A Survey of Clustering Algorithms for Wireless Sensor Networks

Diwa Kumari Sinha[1] Varsha Mishra[2]
[1,2]Department of Computer Science & Engineering
[1,2]L.N.C.T. Indore

Abstract— The past few years have witnessed amplified interest in the possible use of wireless sensor networks (WSNs) in a wide range of request and it has develop into a hot investigate area. Based on network structure, routing protocols in WSNs can be divided into two categories: flat routing and hierarchical or clustering routing. Due to an assortment of compensation, clustering is becoming an active branch of routing technology in WSNs. In this paper, we present a complete and fine grained survey on clustering routing protocols future in the literature for WSNs. Clusters create hierarchical WSNs which include resourceful consumption of limited income of sensor nodes and therefore enlarge network lifetime. The objective of this paper is to present a survey on clustering algorithms reported in the literature of WSNs.

Key words: Base Station, Cluster, Cluster Head, Sink, Wireless Sensor Networks

I. INTRODUCTION

A wireless sensor network (WSN) is a computer network consisting of spatially distributed autonomous devices using sensors to cooperatively monitor physical or environmental conditions, such as temperature, sound, vibration, pressure, movement or pollutant, at dissimilar locations. The growth of wireless sensor networks was initially aggravated by military request such as battlefield observation. Though, wireless sensor networks are now used in various civilian submission areas, together with surroundings and habitat monitoring, healthcare request, home automation, and traffic control [1].

In addition to one or more sensors, each node in a sensor network is typically equipped with a radio transceiver or other wireless communications device, a small microcontroller, and an energy source, usually a battery. The size a single sensor node can vary from shoebox-sized nodes down to devices the size of grain of dust. The cost of sensor nodes is similarly variable, ranging from hundreds of dollars to a few cents, depending on the size of the sensor network and the complexity required of individual sensor nodes.

II. APPLICATIONS

Representative applications of WSNs comprise monitoring, tracking, and controlling. Several of the exact requests are habitat monitoring, object tracking, nuclear reactor controlling, fire discovery, traffic monitoring, etc. In a characteristic request, a WSN is scattered in a region anywhere it is meant to collect data during its sensor nodes [2].

A. Area Monitoring

Area monitoring is a typical application of WSNs. In area monitoring, the WSN is deployed over a region where various phenomenon is to be monitored. As an instance, a big quantity of sensor nodes might be deployed over a battleground to identify enemy intrusion instead of using landmines. When the sensors detect the occasion being monitored (heat, pressure, sound, light, electro-magnetic field, vibration, etc), the event requirements to be reported to one of the base stations, which can take suitable action (e.g., send a message on the internet or to a satellite). Depending on the accurate application, different objective functions will require [3].

B. Software

Energy is the scarcest resource of WSN nodes, and it determines the lifetime of WSNs. WSNs is meant to be deployed in large numbers in various environments, including remote and hostile regions, with ad-hoc communications as key. For this reason, algorithms and protocols need to address the following issues:

C. Middleware

There is a required and significant research efforts presently invested in the design of middleware for WSN's. There are different research efforts in increasing middleware for wireless sensor networks.[2] In universal move toward can be confidential into dispersed database, mobile agents, and event-based.

D. Algorithms

WSNs are self-possessed of a big number of sensor nodes; consequently, an algorithm for a WSN is completely a distributed algorithm. In WSNs the scarcest source is energy, and one of the mainly energy-expensive procedures is data broadcast. For this reason, algorithmic investigate in WSN typically focuses on the study and design of energy aware algorithms for data broadcast from the sensor nodes to the bases stations.

III. DESIGN ISSUES OF WIRELESS SENSOR NETWORK

Regard as a wireless network made up of units that execute together capacity and infrastructure. These units are totally self-governing and are able of recording data from sensors. The mobility of these units is very low but the data forwarding strategy is forceful sufficient to be fault broadminded and to allow intermittent mobility between units.

