach, it assumed the 2DUWSN architecture for a shallow-water coastal area. In this architecture, five major types of elements are included such as a sensor node, mesh node, underwater sink, surface buoy and monitoring center. Sensor node, mesh node

and the sink are surfacing at the ocean bed through a wire while surface buoy is at the ocean surface and is connected to the monitoring system through the internet. Mesh nodes are like cluster head installed with higher memory, more energy and longer transmission range in order to proceed the data aggregation. The communication takes place in three steps: (i) Firstly, each sensor node transmits its data to the closest mesh node; (ii) Secondly, mesh nodes aggregate the data and transmit the aggregated packet to the underwater sink; (iii) Finally, the aggregated packets are transmitted to the surface buoy and through the internet, they are sent to the monitoring system. Although this protocol has proved to be an efficient protocol as exploitation of packet duplications results in high packet delivery ratio, but also it has some serious issues such as the assumption that each mesh node contains information not only about the nearby nodes, but also about the underwater sinks, node ids, node locations etc. Additionally it is assumed that there exists a stationary network which is not pragmatic in the underwater environment because of the ocean currents. Also, replicated copies of the same packet can be generated and transmitted. Therefore, duplication of aggregated packets could be introduced as the packets are transmitted along a redundant path.

6.13. Multipath Virtual Sink Architecture

In underwater sensor networks, reliability, energy consumption and capacity are important for defining the topology of a network. Also, redundancy and robustness are the prime ingredient of the underwater sensor network. Also, there are possibilities such as that a part of a network stops operating while the other party still continues to work. In order to achieve the significant robustness and the redundancy, and keeping all the other constraints in mind, Seah and Tan *et al.* [24], (2006) proposed a multipath virtual sink architecture that ensures that the sensor network will continue to work even if a portion of the network stops working properly. Network topology is an important aspect. Therefore, in this architecture, the whole network is divided into clusters where each cluster includes one or more aggregation points as shown in Figure 11.

Each aggregation point forms a mesh network and is ultimately connected to the underwater sink. It ensures that data packets are received at the sink by assuming that the local sinks are connected through a high speed link in order to fulfill the transmission needs. The connection is sporadic in nature and the bandwidths are small as well. Therefore, each sensor node senses the data and transmits the data as soon as the conditions are favorable for the communication. For the delayed data, the network try to transmits the data packets through multiple paths in order to ensure the high data delivery ratio. The data transmitted by the sensor nodes are delivered to the respective aggregating points where the data is aggregated.

Each and every sink broadcasts a message and as soon as the sink receives the message, it updates its hop count value and then broadcasts it again after making an increment by one. This approach enabled the algorithm to make the architecture more reliable and efficient, but redundancy or transmission is an open research issue as it consumes underwater resources.

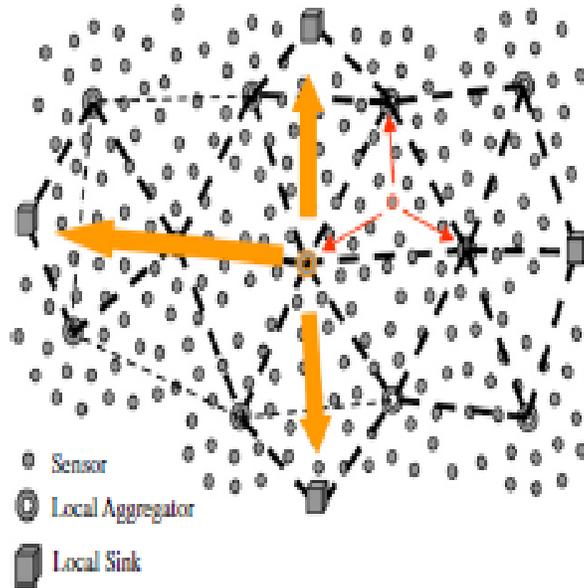


Figure 11. Underwater Network Topology for Multipath Virtual Sink Architecture [30]

