ISSN: 2277-3754
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 1, Issue 4, April 2012

Features of WSN and Data Aggregation techniques in WSN: A Survey

Sushruta Mishra, Hiren Thakkar

Abstract*— Sensor networks consist of several sensor nodes which co-operatively send sensed data to base station. One of the critical constraints of sensor nodes is the power consumption requirement. Sensor nodes carry limited, generally irreplaceable, power sources. In real life scenario when these sensor nodes are deployed in any environment it is possible that these nodes may not be closely located to the base station in fact the nodes themselves may be far off distance from each other. As sensor nodes are battery driven, an efficient utilization of power is essential to reduce data traffic inside sensor networks thus reduce amount of data that need to send to base station thereby enhancing the network lifetime. The main goal of data aggregation algorithms is to gather and aggregate data in an energy efficient manner so that network lifetime is enhanced. In this paper we presents some important aspects of wireless sensor networks related to data aggregation and various techniques of data aggregation .Our aim is to provide a good understanding of data aggregation in WSN and its related issues.

***Index Terms*- Cluster head, Data Aggregation, Fault Tolerance, Simulators, Wireless Sensor Networks.**

I. INTRODUCTION

A sensor network is defined as being composed of a large number of nodes with sensing, processing and communication facilities which are densely deployed either inside the phenomenon or very close to it . Each of these nodes collects data and its purpose is to route this information back to a sink. The network must possess self-organizing capabilities since the positions of individual nodes are not predetermined. Cooperation among nodes is the dominant feature of this type of network, where groups of nodes cooperate to disseminate the information gathered in their vicinity to the user as shown in fig 1. Recent advances in micro-electro-mechanical systems (MEMS) technology, wireless communications, and digital electronics have made possible to develop low-cost, low-power, multifunctional sensor nodes that are small in size and communicate freely in short distances. These tiny sensor nodes, which consist of sensing, data processing, and communicating components, leverage the idea of sensor networks based on collaborative effort of a large number of nodes. Sensor networks represent a significant improvement over traditional sensors, which are deployed in the following two ways [1]:

- Sensors can be positioned far from the actual phenomenon, i.e., something known by sense perception. In this approach, large sensors that use some complex techniques to distinguish the targets from environmental noise are required.
- Several sensors that perform only sensing can be deployed. The positions of the sensors and communications topology are carefully engineered. They transmit time series of the sensed phenomenon to the central nodes where computations are performed and data are fused. The position of sensor nodes need not be engineered or pre-determined. This allows random deployment in inaccessible terrains or disaster relief operations. On the other hand, this also means that sensor network protocols and algorithms must possess self-organizing capabilities.

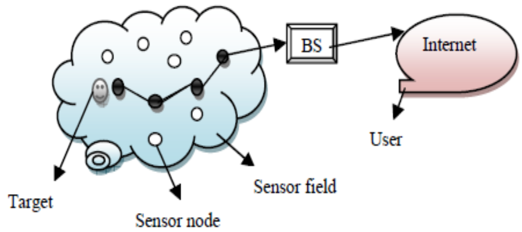

Fig 1 Sensor Nodes in a Sensor Network

As it is shown here there are several sensor nodes scattered randomly and the data content of individual sensor nodes gets collected in the sink. Then through internet the user can view the data collected by the network. [2]

Another unique feature of sensor networks is the cooperative effort of sensor nodes. Sensor nodes are fitted with an on-board processor. Instead of sending the raw data to the nodes responsible for the fusion, sensor nodes use their processing abilities to locally carry out simple computations and transmit only the required and partially processed data. A sensor node is made up of four basic components as shown in the figure a sensing unit, a processing unit, a transceiver unit and a power unit. They may also have application dependent additional components such as a location finding system, a power generator and a mobilizer. As shown in fig 2 Sensing units are usually composed of two subunits: sensors and analog to digital converters (ADCs). The analog signals produced by the sensors based on the observed phenomenon are converted to digital signals by the ADC, and then fed into the processing unit. The processing unit, which is generally associated with a small storage unit, manages the procedures

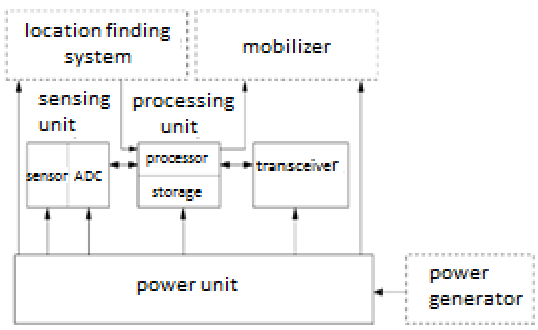

Fig 2 components of WSN

A transceiver unit connects the node to the network. One of the most important components of a sensor node is the power unit. Power units may be supported by a power scavenging unit such as solar cells. There are also other subunits, which are application dependent. Sensor networks may consist of many different types of sensors such as seismic, low sampling rate magnetic, thermal, visual, infrared, acoustic and radar, which are able to monitor a wide variety of ambient conditions that include the following [4]: temperature, humidity, vehicular movement, lightning condition, pressure, soil makeup, noise levels, the presence or absence of certain kinds of objects, ssand size of an object. Sensor nodes can be used for continuous sensing, event detection, even ID, location sensing, and local control of actuators. The concept of micro-sensing and wireless connection of these nodes promises many new application areas. We categorize the applications into military, environment, health, home and other commercial areas.[4]

MILITARY

Wireless sensor networks can be an integral part of military command, control, communications, computing, intelligence, surveillance, reconnaissance and targeting (C4ISRT) systems. The rapid deployment, self-organization and fault tolerance characteristics of sensor networks make them a very promising sensing technique for military C4ISRT. Since sensor networks are based on the dense deployment of disposable and low-cost sensor nodes, destruction of some nodes by hostile actions does not affect a military operation as much as the destruction of a traditional sensor, which makes sensor networks concept a better approach for battlefields. Some of the military applications of sensor networks are monitoring friendly forces, equipment and ammunition; battlefield surveillance; reconnaissance of opposing forces and terrain; targeting; battle damage assessment; and nuclear, biological and chemical (NBC) attack detection and reconnaissance.