Such a real life sensor network is including of hundreds of sensors. The nodes are battery powered, so the first networking challenge is getting data back with minimal energy disbursement, by choosing energy-efficient pathway and by minimizing the routing overhead. The second challenge is to maintain connectivity in case various pods are stimulated to a dissimilar location or fail to contribute owing to lack of power, although overall mobility is likely to be additional limited than in a network of laptops. The third challenge is that sensor networks can be predictable to grow to various thousands of nodes, so any algorithms used in these networks have to be scalable. To conclude, these networks must use manifold pathway when probable, together for redundancy and to distribute the energy expenses of forwarding packets [4]. These necessities

differentiate ad-hoc wireless sensor networks from mobile ad-hoc networks (MANETs) [5]. A sensor network such as the PODS network moreover differs from several of the wireless sensor networks measured in the literature. Although various data can be combined and summarized, additional data, for instance camera images, should be delivered unchanged to a base station. The PODS network is designed to have manifold base stations if probable. In addition, communiqué is not limited to sending data to base stations: communication among human being sensor nodes might be wanted to allow circulated multiplication between nodes in close geographic nearness, to support infrequent announcement from the base stations to the human being nodes, and for a variety of reasons counting fault-tolerance.

There has been a lot of investigate in wireless routing protocols. Obtainable protocols provide dissimilar tradeoffs between the following enviable individuality: fault tolerance, disseminated multiplication, robustness, scalability, and consistency. Wireless protocols future so far for wireless sensor networks is extremely limited, usually focusing on communiqué to a single base station or on merging sensor data. Even as these protocols are appropriate for their intended reason, in this paper we discover the use of procedure developed for MANETs to provide additional universal communication amongst nodes in a sensor network [6].

A. Low Energy Consumption

Protocols for MANETs are designed for communication between laptops. Still while laptops are battery-powered, their power budget far exceeds that of a node in a wireless sensor network. Such nodes are frequently deployed in remote locations. Whether powered by batteries, solar energy, or various additional techniques, reducing energy expenditure lessens the weight or extends the lifetime of the package and makes the sensor easier to conceal. Every node in a wireless sensor network only wants to record, broadcast, and forward data, unlike a laptop which strength have to execute much additional multifaceted tasks. As a result, the computational engine in a sensor node consumes appreciably less energy than a laptop, and infrastructure have to likewise use less energy [7].

B. Low Mobility

Sensor networks are different from MANETs in an extremely significant method and that is in mobility. A MANET is an additional universal case where the participating laptops can moreover be stationary or move haphazardly with a haphazard speed. As nodes inside a MANET move, they move out of range of their neighbors and therefore are no longer able to converse with the old next nodes and come inside range of new nodes. Therefore the mobility introduces the difficulty of fault tolerance. An ideal routing protocol for MANET must be capable to deliver data packets from resource to purpose still while several of the intermediate nodes progress away from their neighbors range. This complicates the design of the routing protocol as this introduces further routing overhead. In before work, one of the authors related the speed of the movement of the nodes to the packet delivery ratio and routing overhead. The packet delivery ratio worsens as speed is amplified for DSR [8], whereas AODV [6] does not degrade as rapidly when mobility amplifies. Nodes in a sensor network the majority of the time are static and with an occasional breaking of a link as the node run out of its energy or are relocated. Sensor networks require the aptitude to re-configure mechanically in case links disappear or new nodes materialize.

C. Self-Configuring Nature

Ad hoc wireless sensor networks are self-configuring in nature. This can be measured an added attribute to the accessible ad hoc nature of the network. The network is flexible to the changing necessities and is capable to diagnose when a link / sensor node goes down and when it comes up. Present are two major schemes to design a wireless sensor network, the address centric scheme and the data centric scheme. The address centric scheme has been used by different routing protocols such as LAR, GSPR, and DREAM etc. In this scheme we assign IP addresses to every sensor node, simplify the procedure of routing. This concept is comparable to that of normal wired networks. A unique IP address will help the resource sensor node to recognize the sensor node to which data have to be routed. Though in [5] a new concept of data centric model is available which is not address oriented. The mechanism and goal of self-configuration in these networks is dissimilar from those of the address centric scheme.