7. Open Research Issues

The growing interest in UWSN is retrospectively attracting many top notch researchers. Underwater sensor networks are becoming highly important for an effective and reliable underwater communication. A number of researchers have worked upon sensor networks as discussed above and faced a number of problems in order to propose an effective communication network. Extreme water condition, hardware limitations, communication issue, *etc.*, are the issues that have to be taken into concern while designing the sensor network architecture. Based on the above survey, it can be clearly devised that continuous node mobility is the most important issue in the underwater sensor network, especially the highly sparse deployed network. Predicting the node mobility is a highly challenging task and many researchers have already proposed algorithms for node mobility predictions such as I.F. Akyldiz [5], Sezaki [36]. However, the proposed algorithms proved out to be efficient for the terrestrial sensor network, but these techniques are not pertinent for underwater sensor network because of the harsh undersea conditions. Node Mobility is an important and open research issue. Additionally, network topology for 3-D underwater architecture also counts to critical issues. Due to the unique characteristics of underwater sensor network, many proposed algorithms for terrestrial network do not hold effective in UWSNs. In accordance with the already proposed algorithms, Zhou *et al.* [37], 2011 and Caruso *et al.* [38], 2008 proposed node prediction algorithms for underwater sensor network but is still worth more investigation.

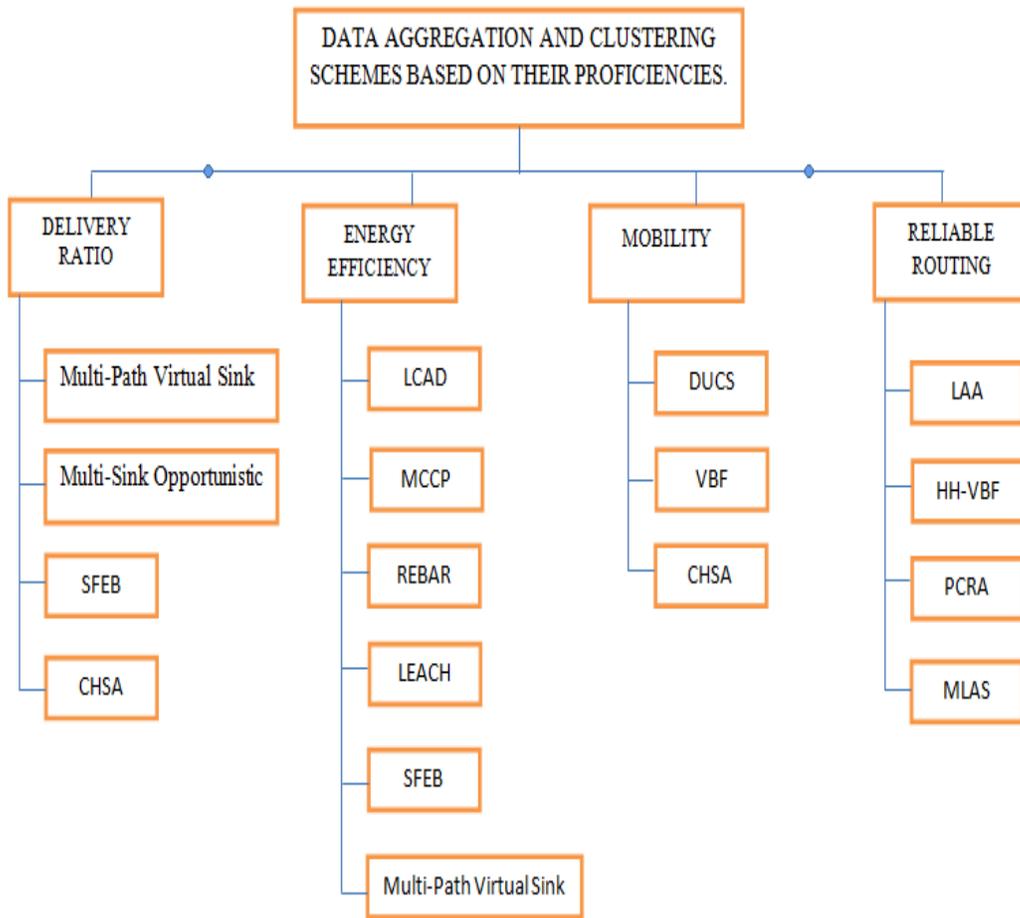


Figure 12. Classification Based on the Proficiencies of the Algorithms Reviewed

Looking at the nexus between the algorithms proposed till date, it can be devised that very little work is published relating to data packet size optimization as in [39] and [40] and packet fragmentation in [41]. Optimized architecture for efficient data collection and transmission of the data, and retrieval of the data in such a challenging environment is also an important open research issue which needs to be resolved. Also, new flow control strategies are required in order to tackle the retrieval issues such as delays. Cross layer approach for the optimization of the main network in order to design a wireless sensor network that can adapt to the changing underwater environments is also open issue. The cross layer approach resulted in the increased performance of WSN by the optimization as set-out above and many techniques have been proposed such as I.F. Akylidiz [5] for the terrestrial wireless sensing networks. As the communication channel in the underwater sensor network is prone to be damaged more quickly than in terrestrial conditions which consequently result in the deterioration of the sensor network, therefore it requires more optimization.