ENVIRONMENT

Some environmental applications of sensor networks include tracking the movements of birds, small animals, and insects; monitoring environmental conditions that affect crops and livestock; irrigation; macro instruments for large-scale Earth monitoring and planetary exploration; chemical/biological detection; precision agriculture; biological, Earth, and environmental monitoring in marine, soil, and atmospheric contexts; forest fire detection; meteorological or geophysical research; flood detection; bio-complexity mapping of the environment; and pollution study. [5-15]

HEALTH

Some of the health applications for sensor networks are providing interfaces for the disabled; integrated patient monitoring; diagnostics; drug administration in hospitals; monitoring the movements and internal processes of insects or other small animals; telemonitoring of human physiological data; and tracking and monitoring doctors and patients inside a hospital[8][16][17].

HOME

As technology advances, smart sensor nodes and actuators can be buried in appliances, such as vacuum cleaners, micro-wave ovens, refrigerators, and VCRs [67]. These sensor nodes inside the domestic devices can interact with each other and with the external network via the Internet or Satellite. They allow end users to manage home devices locally and remotely more easily. The design of smart environment can have two different perspectives, i.e., human-centered and technology centered . For human-centered, a smart environment has to adapt to the needs of the end users in terms of input/output capabilities. For technology-centered, new hardware technologies, networking solutions, and middleware services have to be developed. The sensor nodes can be embedded into furniture and appliances, and they can communicate with each other and the room server. The room server can also communicate with other room servers to learn about the services they offered, e.g., printing, scanning, and faxing. These room servers and sensor nodes can be integrated with existing embedded devices to become self-organizing, self regulated and adaptive systems

OTHER COMMERCIAL

Some of the commercial applications are monitoring material fatigue; building virtual keyboards; managing inventory; monitoring product quality; constructing smart office spaces; environmental control in office buildings; robot control and guidance in automatic manufacturing environments; interactive toys; interactive museums; factory process control and automation; monitoring disaster area; smart structures with sensor nodes embedded inside; machine diagnosis; transportation; factory instrumentation; local control of actuators; detecting and monitoring car thefts; vehicle tracking and detection; and instrumentation of semiconductor processing chambers, rotating machinery, wind tunnels, and anechoic chambers.

II DESIGN ISSUES IN WSN [19] [20] [21] [25]

A sensor network design is influenced by many factors. These factors are important because they serve as a guideline to design a protocol or an algorithm for sensor networks. In addition, these influencing factors can be used to compare different schemes.

FAULT TOLERANCE

Some sensor nodes may fail or be blocked due to lack of power, have physical damage or environmental interference. The failure of sensor nodes should not affect the overall task of the sensor network. This is the reliability or fault tolerance issue. Fault tolerance is the ability to sustain sensor network functionalities without any interruption due to sensor node failures.[19][20][25]

SCALABILITY

The number of sensor nodes deployed in studying a phenomenon may be in the order of hundreds or thousands. Depending on the application, the number may reach an extreme value of millions. The new schemes must be able to work with this number of nodes. They must also utilize the high density nature of the sensor networks. The density can range from few sensor nodes to few hundred sensor nodes in a region, which can be less than 10 m in diameter [11]. The density can be calculated according to [8]:

$$\mu(R) = (N\pi R^2)/A$$

Where N is the number of scattered sensor nodes in region A; and R, the radio transmission range. Basically this gives the number of nodes within the transmission radius of each node in region A.

PRODUCTION COSTS

Since the sensor networks consist of a large number of sensor nodes, the cost of a single node is very important to justify the overall cost of the networks. As a result, the cost of each sensor node has to be kept low. The cost of a sensor node should be much less than 1$ in order for the sensor network to be feasible [25] Note that a sensor node also has some additional units such as sensing and processing units. In addition, it may be equipped with a location finding system, mobilizer or power generator depending on the applications of the sensor networks. As a result, the cost of a sensor node is a very challenging issue given the amount of functionalities with a price of much less than a dollar.

HARDWARE CONSTRAINTS

A sensor node is made up of four basic components as shown in Fig 2: a sensing unit, a processing unit, a transceiver unit and a power unit. They may also have application dependent additional components such as a location finding system, a power generator and a mobilizer. Sensing units are usually composed of two subunits: sensors and analog to digital converters (ADCs). Most of the sensor network routing techniques and sensing tasks require the knowledge of location with high accuracy. Thus, it is common that a sensor node has a location finding system. A mobilize may sometimes be needed to move sensor nodes when it is required to carry out the assigned tasks. All of these subunits may need to fit into a matchbox-sized module [39]. The required size may be smaller than even a cubic centimeter [69] which is light enough to remain suspended in the air. Apart from the size, there are also some other stringent constraints for sensor nodes. These nodes must [42]

- Consume extremely low power,
- Operate in high volumetric densities,
- Have low production cost and be dispensable,
- Be autonomous and operate unattended,
- Be adaptive to the environment.