D. Multipath Desirable

There are a number of different path strategies. One that is very common is shortest path where one copy of the message is in the network at any time [9]. At the other extreme is the flooding based approach where the message is flooded through the whole network area. A good example of this approach is the Multi-path On-demand Routing (MOR) Protocol [10] which is a on-demand, load balancing routing protocol designed for the PODS project at the University of Hawaii at Manoa. MOR may require as little as one network flood to establish necessary routes and its energy efficient and robust in low mobility and low energy networks such as PODS. Broadcasting usually solves the routing in highly mobile conditions but considering our requirement for a general sensor network for PODS this is undesirable. The compromise between these two approaches is a multipath strategy, where data packets are routed through a few distinct paths and successive packets follow different paths whenever possible. This not only provides robustness to the network using multiple paths but also helps in distribution of the energy requirement of the network evenly across the network. In [1] A.Nasipuri et al., prove that the use of multiple paths in DSR can keep correct end to end connections, but they did not study the performance improvement on network load balancing. M.R Perlman et al., demonstrate [10] that multipath routing can balance loads. They propose a diversity injection method to find more node-disjoint paths compared to DSR. However, their work is based on multiple channel networks, which are contention free but may not be available in some applications. [11] applies the multipath strategy to DSR's source routing technique and achieves some scalability under mobile conditions. However the energy distribution component of the multipath strategy has not been adequately explored in the paper.

E. Scalability

An ideal routing protocol for a MANET should be scalable. This means that as the size of the network increases or the number of nodes increases the routing protocol should be able to adapt to the changes and provide consistent performance based on the parameters that we have discussed earlier. [2] describes three methods, which have been used by researchers to provide scalability to a routing protocol for MANETs. The first method uses hierarchy to provide scalability. The second way to provide scalability is caching. The third way to provide scalability is using geographic information. Using hierarchy to provide scalability is the most widely deployed approach to scale routing as the number of destinations increases. Two main strategies used to combine nodes location and hierarchical network structures are the Zone Based Routing and the Dominating Set Routing. Online power-aware routing routing [11] schemes are example of Zone Based Routing and GRID is an example of dominating set routing

IV. LITERATURE SURVEY

In the probabilistic clustering algorithms, each sensor node is assigned with a node ID or some probability value used to determine the initial CH. The probabilities at first assigned to each node often serve as the primary criterion, for nodes present in network, in order to decide individually their election as CHs; however other secondary criteria may also be considered either during Selection process i.e., the residual energy or during the cluster formation process i.e., the communication cost in order to achieve better energy consumption and network lifetime.

Clustering algorithms in this category shows faster convergence in addition to energy efficient network utilization, efficient load balancing and low message overheads and reduced volume of exchanged messages.

A. CACC: Clustering Algorithm based on Cell Combination [12]

In this paper author proposed a clustering algorithm which based on cell mixture for the networks. Sensor nodes are detached compactly and the energy of sensor nodes is forever limited. In this clustering algorithm, the monitoring region is divided into hexagonal cells by bearing in mentality the geographic location information of nodes. Each cluster consists of at least seven hexagon cells. Nodes with the comparable cluster personality form a cluster and the cluster head in every cluster is designated from the central cell of each cluster. The shape of the cells regard as practically circular in order to development channel reuse and energy capability.

B. VAP-E: Energy-Efficient Clustering-Virtual Area Partition [13]

In this authors proposed an energy practical clustering algorithm which based on virtual area partition in assorted networks environment anywhere the maximal transmit power of each node strength be different. Authors found that VAP-E can balance the load between clusters, progress the energy capability of sensor nodes, prolong the lifetime of networks, and development the capability of communications. Authors additionally estimate this algorithm with respect to LEACH and LEACH-E and found that VAP-E can progress the eternalness period and network life time with the comparable duplicate condition.

C. CFL: Clustering for Localization [14]

Authors proposed a clustering algorithm which uses a combined burden function and tries to divide the sensor nodes so that a minimum number of clusters with maximum number of sensor nodes in each cluster may be achieved. The weight functions at each sensor node, which is a arrangement of different parameters counting: permanent energy, number of neighbors and transmit power. Basically CFL clustering algorithm is designed for localization in WSNs. It is unable to work when the allocation of sensor nodes are not good.

D. FoVs: Overlapped Field of View [15]

Authors proposed a clustering algorithm for wireless multimedia sensor networks based on overlapped Field of View (FoV) areas. The main contribution of this algorithm is verdict the connection polygon and computing the overlapped areas to start clusters and terminate cluster membership. For dense networks, overlapping FoVs causes wasting power of the system because of superfluous sensing of the area. The aim of the clustering method is prolonging network lifetime and energy protection.

