IJRSET JANUARY 2021 Volume 8, Issue 1 Pages: 46-53

ISSN 2394-739X

International Journal for Research in Science Engineering and Technology

A REVIEW ON DATA AGGREGATION IN WIRELESS SENSOR NETWORK

[1] A. Vimalathithan, [2] Dr. A. Suresh
[1] Ph.D Research Scholar (PT), [2] Head, Dept. of Computer Science
[1]Periyar University, Salem-11, [2]Sona College of Arts and Science, Salem-5.

ABSTRACT- Wireless sensor networks (WSN) include sensor nodes. These networks have colossal application in common environmental factors checking, calamity the executives, security and military, etc.This paper presents an expansive outline of total procedures that can be used in circled approach to improve lifetime and energy protection of wireless sensor networks. Data total is unbelievably fundamental method in wireless sensor networks. With the help of data total, we decline the energy utilization by discarding abundance. In this paper investigates about data conglomeration and its distinctive energy-proficient method used for data accumulation in Wireless sensor organization.

Keywords: [Data Aggregation, Energy Harvesting, Cluster, Disaster Management, Wireless Sensor Networks.]

1. INTRODUCTION

Wireless sensor network (WSN) infers a get-together of spatially scattered and gave sensors for checking and recording the conditions of being of the climate and coordinating the amassed information at a focal area. WSNs measure ecological conditions like temperature, sound, contamination levels, dampness, wind, etc. These look like wireless uniquely named networks as in they depend upon wireless connectivity and unconstrained improvement of networks with the target that sensor information can be dispatched wirelessly.

WSNs are spatially spread autonomous sensors to screen physical or natural conditions, for instance, temperature, sound, pressure, etc., and to pleasingly go their information through the network to a principal area. The more present-day networks are bi-directional, both social gathering information from circumnavigated sensors and drawing in charge of sensor activity. The improvement of wireless sensor networks was persuaded by military applications, for example, combat zone observation; today such networks are utilized in different mechanical and customer applications, for example, modern cycle checking and control, machine wellbeing checking, etc.

In WSN data aggregation is a compelling method to save the restricted assets. The principal objective of data aggregation algorithms is to assemble and total data in an energy effective way so that network lifetime is improved. Wireless sensor networks have restricted computational force and restricted memory and battery power, this prompt expanded unpredictability for application engineers and regularly brings about

applications that are firmly combined with network protocols. In this paper, a data aggregation structure on wireless sensor networks is introduced and a study on different energy-effective calculation for data aggregation. The system fills in as a middleware for collecting data estimated by number of nodes inside a network.

1.1 DATA AGGREGATION

The data aggregation is a procedure used to tackle the collapse and cover issues in data centric routing. Data coming from multiple sensor nodes are amassed as though they are about a similar property of the phenomenon when they arrive at the equivalent routing hub in transit back to the sink. Data aggregation is a generally utilized strategy in wireless sensor networks. The security issues, data confidentiality and integrity, in data aggregation become essential when the sensor network is conveyed in a threatening climate. Data aggregation is a cycle of accumulating the sensor data utilizing aggregation draws near. The security issues, data mystery and uprightness, in data aggregation become fundamental when the sensor network is passed on in an undermining climate. Data aggregation is a pattern of accumulating the sensor data using aggregation moves close. The general data aggregation algorithm fills in as shown in the under figure1.

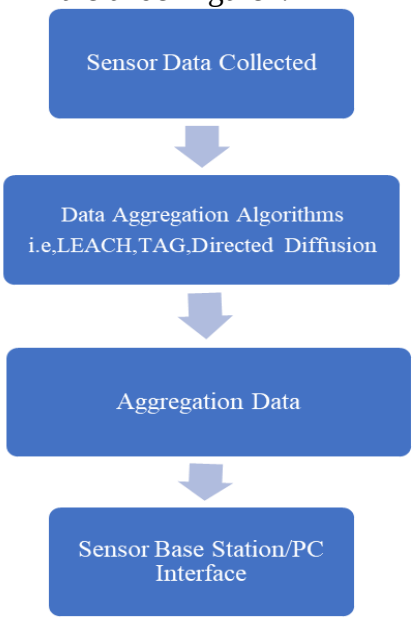

Figure 1. Architecture of Data Aggregation

1.2 Data Aggregation Based Network

i. Flat Networks: In flat networks, every sensor node assumes a similar part and is outfitted with roughly a similar battery power. In such networks, data collection is cultivated by data driven directing where the sink for the most part communicates a question message to the sensors, for instance, by means of flooding and sensors which have data coordinating the inquiry send reaction messages back to the sink. The decision of a specific correspondence protocol relies upon the particular application close by.