NETWORK TOPOLOGY

Sheer numbers of inaccessible and unattended sensor nodes, which are prone to frequent failures, make topology maintenance a challenging task. Hundreds to several thousands of nodes are deployed throughout the sensor field. They are deployed within tens of feet of each other [39]. The node densities may be as high as 20 nodes/m3 [77]. Deploying high number of nodes densely requires careful handling of topology maintenance. the schemes for initial deployment must

- Reduce the installation cost,
- Eliminate the need for any pre-organization and pre-planning,
- Increase the flexibility of arrangement, and
- Promote self-organization and fault tolerance. After deployment, topology changes are due to change in sensor nodes' position,
- Reach ability (due to jamming, noise, moving obstacles, etc.),
- Available energy,
- malfunctioning

It is also possible to have sensor networks with highly mobile nodes. Besides, sensor nodes and the network experience varying task dynamics, and they may be a target for deliberate jamming. Therefore,

Sensor network topologies are prone to frequent changes after deployment. Additional sensor nodes can be re-deployed at any time to replace the malfunctioning nodes or due to changes in task dynamics. Addition of new nodes poses a need to re-organize the network. Coping with frequent topology changes in an ad hoc network that has myriads of nodes and very stringent power consumption constraints requires special routing protocols.

ENVIRONMENT

Sensor nodes are densely deployed either very close or directly inside the phenomenon to be observed.

Therefore, they usually work unattended in remote geographic areas. They may be working

- In busy intersections,
- In the interior of a large machinery,
- At the bottom of an ocean,
- Inside a twister,
- On the surface of an ocean during a tornado,
- In a biologically or chemically contaminated field,
- In a battlefield beyond the enemy lines,
- In a home or a large building,
- In a large warehouse,
- attached to animals,
- attached to fast moving vehicles, and
- in a drain or river moving with current.

This list gives us an idea about under which conditions sensor nodes are expected to work. They work under high pressure in the bottom of an ocean, in harsh environments such as a debrisor a battlefield, under extreme heat and cold such as in the nozzle of an aircraft engine or in arctic regions, and in an extremely noisy environment such as under intentional jamming.

POWER CONSUMPTION

The wireless sensor node, being a micro-electronic device, can only be equipped with a limited power source (<0.5 Ah, 1.2 V). In some application scenarios, replenishment of power resources might be impossible. Sensor node lifetime, therefore, shows a strong dependence on battery lifetime. In a multihop ad hoc sensor network, each node plays the dual role of data originator and data router. The disfunctioning of few nodes can cause significant topological changes and might require re-routing of packets and re-organization of the network. Hence, power conservation and power management take on additional importance. It is for these reasons that researchers are currently focusing on the design of power-aware protocols and algorithms for sensor networks.

ISSN: 2277-3754
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 1, Issue 4, April 2012

Power consumption can hence be divided into three domains: sensing, communication, and data processing.

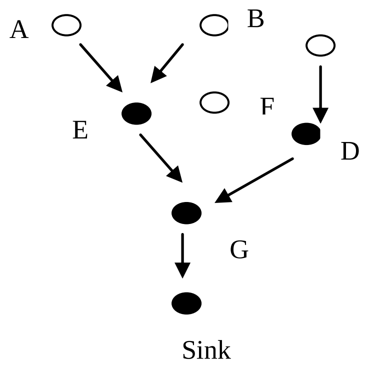

Fig 3 aggregation of data

Communication

Of the three domains, a sensor node expends maximum energy in data communication. This involves both data transmission and reception It is important that in this computation we not only consider the active power but also the start-up power consumption in the transceiver circuitry. The start-up time, being of the order of hundreds of micro-seconds, makes the start-up power non-negligible.

Data processing

Energy expenditure in data processing is much less compared to data communication. local data processing is crucial in minimizing power consumption in a multi-hop sensor network. A sensor node must therefore have built-in computational abilities and be capable of interacting with its surroundings.

III. DATA AGGREGATION [3][19][22][23]

Data gathering is defined as the systematic collection of sensed data from multiple sensors to be eventually transmitted to the base station for processing [8]. Since sensor nodes are energy constrained, it is inefficient for all the sensors to transmit the data directly to the base station. Data generated from neighboring sensors is often redundant and highly correlated. In addition, the amount of data generated in large sensor networks is usually enormous for the base station to process. Hence, we need methods for combining data into high quality information at the sensors or intermediate nodes which can reduce the number of packets transmitted to the base station resulting in conservation of energy and bandwidth. This can be accomplished by data aggregation.[32][34] Data aggregation is defined as the process of aggregating the data from multiple sensors to eliminate redundant transmission and provide fused information to the base station. Data aggregation usually involves the fusion of data from multiple sensors at intermediate nodes and transmission of the aggregated data to the base station (sink). An example illustrating data aggregation in sensor network is shown in fig 3. As it is shown data from nodes A and B gets aggregated at node E .Similarly data from nodes C and D are aggregated at node F and finally at node G resultant aggregation occurs and the result is transmitted to the sink.

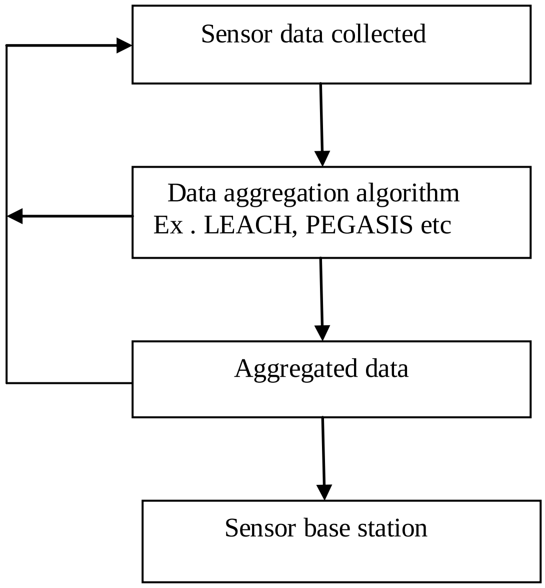

Fig 4 Data aggregation flowchart

IV PERFORMANCE METRICS [22]

The design of SN is a challenge because many influencing factors such as fault tolerance, scalability, production cost, operating environment, network topology, hardware constraints, transmission media, power consumption and others have to be considered. The performance of the network is then measured based on quantifiable parameters called performance metrics.