8. Conclusion

In this paper, we presented an overview of the state of the art of data aggregation and clustering schemes in underwater wireless sensor network. Energy constraint, node mobility and various other parameters are taken into concern, and numerous strategies and techniques are proposed in order to fulfil the requirements. The development of clustering and aggregating techniques is regarded as an essential research area, which helps making the network more efficient and reliable. Also, the protocols are compared with different parameters in Table 1. The paper summarizes the potential applications, architectures and is compared to terrestrial networks. Finally, we have discussed various clustering and data aggregation schemes proposed for UWSNs highlighting the advantages as well as the reliability issue for each proposed scheme. Additionally, protocols are classified on the basis of their proficiencies in Figure 12. The ultimate objective of this paper is to encourage researcher's efforts to lay down and to develop a solid foundation for the development of new advanced schemes by providing the basis of aggregating and clustering schemes proposed to date.

Table 1. Comparison of Underwater Sensor Networks Clustering Schemes and Data Aggregation Protocols

DATA AGGREGATION PROTOCOLS AND CLUSTERING SCHEMES.	Bandwidth Efficiency	Energy Efficiency	Performance	Delivery Ratio	Year of Introduction
PCRA (Aarti et al., 2013)	Low	High	Moderate	Moderate	2013
VBF (Xie et al., 2006)	Moderate	Moderate	Low	Low	2006
HH-VBF (Nicolaou et al., 2007)	Moderate	Low	Moderate	Moderate	2007
SFEB (Chao et al., 2009)	Moderate	High	High	High	2009
LEACH (Heinzelman et al., 2002)	Low	Low	Moderate	Low	2002
LAA (Dongseung et al. 2011)	Moderate	Low	Moderate	High	2011
REBAR (Jinming et al., 2007)	Moderate	High	Moderate	Moderate	2007
LCAD (Anupama et al., 2008)	Moderate	Moderate	Low	Moderate	2008

MLAS (Zuodong et al., 2011)	High	Moderate	Moderate	High	2011
MCCP (Pu et al., 2007)	Moderate	High	Moderate	Low	2007
DUCS (Domingo and prior, 2007)	Moderate	Moderate	Low	Moderate	2007
CHSA (Chenn et. al 2010)	Low	Moderate	Low	Moderate	2010
MVSA (Seah and Tan, 2006)	Moderate	Low	Moderate	Moderate	2006
MSOP (Tonghong et al. 2008)	Moderate	Moderate	Moderate	High	2008