ii. Diffusion: Directed diffusion (DD) might be a well-known data total worldview for wireless gadget networks. It's an information driven and application mindful worldview, inside the feeling that everybody data produced by sensor nodes is called by property estimation sets. Such a plan joins the data returning from entirely unexpected sources on the way to the sink by dispensing with repetition and limiting the measure of transmissions. During this implies, it saves the energy consumption and will build the organization life expectancy of WSNs. during this topic typically base station broadcast the message to the intrigued supply hub. Along these lines each hub gets interest. These interest's layout the quality worth like name of article. Each hub gets the interest will reserve it for some time in the future. since the interest is communicated by the organization bounce by jump, angle square measure arrangements to draw data fulfilling the question toward the mentioning hub. An angle might be an answer connect to the closer from that the interest was gotten.

iii. Cluster-Based Networks for data aggregation: This Wireless sensor network is asset requirement that is the reason sensor can't straightforwardly communicate data to the base station. In which all normal sensors can send data packet to a cluster head (nearby aggregator) which aggregates data packet from all the ordinary sensors in its cluster and sends the succinct summary to the base station. With the assistance of the plan creator

save the energy of the sensors. Filter Low energy adaptive clustering has been proposed to arrange a sensor network into a bunch of clusters so the energy consumption can be occasion disseminated among all the sensor nodes.

iv. Chain –Based Networks for Data Aggregation: In which every sensor sends data to the nearer neighbor. Force Efficient Data-Gathering Protocol for Sensor Information Systems PEGASIS) is kind of chain-based data aggregation. In PEGASIS, all sensors are organized into a straight in for data aggregation. The nodes can shape a chain by utilizing a voracious algorithm or the sink can choose the chain in a concentrated way. In the Greedy chain development accepts that all sensors have comprehensive information on the organization. The farthest node from the sink starts chain development and, at each progression, the nearest neighbor of a node is chosen as its replacement in the chain. In every data-assembling cycle, a node gets data bundle from one of its neighbors, totals the data with its own, and sends the totals data parcel to its other neighbor along the chain. In the long run, the pioneer node in that is like bunch head sends the accumulated data to the base station.

1.3 Energy Harvesting

Energy harvesting alludes to interface energy from the climate or other energy sources (body heat, foot strike) and convert it to electrical energy. The related electrical energy controls the sensor nodes. If the reaped energy source is tremendous and on occasion/reliably available, a sensor hub can be fueled interminably. In this, all energy sources ought to be manhandled to remove the available energy; among the others, the sunlight based one is regularly the awesome open-airapplications for the amazing thickness give and exploitable all through sun-based cells. Figure 2 shows the energy harvesting.

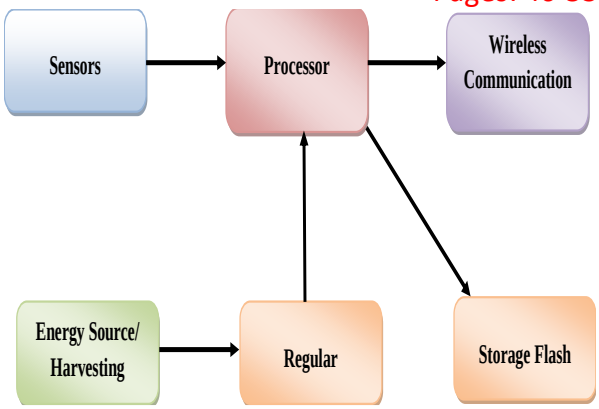

Figure2.Energy Harvesting

1.4 WSN Routing Design Issues

A bunch of unmistakable factors should be viewed as when planning WSN routing protocols. They rely upon routing plan, network engineering and hub attributes. The key ones are clarified underneath:

i. Network architecture: It essentially affects routing to find and set up course that are utilized to advance packets. Data packets are steered to more significant levels of progressive system, for example, bunch heads when WSN is various leveled, though they are straightforwardly/by implication sent to the sink in level WSN.