Network Lifetime: Network lifetime is defined as the number of data aggregation rounds till x % of sensors die where x is specified by the system designer. For instance, in applications where the time that all nodes operate together is vital, lifetime is defined as the number of rounds until the first sensor is drained of its energy.

Data accuracy: The definition of data accuracy depends on the specific application for which the sensor network is designed. For instance, in a target localization problem, the estimate of target location at the sink determines the data accuracy.

Latency: Latency is defined as the delay involved in data transmission, routing and data aggregation. It can be measured as the time delay between the data packets received at the sink and the data generated at the source nodes.

Energy Efficiency: The functionality of the sensor network should be extended as long as possible. In an ideal data aggregation scheme, each sensor should have expended the same amount of energy in each data gathering round. A data aggregation scheme is energy efficient if it maximizes the functionality of the network. If we assume that all sensors are equally important, we should minimize the energy consumption of each sensor. This idea is captured by the network lifetime which quantifies the energy efficiency of the network.

Bandwidth*, **Capacity** and **Throughput**:* These indicate the capacity of data which can be sent over a link within a given time, however since the data size is very small bandwidth rarely matters.

Hop Count: No of hop in communication determine the cost of path, and eventually the energy consumed in the process.

Signal Strength: SNR as an indication for the link quality and the distance between two nodes is helpful to compute and determine the nodes and their reach ability during the communication process.

V. DATA AGGREGATION BASED NETWORKS

1. Flat networks
2. Hierarchical networks

FLAT NETWORKS

Flat networks plays very important role in wireless sensor network, in which each sensor node have a equal battery power and plays the same type of role in a network. In such type of networks, data aggregation has to be done in data centric routing manner, where the sink generally sends a data packet to the sensor nodes, such as, flooding. In the flooding sensors which have data matching the data packet and transmit response data packet back to the sink.

HIERARCHICAL NETWORKS

All the communication and computation burden at the sink in flat network, that's why lot of energy is consumed. In the hierarchical network, In which data aggregation data has to be done at special nodes, with the help of these special node we can reduce the number of number of data packet transmitted to the sink. So with this network improves the energy efficiency of the whole network [4].

Table 1 Hierarchical Network vs. Flat Networks

Hierarchical network	Flat Network
Data aggregation performed by cluster heads or leader node	Data aggregation is performed by different nodes along the multi-hop path
Overhead involved in cluster or chain formation throughout the network	Data aggregation routes are formed only in regions that have data for transmission
Even if one cluster head fails, the network may still be operation	The failure of sink node may result in the breakdown of entire network
Lower latency in involved since sensor nodes perform short rang transmission to the cluster head	Higher latency is involved in the data transmission to the sink via multihop path.
Routing structure is simple but not necessarily optimal	Optimal routing can guaranteed with additional overhead
Node heterogeneity can exploited by assigning high energy nodes as cluster heads	Does not utilize node heterogeneity for improving energy efficiency.

Table 2 protocols for Hierarchical and Flat networks

Protocol Name	Flat Network	Hierarchical Network
LEACH		▪
PEGASIS		▪
HIERARCHICAL PEGASIS		▪
SPIN	▪	
DIRECTED DIFFUSION	▪	
TEEN		▪
APTEEN		▪
ROUMAR ROUTING	▪	
GRADIENT BASED ROUTING	▪	
ENERGY AWARE ROUTING FOR CLUSTER BASED SENSOR NETWORK		▪
CADR(CONSTRAINED ANISOTROPIC DIFFUSION ROUTING)	▪	
ACQIRE(ACTIVE QUERY FORWARDING IN SENSOR NETWORK)	▪	
ENERGY AWARE ROUTING	▪	

VI ARCHITECTURES OF DATA AGGREGATION [22]

Based on various applications and requirements there are several existing architectures for data aggregation.[21][22]

- Centralized
- Decentralized
- Cluster based
- Tree based
- Grid
- Chain

CENTRALIZED ARCHITECTURE

Centralized architecture is very simplest architecture of wireless sensor network. In which we can apply data fusion process. Means each sensor nodes sense a data and transmit to the one central node, called central processor fusion node .this central processor fuse the reports collected by all sensor nodes. In this architecture central node have a responsibility of whole network. The basic advantage of this architecture is it can be easily detected erroneous report of information which is taken by the entire wireless sensor network. The disadvantage is that inflexible to sensor changes and the workload is concerned at a single point.