References

- [1] S. Peter, K. Piotrowski and P. Langendoerfer, "On concealed data aggregation for WSNs", In Proc. of 4th IEEE Consumer Communi. Networking Conf. (CCNC), (2007), pp. 192 -196.
- [2] Field Test: Second field test for the AOSN program, Monterey Bay, August 2003, Available from <http://www.mbari.org/aosn/MontereyBay2003/MontereyBay2003Default.htm>.
- [3] W.K.G. Seah and H.X. Tan, "Multipath Virtual Sink Architecture for Underwater Sensor Networks", in Proc. of 2002. IEEE OCEANS AsiaPacific Conference, Singapore, (2006) May 16-19.
- [4] I. F. Akyildiz, D. Pompili and T. Melodia, "Underwater Acoustic Sensor Networks: Research Challenges", Ad Hoc Networks (Elsevier), vol. 3, (2005) March, pp. 257-279.
- [5] K. Ovaliadis, N. Savage and V. Kanakar, "Energy Efficiency in Underwater Sensor Network", Journal of Engineering Science and Technology Review, vol. 3, (2010) January, pp. 151-156.
- [6] I. F. Akyildiz, D. Pompili and T. Melodia, "State-of-the-Art in Protocol Research for Underwater Acoustic Sensor Networks", ACM Mobile computing and communication Review, vol. 11, no. 4, (2007) October.
- [7] A. Sanjiv Kumar Tomar, "A Parametric Chain based Rounting Approach for Underwater Sensor Network", International Journal of Computer Trends and Technology (IJCTT), vol. 4, no. 5, (2013) May.
- [8] K.-P. Shin, S.-Y. Wang, H.-C. Chen and B.-J. Liu, "On target coverage in wireless heterogeneous sensor networks with multiple sensing units", Proc. of the twelfth IEEE Symposium on computers and communications, ISCC2007, vol. 90, (2007) July 1-4, pp. 185.
- [9] P. Xie, J.-H. Cui and L. Lao, "VBF: vector-based forwarding protocol for underwater sensor networks", Networking technologies, services, and protocols: performance of computer and communication networks; mobile and wireless communications systems. Berlin/Heidelberg: Springer, (2006), pp. 1216-21.
- [10] N. Nicolaou, A. See, P. Xie, J. H. Cui and D. Maggiorini, "Improving the robustness of location based routing for under water sensor networks", Proc. of the IEEE Oceans, Europe, (2007) June, pp. 1-6.
- [11] W. B. Heinzelman, A. P. Chandrakasan and H. Balakrishnan, "An application specific protocol architecture for wireless micro sensor networks", IEEE Trans. Wireless Communication, vol. 1, no. 4, (2002) October, pp. 660-670.
- [12] L. Clare, G. Pottie and J. Agre, "Self-organizing distributed sensor networks", Proc. SPIE Conf. Unattended Ground Sensor Technologies and Applications, Orlando, FL, vol. 3713, (1999) April, pp. 229-237.
- [13] C.-J. Huang, Y.-W. Wang, C.-F. Lin, Y.-T. Chen, H.-M. Chen, H.-Y. Shen, Y.-J. Chen, I.-F. Chen, K.-W. Hu and D.-X. Yang, "A self-healing clustering algorithm for underwater sensor networks", Springer, Cluster Computing, vol. 14, (2011) July, pp. 91-99.
- [14] O. Younis and S. Fahmy, "HEED: A Hybrid, energy efficient, distributed clustering approach for ad hoc sensor networks", IEEE Trans. Mobile Computing, vol. 3, no. 4, (2004) October-December, pp. 366-379.