ii. Node capabilities: Nodes capability and usefulness impacts routing plan and execution. For instance, source nodes may advance data bundles to the middle nodes, for example, bunch heads which can act in-network data aggregation.

iii. MAC protocol design:MAC protocols influence the directing presentation as they are answerable for wireless link availability. For instance, the link availability and therefore correspondence network may be affected by MAC protocols. Also, energy preservation can be improved in WSN if MAC protocols dispense with inactive listening energy utilization. They permit hubs to awaken when they need to send or get network bundles and afterward rest on the off chance that they don't have anything to do.

iv. Data delivery model: Routing is affected by the data conveyance models that are constant, question driven, and occasion driven or half and half. For instance, single way routing isn't suggested in nonstop data conveyance as transmitting all the parcels constantly through a solitary/same way can deplete the energy of hubs being utilized (bottleneck). Attributable to this, multi-way or various leveled routing protocols are used to advance the network traffic through a bunch of variation ways or halfway hubs that can take out repetitive data.

v. Node placement: Node position is established in the network applications and additionally the customer necessities and can possibly impact be directing availability and inclusion. Sensor nodes can be put in two plans: deterministic and non-deterministic. In the previous, the sensor nodes get physically positioned and pre-decided courses are utilized to report the network traffic, while they are haphazardly dissipated and the ways are dynamically framed in last mentioned.

2. DATAAGGREGATION TECHNIQUES

In average WSNs, sensor nodes are typically asset compelled and battery-restricted. To save assets and energy, data should be totaled to try not to overwhelm measures of traffic in the network. There has been broad work on data aggregation plans in sensor networks. The point of data aggregation is that kills repetitive data transmission and upgrades the lifetime of energy in WSNs. Data aggregation is the cycle of one or a few sensors at that point gathers the location result from other sensor. The gathered data should be handled by sensor to diminish transmission burden before they are communicating to the BS or sink.

The WSN comprise of three sorts of nodes: Simple regular sensor nodes, aggregator hub and querier. Regular sensor nodes sense data packet from the climate and ship off the aggregator nodes. These aggregator nodes gather data from multiple sensor nodes of the network.

Totals the data packet utilizing some collection capacities like entirety, normal, check, max, and min and afterward sends totals result to upper aggregator hub or the querier hub that creates the question. It very well may be the BS or at times an outer client who has consent to interface with the network. Data transmission between sensor nodes, aggregators and the querier burns-through a great deal of energy in WSNs.

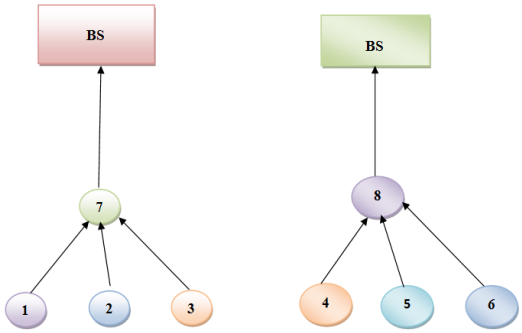

Figure 3. Data Aggregation Model and Non-data Aggregation Model

2.1 Data Analysis

Ang et al. utilized analytic ways to deal with ascertain the energy consumption of the nodes and the ideal number of clusters for two mobile data assortment models: MULE (multichip model) and SENMA (single jump model) explored different avenues regarding diverse huge scope network situations in a short pass of time. Likewise, and to limit the energy consumption in huge scope WSNs, creators proposed multi group models for deciding the ideal number of clusters. Saneja and Rani mean to address the versatility and the connection restrictions of enormous data in wireless sensor networks for the discovery of flawed sensors.

For this, the creators proposed an exception adaptable to big data location approach dependent on relationship and dynamic SMO (Sequential Minimal Optimization) regression. In view of Hadoop Map Reduce system, the proposed approach intends to and out the emphatically related credits and to proficiently distinguish the anomalous nodes, lessening

then the quantity of bogus alerts. Ejaz et al. reviewed the new proposed outline works identified with big data analytics for IoT (Internet of Things).The works mainly intend to conquer the difficulties of breaking down enormous measure of data. The creators additionally investigated the big IoT-produced data preparing and analytics stages, and considered the IoT big data and analytics necessities. In light of significant boundaries, the creator's taxonomies the IoT big data and analytics arrangements.