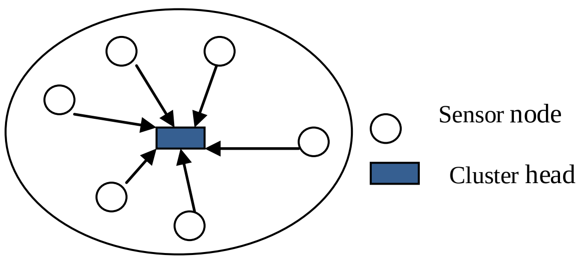

Fig 5 Centralized architecture

ISSN: 2277-3754
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 1, Issue 4, April 2012

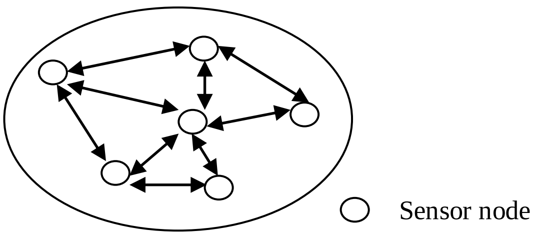

Fig 6 Decentralized architecture

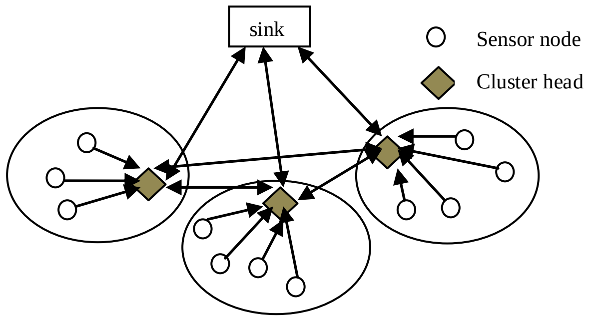

Fig 8 Cluster based architecture

DECENTRALIZED ARCHITECTURE

The decentralized architecture of wireless sensor network, there is no single centralized node that makes decisions on behalf of all the sensor nodes. Data fusion occurs locally at each node on the basis of local observations and the information obtained from neighbouring nodes. in which all sensor nodes are connected to each other on the observation .The advantage of this architecture are scalable and tolerant to the addition or loss of sensing nodes or dynamic changes in the network.

CLUSTER BASED ARCHITECTURE

Wireless sensor network is resource constraint that's why sensor cannot directly transmit data to the base station. In which all regular sensors can send data packet to a cluster head (local aggregator) which aggregates data packet from all the regular sensors in its cluster and sends the concise digest to the base station. With the help of the scheme we save the energy of the sensors. In energy-constrained sensor networks of large size, it is inefficient for sensors to transmit the data directly to the sink. In such scenarios, sensors can transmit data to a local aggregator or cluster head which aggregates data from all the sensors in its cluster and transmits the Concise digest to the sink. There are some issues involved with the process of clustering in a wireless sensor network. First issue is, how many clusters should be formed that could optimize some performance parameter. Second could be how many nodes should be taken in to a single cluster. Third important issue is the selection procedure of cluster-head in a cluster.

.

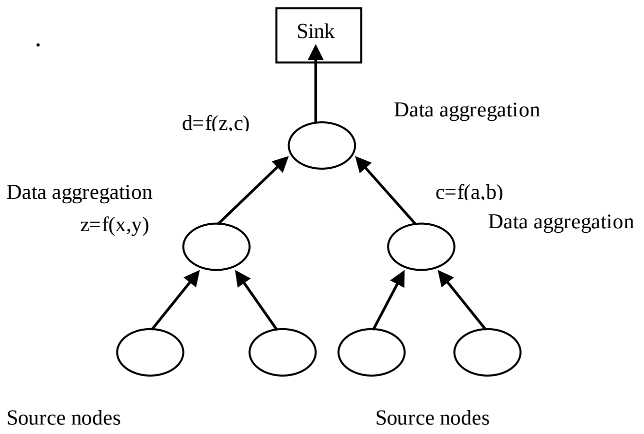

Fig 7 Tree based architecture

Another issue is that user can put some more powerful nodes, in terms of energy, in the network which can act as a cluster-head and other simple node work as cluster-member only

TREE BASED ARCHITECTURE

In the tree-based approach perform aggregation by constructing an aggregation tree, which could be_a minimum spanning tree, rooted at sink and source nodes are considered as leaves. Each node has a parent node to forward its data. Flow of data starts from leaves nodes up to the sink and therein the aggregation done by parent nodes. In which all nodes are organized in form of tree means hierarchical, with the help of intermediate node we can perform data aggregation process and data transmit leaf node root node. One of the main aspects of tree-based networks is the construction of an energy efficient data-aggregation tree.

CHAIN BASED ARCHITECTURE

In which each sensor sends data to the closer neighbor. All sensors are structured into a linear chain for data aggregation. The nodes can form a chain by employing a greedy algorithm or the sink can decide the chain in a centralized manner. In the Greedy chain formation assumes that all sensors have inclusive knowledge of the network. The farthest node from the sink initiates chain formation and, at each step, the closest neighbor of a node is selected as its successor in the chain. In each data-gathering round, a node receives data packet from one of its neighbors, aggregates the data with its own, and sends the aggregates data packet to its other neighbor along the chain. Eventually, the leader node in the are similar to cluster head sends the aggregated data to the base station.