- [15] C.-M. Chao and T.-Y. Hsiao, "Design of Structure-Free and Energy-Balanced Data Aggregation in Wireless Sensor Networks", 11th IEEE International Conference on High Performance Computing and Communications, (2009) pp. 222-229.
- [16] D. Shin, S. Lee and D. Kim, "Location-based k-ACK Aggregation Method for Underwater Sensor Networks", IEEE Oceans, (2011), pp. 1-5.
- [17] C. Jinming, W. Xiaobing and C. Guihai, "REBAR: Reliable and Energy Balanced Routing Algorithm for Underwater Sensor Networks", Proc. of the 7th IEEE International Conference on Grid and Cooperative Computing, (2008), pp. 349-355.
- [18] K. R. Anupama, A. Sasidharan and S. Vadlamani, "A location-based clustering algorithm for data gathering in 3-D underwater wireless sensor networks", Proc. of the International Symposium on Telecommunications, Tehran, (2008), pp. 343-348.
- [19] Z. Wu, C. Tian, H. Jiang and W. Liu, "Minimum-latency Aggregation Scheduling in Underwater Sensor Networks", IEEE Communications Society, pp. 1-5, (2011) June 5-9.
- [20] W. Pu, L. Cheng and Z. Jun, "Distributed minimum cost clustering protocol for underwater sensor networks (UWSNs)", Proc. of the IEEE International Conference on Communications, Scotland, (2007), pp. 3510-3515.
- [21] M. C. Domingo and R. Prior, "A distributed clustering scheme for underwater wireless sensor networks", Proc. of the IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications, (2007) September 3-7, pp. 1-5.
- [22] L. Tonghong, "Multi-sink opportunistic routing protocol for underwater mesh network", Proc. of the International Conference on Communications, Circuits and Systems, (2008).
- [23] W. K. G. Seah and H. P. Tan, "Multipath virtual sink architecture for wireless sensor networks in harsh environments", Proc. of the First International Conference on Integrated Internet Ad Hoc and Sensor Networks. (France), (2006).
- [24] J. Yick, B. Mukherjee and D. Ghosal, "Wireless Sensor Network Survey", Computer Networks, Elsevier, vol. 52, (2008), pp. 2292-2330.
- [25] J.-H. Cui, J. Kong, M. Gerla and S. Zhou, "The challenges of building scalable mobile underwater wireless sensor networks for aquatic applications", IEEE Network, vol. 20, no. 3, (2006), pp. 12-18.
- [26] Z. Zhou, "Efficient multipath communication for time-critical applications in underwater acoustic sensor networks", IEEE/ACM Transactions on Networks, vol. 19, no. 1, (2011), pp. 28-41.
- [27] J. Heidemann, "Research challenges and applications for underwater sensor networking", Proc. of the IEEE Wireless Communications and Networking Conference, vol. 4, (2006) April 3-6, pp. 1-5.
- [28] M. C. Domingo and R. Prior, "Energy analysis of routing protocols for underwater wireless sensor networks", Computer Communication Journal, vol. 31, no. 6, (2008), pp. 1227-38.
- [29] M. Ayaz, I. Baig, A. Abdullah and I. Faye, "A survey on routing techniques in underwater wireless sensor networks", Elsevier Journal of Network and Computer Applications, vol. 34, no. 6, (2011) June, pp. 1908-1927.
- [30] Y. Li, Fall 2010 - GSU by Robert Persuad, "LEACH Protocol for Wireless Sensor Networks", Wireless Sensor Networks, CSC, 8823, (2010).
- [31] Information Science Institute, "Transmission Control Protocol", RFC793, (1981) September.
- [32] D. Shin, D. Hwang and D. Kim, "DFR: An efficient directional flooding-based routing protocol in underwater sensor networks", Wireless Communications and Mobile Computing (Online), (2011) January.
- [33] J. M. Jomet, M. Stojanovic and M. Zorzi, "Focused beam routing protocol for underwater acoustic networks", Proc. of ACM International Workshop Underwater Networks, (2008), pp. 75-82.
- [34] C. Jun-Hong, "The challenges of building mobile underwater wireless networks for aquatic applications Network", IEEE Network, vol. 20, no. 3, (2006), pp. 12-8.
- [35] W. Creixell and K. Sezaki, "Routing protocol for ad hoc mobile networks using mobility prediction", International Journal Ad Hoc Ubiquitous Computing, vol. 2, no. 3, (2007), pp. 149-56.
- [36] Z. Zhou, Z. Peng, J.-H. Cui, Z. Shi and A. C. Bagtzoglou, "Scalable localization with mobility prediction for underwater sensor networks", IEEE Transactions on Mobile Computing; vol. 10, no. 3, (2011), pp. 335-48.
- [37] A. Caruso A., "The meandering current mobility model and its impact on underwater mobile sensor networks". In Proc. of the IEEE 27th conference on Computer Communications, (2008) April 13-18.
- [38] M. Stojanovic, "Optimization of a data link protocol for an underwater acoustic channel", Proc. of the Oceans, Europe, vol. 2, (2005) June, pp. 68-73.
- [39] S. Basagni, "Choosing the packet size in multi-hop underwater networks", Proc. of the IEEE Oceans, Sydney, vol. 1, (2010) May 24-27, pp. 1-5.
- [40] S. Basagni, C. Petrioli, R. Petrocchia and M. Stojanovic, "Optimizing network performance through packet fragmentation in multi-hop underwater communications", Proc. of the IEEE Oceans, Sydney, vol. 1, (2011), pp. 1-7.

- [41] M. Stojanovic, J. A. Catipovic and J. G. Proakis, "Phase-coherent digital communications for underwater acoustic channels", *IEEE Oceans*, vol. 19, no. 1, pp. 100–111, 1994.
- [42] E. Cayirci, H. Tezcan, Y. Dogan and V. Coskun, "Wireless sensor networks for underwater surveillance systems", *Ad Hoc Networks*, in press; doi: 10.1016/j.adhoc.2004.10.008.