2.2Correctness/Reliability Enhancement

WSN nodes reliably discover their applications in unforgiving and frail conditions where nodes are powerless against disappointments and detected data may suffer transmission misfortunes, inconsistencies and mistakes. This requires joining of a system which can recognize and dispose of misleading qualities to improve rightness of results. Particular aggregation techniques are arranged that incorporates rounds of correspondence between adjoining nodes for sharing their perceptions to show up at concurrence on discovery of the event of premium.

Each node utilizes spatial redundancy present in the network and updates its decision reliant on dominant part projects a ballot. As can be handily envisioned, such methods bring about precision energy protection tradeoff and proportion less energy when higher exactness is required. Similar example can be seen for aggregation idleness. Another sort of algorithms in this class tries to improve constancy by using trust model. Each node consigns trust rating to its neighbor’s dependent on past records and in like manner decides to consider or disregard neighbor's perceptions.

2.3 Multilevel Clustering

The clustering of the coarser issue is arranged back level-by level to get a predominant clustering of the main issue by refining the center particular clustering got at various levels. A benchmark using various informational indexes accumulated from a collection of spaces is used to investigate the adequacy of the hierarchical methodology against its single-level counterpart.

3. THE PCCDC PROTOCOL

In this segment, the creator talks about the itemized framework of our proposed PCCDC protocol for WSNs. PCCDC is a protocol that considers portable hubs with homogeneous nature which are coordinated dynamically into clusters. It is equipped for distinguishing and lightening clog at intra-and intercluster level utilizing straight and parallel criticism, individually. It additionally utilizes a proper line model in every hub to stay away from focused on bundle drop or postponement during clog. The suspicions made for the protocol configuration are introduced underneath.

(i)Nodes are conveyed haphazardly in the sensor field with an equivalent arrangement of energy level.

(ii)Nodes are permitted to move around the sensing field at various speed to give fundamental inclusion and availability.

(iii)Each cluster ought to have just a single CH. Subsequently the quantity of CHs consistently decides the quantity of clusters.

(iv)CMs are equipped for changing their transmission power to come to their individual CH during a particular round.

(v)BS is constantly arranged external the sensing field which is versatile in nature and has the most noteworthy energy when contrasted and the wide range of various nodes in the WSN.

3.1 Framework of PCCDC

Static WSN gives fractional inclusion of the sensing field. To give total inclusion, the proposed PCCDC protocol utilizes versatile nodes which move around the sensing field to notice explicit occasion that fulfills certain prespecified conditions. PCCDC is an application explicit protocol which thinks about two application boundaries, flooding (flood of water) and temperature, which brings about a multiclass traffic. Flooding is considered to happen when the temperature transcends the edge esteem continually in ice

covered mountains. Since flooding happens quickly, its traffic ought to be accounted for promptly to the base station. The packets for this situation ought to be with less deferral. This is called ongoing packets. The continuous packets require particular help when contrasted with the other. Then again sensing temperature and answering to the base station happen occasionally and are alluded to as non-ongoing traffic. In the event of continuous traffic event, traffic streams meet with each other and locale around the crossing point makes a blockage area of interest.

3.2 PCCDC Mechanism

In this subsection, features of PCCDC protocol are given in agreement with blockage recognition and control. It additionally gives an application outline, clustering philosophy, queuing model, and energy utilization model.

3.2.1 PCCDC: An Application Specific Protocol. PCCDC is intended to help and detect two distinctive application boundaries. Every node is furnished with two kinds of sensors, one to screen flooding and the other to detect temperature.