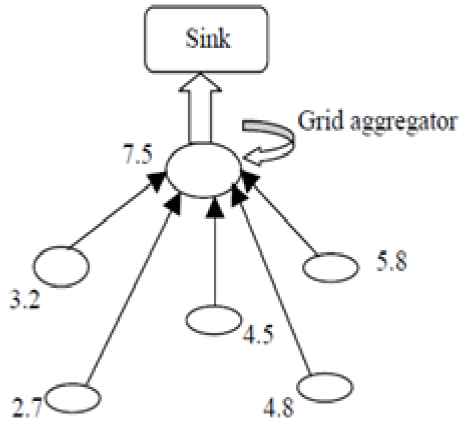

Fig 9 Grid based architecture

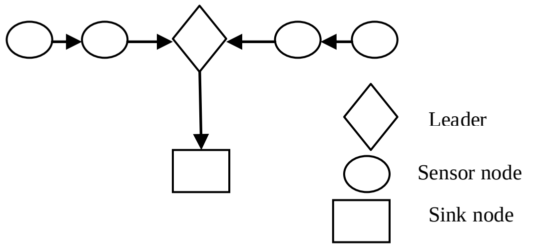

Fig 10 Chain based architecture

GRID BASED ARCHITECTURE

In which a set of sensors is assigned as data aggregators in fixed regions of the sensor network. The sensors in a grid send the data packet directly to the aggregator of that grid. Hence, the sensors within a grid do not communicate with each other. In-network aggregation is similar to grid-based data aggregation with two major differences; each sensor within a grid communicates with its neighboring node. Any node within a grid can assume the role of aggregator node in terms of rounds until the last node dies. This is similar to cluster-based data aggregation in which the cluster heads are fixed. In in-network aggregation, the sensor with the most critical information aggregates the data packets and sends the fused data to the sink. Each sensor transmits its signal strength to its neighbors. If the neighbour has higher signal strength, the sender stops transmitting packets. After getting data packets from all the neighbours, the node that has the maximum signal strength becomes the data aggregator. The in-network aggregation scheme is best suitable for environments where events are highly localized.

VII. ADVANTAGE OF AGGREGATION [12] [21]

Data aggregation offers several advantages.

- With the help of data aggregation process we can enhance the robustness and accuracy of information which is obtained by entire network.
- Certain redundancy exists in the data collected from sensor nodes thus data fusion processing is needed to reduce the redundant information
- Another advantage is those reduces the traffic load and conserve energy of the sensors

VIII. DISADVANTAGE OF AGGREGATION [12] [21]

Data aggregation process has its own disadvantages also.

- The cluster head means data fusion nodes send fuse these data to the base station .this cluster head or fusion node may be attacked by malicious attacker.
- If a cluster head is compromised, then the base station (sink) cannot be ensure the correctness of the fusion data that has been send to it.
- Another drawback in existing systems are several copies of the fusion result may be sent to the base station (sink) by uncompromised nodes .It increase the power consumed at these nodes.

Table 3 Protocols based on different architectures

Protocol	Organization type	Objectives	Characteristics
LEACH	cluster	Network lifetime: number of nodes that are alive, latency	Randomized cluster head rotation, non-uniform energy drainage across different sensors.
HEED	cluster	Lifetime: number of rounds until the first node death	Assumption: Multiple power levels in sensors. Cluster heads are well distributed. Achieves better performance than LEACH
PEGASIS	chain	Lifetime: average energy expended by a node	Global knowledge of the network is required. Considerable energy savings compared to LEACH.
Hierarchical chain based protocols	chain	Energy× delay	Binary chain based scheme is eight times better than LEACH and the three level scheme is 5 times better than PEGASIS.
EADAT	tree	Lifetime: number of alive sensors at the end of simulation time	Sink initiated broadcasting approach. It is not clear how to choose the threshold power (P_{th}) for broadcasting help messages. No comparisons made with other existing aggregation algorithms.
PEDAP-PA	tree	Lifetime: time until the death of last node	Minimum spanning tree based approach. Achieves two times performance improvement compared to LEACH, PEGASIS.

IX. SIMULATORS [26]

Many network details in WSNs are not finalized and standardized. Building a WSNs test bed is very costly. Running real experiments on a test bed is costly and difficulty. Besides, repeatability is largely compromised since many factors affect the experimental results at the same time. It is hard to isolate a single aspect. Moreover, running real experiments are always time consuming. Therefore, WSNs simulation is important for WSNs development. Protocols, schemes, even new ideas can be evaluated in a very large scale. WSNs simulators allow users to isolate different factors by tuning configurable parameters. Consequently, simulation is essential to study WSNs, being the common way to test new applications and protocols in the field. This leads to the recent boom of simulator development.

DIFFERENT SIMULATORS

NS2 is the abbreviation of Network simulator version two, which first been developed by 1989 using as the REAL network simulator. NS-2 is a discrete event network simulator built in Object-Oriented extension of Tool Command Language and C++. NS-2 can support a considerable range of protocols in all layers. For example, the ad-hoc and WSN specific protocols are provided by NS-2. Secondly, the open source model saves the cost of simulation, and online documents allow the users easily to modify and improve the codes. However, this simulator has some limitations. Firstly, people who want to use this simulator need to familiar with writing scripting language and modeling technique; the Tool Command Language is somewhat difficulty to understand and write. Secondly, sometimes using NS-2 is more complex and time-consuming than other simulators to model a desired job. Thirdly, NS-2 provides a poor graphical support, no Graphical User Interface (GUI); the users have to directly face to text commands of the electronic devices. Fourthly, due to the

ISSN: 2277-3754
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 1, Issue 4, April 2012

Continuing changing the code base, the result may not be consistent, or contains bugs.

TOSSIM

It is an emulator specifically designed for WSN running on TinyOS, which is an open source operating system targeting embedded operating system. TOSSIM is a bit-level discrete event network emulator built in Python, a high-level programming language emphasizing code readability, and C++. People can run TOSSIM on Linux Operating Systems or on Cygwin on Windows. TOSSIM also provides open sources and online documents. To the merits, the open source model free online document save the emulation cost. Also, TOSSIM has a GUI, TinyViz, which is very convenience for the user to interact with electronic devices because it provides images instead of text commands. TOSSIM is a very simple but powerful emulator for WSN. Each node can be evaluated under perfect transmission conditions, and using this emulator can capture the hidden terminal problems. As a specific network emulator, TOSSIM can support thousands of nodes simulation. This is a very good feature, because it can more accurately simulate the real world situation. Besides network, TOSSIM can emulate radio models and code executions. this emulator still has some limitations. Firstly, TOSSIM is designed to simulate behaviors and applications of TinyOS, and it is not designed to simulate the performance metrics of other new protocols. Therefore, TOSSIM can not correctly simulate issues of the energy consumption in WSN; people can use PowerTOSSIM, another TinyOS simulator extending the power model to TOSSIM, to estimate the power consumption of each node. Secondly, every node has to run on NesC code, a programming language that is event-driven, component-based and implemented on TinyOS, thus TOSSIM can only emulate the type of homogeneous applications. Thirdly, because TOSSIM is specifically designed for WSN simulation, motes-like nodes are the only thing that TOSSIM can simulate. In sum, TOSSIM as an emulator of WSN contains both advantages and disadvantages.

EMSTAR

EmStar is an emulator specifically designed for WSN built in C, and it was first developed by University of California, Los Angeles. EmStar is a trace-driven emulator running in real-time. People can run this emulator on Linux operating system. This emulator supports to develop WSN application on better hardware sensors. Besides libraries, tools and services, an extension of Linux microkernel is included in EmStar emulator. To the merits, firstly, the modular programming model in EmStar allows the users to run each module separately without sacrificing the reusability of the software. EmStar has a robustness feature that it can mitigate faults among the sensors, and it provides many modes make debug and evaluate much easier. it can not support large number of sensors simulation, and the limited scalability will decrease the reality of simulation, shown in Figure 5. In addition, EmStar is can only run in real time simulation.

Moreover, this emulator can only apply to iPAQ-class sensor nodes and MICA2 motes

OMNET++

OMNeT++ is a discrete event network simulator built in C++. OMNeT++ provides both a noncommercial license, used at academic institutions or non-profit research organizations, and a commercial license, used at "for-profit" environments. This simulator supports module programming model. Users can run OMNeT++ simulator on Linux Operating Systems, Unix-like system and Windows. OMNeT++ is a popular non-specific network simulator, which can be used in both wire and wireless area. Most of frameworks and simulation models in OMNeT++ are open sources. To the merits, firstly, OMNeT++ provides a powerful GUI. This strong GUI makes the tracing and debugging much easier than using other simulators. Although initial OMNeT++ do not support the module library which is specifically used for WSNs simulation, with the consciously contribution of the supporting team, now OMNeT++ has a mobility framework. This simulator can support MAC protocols as well as some localized protocols in WSN. People can use OMNeT++ to simulate channel controls in WSNs. In addition, OMNeT++ can simulate power consumption problems in WSNs. the number of available protocols is not larger enough. In addition, the compatible problem will rise since individual researching groups developed the models separately, this makes the combination of models difficult and programs may have high probability report bugs.

J-SIM

J-Sim is a discrete event network simulator built in Java. This simulator provides GUI library, which facilities users to model or compile the Mathematical Modeling Language, a “text-based language” written to J-Sim models. J-Sim provides open source models and online documents. This simulator is commonly used in physiology and biomedicine areas, but it also can be used in WSN simulation. In addition, J-Sim can simulate real-time processes. To the merits, firstly, models in J-Sim have good reusability and interchangeability, which facilities easily simulation. Secondly, J-Sim contains large number of protocols; this simulator can also support data diffusions, routings and localization simulations in WSNs by detail models in the protocols of J-Sim. J-Sim can simulate radio channels and power consumptions in WSNs. Thirdly, J-Sim provides a GUI library, which can help users to trace and debug programs. The independent platform is easy for users to choose specific components to solve the individual problem. Fourth, comparing with NS-2, J-Sim can simulate larger number of sensor nodes, around 500, and J-Sim can save lots of memory sizes. However, this simulator has some limitations. The execution time is much longer than that of NS-2. Because J-Sim was not originally designed to simulate WSNs, the inherently design of J-Sim makes users hardly add new protocols or node components.

AVRORA

Avrora is a simulator specifically designed for WSNs built in Java. . Avrora provides a wide range of tools that can be used in simulating WSNs. Avrora also supports energy consumption simulation. This simulator provides open sources and online documents. However, this simulator has some drawbacks. It does not have GUI. In addition, Avrora can not simulate network management algorithms because it does not provide network communication tools. To the merits, firstly, Avrora is an instruction-level simulator, which removes the gap between TOSSIM and ATEMU. The codes in Avrora run instruction by instruction, which provides faster speed and better scalability. Avrora can support thousands of nodes simulation, and can save much more execution time with similar accuracy. Avrora provides larger scalability than ATEMU does with equivalent accuracy; Avrora provides more accuracy than TOSSIM does with equivalent scales of sensor nodes. Unlike TOSSIM and ATEMU, Avrora is built in Java language, which provides much flexibility. Avrora can simulate different programming code projects, but TOSSIM can only support TinyOS simulation.

X. CONCLUSION

We have presented a precise survey on various aspects of wireless sensor networks and different architectures of data aggregation all of them focus on optimizing important performance measures such as network lifetime, data latency, data accuracy and energy consumption. We have described the main features of wireless sensor networks with required diagrammatic illustrations. We also discussed various simulators available in sensor networks with its merits and demerits.

REFERENCES

[1] C. Intanagonwiwat, R. Govindan, D. Estrin, Directed diffusion: a scalable and robust communication paradigm for sensor networks, Proceedings of the ACM Mobi- Com'00, Boston, MA, 2000, pp. 56–67.

[2] Wireless sensor networks: a survey I.F. Akyildiz, W. Su*, Y. Sankarasubramaniam, E. Cayirci Broadband and Wireless Networking Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA.

[3] Data Aggregation in Wireless Sensor Network Nandini. S. Patil, Prof. P. R. Patil B.V. Bhoomaraddi College of Engineering and Technology, Hubli 580031, India, Visvesvaraiya Technological University Belgum-590014, India.