3.2.2 PCCDC: Clustering Operation. PCCDC is a dynamic group-based protocol, which is iterative and gives various arrangements of bunches in each round. Each node keeps an update table. The update table holds data identified with blockage of the adjoining nodes inside the group. Every node processes its own weight and floods it to the neighbors which thus update their update table. PCCDC computes the heaviness of node i which is given in the (1) and (2). Think about the accompanying:

$$\mathbf{W_i = \alpha * \Delta_i + \beta * E_{Ri} + \gamma * Q_{Li}}, \qquad (1)$$

Where, $\alpha, \beta,$ and γ are nonnegative weight coefficients. It is equal to 0.33. Consider the following:

$$\Delta_i = |d_i - N|$$

Where di is the degree of the hub i by tallying its neighbors, N is the greatest size of a cluster as far as number of nodes, Δi is the degree contrast for the hub i, ERi is the lingering energy for the hub i, and QLi is the current line length of the hub i.

The nodes which have more weight in a specific district will go about as a CH hub for the current round. Each cluster is administered by a CH, which is liable for gathering data delivered by its CMs and identify clog.

3.2.4. PCCDC: Lingering Energy Computation. The energy required for directing packets relies upon the lingering energy at the connection and at the node. It is the amount of energy at connection and complementary energy accessible at the node. This equal capacity consistently doles out higher qualities to nodes having low lingering energy.

3.2.5. PCCDC: Queue Model and Its Operation. The queue model in every sensor node assumes a conspicuous part in the speedy conveyance of packets. Every sensor node has a queue to hold packets to be communicated. Queue flood happens when a node gets packets with a higher data rate than it can communicate. To beat the present circumstance, a congestion control (Algorithm 1) can either decrease the data rate or specifically drop the packets which have lesser significance or reroute the packets through substitute way, which improves the throughput at the BS.

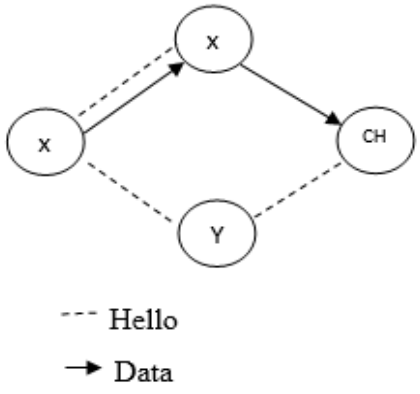

Figure 4: Residual energy computation

(a)Algorithm for queue model in CM

Input: P_{in} -incoming packet

Output: queue in which packet to be stored
Variables:
Q_{TL} -Lower queue threshold
Q_{TH} -Upper queue threshold
Q_L -Current queue length
For each CM Mode
For every packet to queue in
Check the current queue length Q_L against Q_{TL} and Q_{TH}
If ($Q_L < Q_{TL}$)
Queue in both P_L and P_h, Queue act as FCFS
Else If ($Q_L < Q_{TL}$) && ($Q_L < Q_{TH}$)
Queue in P_h and dropP_L randomly
Else
Drop both P_L and P_h
Endif

(b) Algorithm for queue model in CH

For each CM Mode
Check the current queue length Q_L against Q_{TL} and Q_{TH}
If ($Q_L < Q_{TL}$)
Queue in P_L at tail selectively,
Queue in $P_h$at the head
Queue act as Priority queue
Else If ($Q_L < Q_{TL}$) && ($Q_L < Q_{TH}$)
Queue in P_h selectively and dropP_L randomly
Else
Drop both P_L and P_h
Endif End

Algorithm for processing packets in CM and CH nodes

The proposed protocol decreases the energy consumption and yet brings about a slight time delay because of dynamic clustering during crisis circumstances. To address this time delay, it is proposed to deal with the data transmission period of PCCDC in two unique circumstances. During ordinary circumstance CHs alone partake in data transmission to BS and under basic circumstance data can be sent through a base number of nodes which are encompassing the occasion. The proposed protocol could generally be utilized in recognizing cataclysmic events, for example, floods that happen in stream starting from ice caped mountains and in this way forestalling life misfortune. Sensor nodes utilize wireless correspondence to team up to satisfy their assignments like detailing occasion to a concentrated area called a base station (BS) for additional preparing to get important and needful data. The BS goes about as a door between sensor hub and end client and permits a client to beneficially detect and screen from a good way.