[4] D. Estrin, R. Govindan, J. Heidemann, S. Kumar, Next century challenges: scalable coordination in sensor networks, ACM MobiCom'99, Washingtion, USA, 1999, pp. 263–270.

[5] J. Agre, L. Clare, An integrated architecture for cooperative sensing networks, IEEE Computer Magazine (May 2000) 106–108.

[6] M. Bhardwaj, T. Garnett, A.P. Chandrakasan, Upper bounds on the lifetime of sensor networks, IEEE International Conference on Communications ICC'01, Helsinki, Finland, June 2001.

[7] P. Bonnet, J. Gehrke, P. Seshadri, Querying the physical world, IEEE Personal Communications (October 2000) 10–15.

[8] N. Bulusu, D. Estrin, L. Girod, J. Heidemann, and Scalable Coordination for wireless sensor networks: self-configuring localization systems, International Symposium on Communication Theory and Applications (ISCTA 2001), Ambleside, UK, July 2001.

[9] A. Cerpa, D. Estrin, ASCENT: adaptive self-configuring Sensor networks topologies, UCLA Computer Science Department Technical Report UCLA/CSDTR-01-0009, May 2001.

[10] A. Cerpa, J. Elson, M. Hamilton, J. Zhao, and Habitat monitoring: application driver for wireless communications technology, ACM SIGCOMM'2000, Costa Rica, April 2001.

[11] S. Cho, A. Chandrakasan, Energy-efficient protocols for Low duty cycle wireless micro sensor, Proceedings of the 33rd Annual Hawaii International Conference on System Sciences, Maui, HI Vol. 2 (2000), p. 10.

[12] C. Intanagonwiwat, R. Govindan, D. Estrin, Directed diffusion: a scalable and robust communication paradigm for sensor networks, Proceedings of the ACM Mobi- Com'00, Boston, MA, 2000, pp. 56–67.

[13] C. Jaikaeo, C. Srisathapornphat, C. Shen, Diagnosis of sensor networks, IEEE International Conference on Communications ICC'01, Helsinki, Finland, June 2001.

[14] B. Warneke, B. Liebowitz, K.S.J. Pister, Smart dust: communicating with a cubic-millimeter computer, IEEE Computer (January 2001) 2–9.

[15] http://www.fao.org/sd/EIdirect/EIre0074.htm.

[16] J.M. Kahn, R.H. Katz, K.S.J. Pister, Next century challenges: mobile networking for smart dust, Proceedings of the ACM MobiCom'99, Washington, USA, 1999, pp. 271–278.

[17] N. Noury, T. Herve, V. Rialle, G. Virone, E. Mercier, G. Morey, A. Moro, T. Porcheron, Monitoring behavior in home using a smart fall sensor, IEEE-EMBS Special Topic Conference on Micro technologies in Medicine and Biology. October 2000, pp. 607–610.

[18] E.M. Petriu, N.D. Georganas, D.C. Petriu, D. Makrakis, V.Z. Groza, Sensor-based information appliances, IEEE Instrumentation and Measurement Magazine (December 2000) 31–35.

[19] D. Nadig, S.S. Iyengar, A new architecture for distributed sensor integration, Proceedings of IEEE Southeastcon'93, Charlotte, NC, and April 1993.

[20] C. Shen, C. Srisathapornphat, C. Jaikaeo, Sensor information networking architecture and applications, IEEE Personal Communications, August 2001, pp. 52–59.

[21] G. Hoblos, M. Staroswiecki, A. Aitouche, Optimal design of fault tolerant sensor networks, IEEE International Conference on Control Applications, Anchorage, AK, September 2000, pp. 467–472.

[22] Architectural Based Data Aggregation Techniques in Wireless Sensor Network: A Comparative Study Kiran Maraiya#, Kamal Kant*, Nitin Gupta# Department of Computer Science and Engineering National Institute of Technology Hamirpur, Hamirpur (H.P.) INDIA , Department of Computer Science and Engineering ASET, Amity University, Noida (U.P) ,India.

[23] Data aggregation techniques in sensor networks: A survey Ramesh Rajagopalan Syracuse University, Department of Electrical Engineering and Computer Science, rarajago@syr.edu Pramod K. VarshneySyracuse University, Department of Electrical Engineering and Computer Science, varshney@syr.edu.

ISSN: 2277-3754
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 1, Issue 4, April 2012

[24] Ismail H. Kasimoglui Ian .F. Akyildiz. Wireless sensor and actor: research challenges. *(Elsevier) Journal*, 2(38):351{367, 2004.

[25] J. Rabaey, J. Ammer, J.L. da Silva Jr., D. Patel, Pico- Radio: ad-hoc wireless networking of ubiquitous low energy sensor/monitor nodes, Proceedings of the IEEE Computer Society Annual Workshop on VLSI (WVLSI'00), Orlanda, Florida, April 2000, pp. 9–12.

[26] Fei Yu "a survey of wireless sensor network simulation tools"feiyu8643@gmail.com.

AUTHOR BIOGRAPHY

Sushruta Mishra received his B.Tech degree in 2009 from BPUT, Rourkela, Orissa, India. Currently he is pursuing his M.Tech degree from Department of CSE, IIIT Bhubaneswar, Odisha, India. His area of research is Wireless Sensor Networks.
E-mail: mishra.sushruta@iiit-bh.ac.in

Hiren Thakkar received his B.E degree (Computer Engineering) in 2009 from Sardar Patel University, New V V nagar, Gujarat, India. Currently he is pursuing his M.Tech degree from Department of CSE, IIIT Bhubaneswar, Odisha, India. His area of research is Wireless Sensor Networks.
E-mail: hiren.iiit@iiit-bh.ac.in