CONCLUSION

In this paper we have presented an unmistakable review of in network aggregation strategies for wireless sensor networks. One of the principal plan plots for sensor network structures is to expand network lifetime by achieving energy proficiency. The PCCDC protocol gives total inclusion and network utilizing versatile nodes. To reduce energy consumption, nodes are coordinated into groups. PCCDC bunches have two stages, in particular, arrangement and data transmission stage. PCCDC allots need to continuous packets; the quantity of ongoing packets getting dropped during congestion is reduced extensively on the grounds that PCCDC handles congestion at both intra-and intercluster level independently. A suitable queue model in CM and CH additionally dodges bundle drop because of congestion.

REFERENCE

[1]. Anindita Ray, Debashis De Department of Computer Science & Engineering, West Bengal University of TechnologyBF-142, Sector 1, Salt Lake City, Kolkata 700064, West Bengal, India.
[2]. Saeid Pourroostaei Ardakani Allameh Tabataba'i University, Tehran, Iran International Journal of Computer Networks & Communications (IJCNC) Vol.9, No.2, March 2017.
[3]. PREETI SETHI, DIMPLE JUNEJA, and NARESH CHAUHAN. A mobile agent-based event drivenroute discovery protocol in wireless sensor network: Aerdp. International Journal of EngineeringScience and Technology (IJEST), 3(12):8422–9, December 2011.
[4]. Xiangbin Zhu and Wenjuan Zhang. A mobile agent-based clustering data fusion

algorithm in WSN.International Journal of Electrical and Computer Engineering, 5(5):227–280, 2010.
[5]. Sukhdeep Kaur and Abhinav Hans , 2DAV Institute of Engineering and Technology Jalandhar, Punjab.International Journal of Future Generation Communication and NetworkingVol.9,No.8 (2016), pp. 179-186 .
[6]. Y. Wu and W. Liu, “Routing protocol based on genetic algorithm for energy harvesting-wireless sensor networks”, IET Wireless Sensor Systems 3.2, (2013), pp. 112-118.
[7]. D. Mantri, N. R. Prasad and R. Prasad, “Grouping of clusters for efficient data aggregation (GCEDA) in wireless sensor network”, Advance Computing Conference (IACC), 2013 IEEE 3rd International, IEEE, (2013).
[8]. B. S. Mathapati, S. R. Patil and V. D. Mytri, “Energy Efficient Reliable Data Aggregation Technique for Wireless Sensor Networks”, Computing Sciences (ICCS), 2012 International Conference on. IEEE, (2012).
[9]. Sabrina Boubiche, DjallelEddineBoubiche, Azeddine Bilami1,AndHomero Toral-Cruz21Lastic Laboratory, Department of Computer Science, University of Batna, Batna 05000, Algeria2Department of Sciences and Engineering, University of Quintana Roo, Chetumal 77019, Mexico 2018.
[10]. U. Sivarajah, M. M. Kamal, Z. Irani, and V.Weerakkody, ``Critical analysis of big data challenges and analytical methods," J. Bus. Res., vol. 70,pp. 263_286, Jan. 2017.
[11]. T. Dr AbdulRazak, R. Rajakumar, and M. Rameeja, ``Improving wireless sensor network performance using bigdata and clustering approach," Int.J. Sci. Res. Publ., vol. 4, pp. 1_7, Aug. 2014.
[12]. G. S. Kunal et al., ``An ef_cient EM-algorithm for big data in wireless sensor network using mobile sink," Int. J. Comput. Sci. Inf. Technol., vol. 7, no. 5, pp. 2201_2205, 2016.
[13]. J. Zhou, Y. Zhang, Y. Jiang, C. L. P. Chen, and L. Chen, ``A distributed k-means clustering algorithm in wireless sensor networks," in Proc. Int. Conf. Inform. Cybern. Comput. Social Syst. (ICCSS), Aug. 2015pp. 26_30. Doddappa Kandakur M.Tech Student, Dept. of CSE SIT, Tumkur, International Journal of Engineering and Techniques - Volume 1 Issue 4, July-Aug 2015.
[14]. D. Wu, B. Yang, and R. Wang, ``Scalable privacy-preserving bigdata aggregation mechanism," Digit. Commun. Netw., vol. 2, no. 3,pp. 122_129, 2016.
[15]. J. Li, S. Guo, Y. Yang, and J. He, ``Data aggregation with principal component analysis in big data wireless sensor networks," in Proc. 12thInt. Conf. Mobile Ad-Hoc Sensor Network., Dec. 2016, pp.45_51.