E. KOCA: K-Hop Overlapping Clustering Algorithm [16]

Authors proposed a clustering algorithm based on K-hop overlapping which is used to overcome the complexity of overlapping multi-hop clustering for WSNs. Goal of KOCA algorithm is produce connected overlapping clusters that cover the entire sensor network with a detailed average overlapping degree. Authors furthermore found that KOCA manufacture regarding equal-sized clusters, which permit also, dispense the load frequently over different clusters. In KOCA, clustering formation conclude in an inflexible time besides of the network size. Under disagreement and severe errors, up to 10 percent, KOCA communiqué overhead is condensed owing to the dropped packets. Author's duplicate results show that clusters are concerning equal in size. This is require achieving load balancing among dissimilar clusters.

F. PEZCA: Power-Efficient Zoning Clustering Algorithm [17]

Authors proposed a Power-Efficient Zoning Clustering Algorithm (PEZCA) which uses two algorithms: traditional LEACH (Low-Energy Adaptive Clustering Hierarchy) and PEGASIS (Power-Efficient Gathering in Sensor Information Systems). In this algorithm, base station observe as at a center of the scenario and the scenario area is divided into various fan shaped regions and the clusters closer to the base station have smaller sizes than personnel additional than missing from the base station. Consequently CHs (cluster heads) nearest to the BS (base station) can preserve further energy for inter-cluster data transmit. PEZCA provide further balance in energy disbursement and life time of network evaluation to LEACH.

G. VoGC: Voting-on-Grid clustering [18]

In this author combined voting method and clustering algorithm, and inhabited new clustering schemes for secure localization of sensor networks. Authors moreover found

that the newly future move toward have very good appearance on localization correctness and the detection rate of malicious beacon signals. In this scheme, malicious inspiration signals are filtered out according to the clustering result of intersection of location reference circles. Authors used a voting-on-grid (VOGC) method instead of conservative clustering algorithms to diminish the computational cost and found that the scheme can present good localization accuracy.

H. PDCH: Pegasis Algorithm Improving Based on Double Cluster Head [19]

Authors proposed an algorithm based on hierarchical chain topology and this algorithm using bottom level cluster head and super level cluster head to get better the load balance. In the hierarchical structure, base station (BS) is the center of a circle. The BS determination predefines the number of levels and each node's distance to BS determined the level which it belongs to. Each node receives the signal from the BS, then according to the signal might to identify the distance to BS. PDCH outperform to PEGASIS algorithm and it is moreover useful for big networks.

I. PEGASIS: Power-Efficient Gathering in Sensor Information System [20]

By this author proposed algorithm PEGASIS that is a chain based protocol provide development over LEACH algorithms. In PEGASIS, every node converse only with a close national and takes turns broadcast to the base station, thus reducing the amount of energy spent per round. Using greedy algorithm, the nodes will be prearranged to form a chain, behind that BS can compute this chain and transmit it to all the sensor nodes. Energy saving in PEGASIS over LEACH takes place by various stages: First, in the local data assembly, the distances that the majority of the sensor nodes broadcast are much less evaluate to broadcast to a cluster-head in LEACH. Second, only one node broadcast to the BS in every round of communiqué. PEGASIS outperforms LEACH by limiting the number of broadcast, abolish the transparency of vibrant.

V. Conclusion

In this paper we had given the survey of different clustering algorithms used in WSNs. The need of clustering in WSNs and method of approach of different algorithms are discussed with their merits and short comings. In wireless sensor networks, the energy limitations of nodes play a crucial role in designing any protocol for implementation. It is proved from literature that clustering is useful in WSNs.

References

[1] A.Nasipuri and S.R Das, "On demand Multipath Routing for Mobile Ad Hoc Networks", Proceedings of the 8th International Conference on Computer Communications and Networks (IC3N), Boston, 1999
[2] B.Karp and HT.Kung, "GPSR: Greedy peimeter stateless routing for wireless networks", Proc. MOBICOM, August 2000
[3] B. Krishnamachari et al., "Modeling Data-Centric Routing in Wireless sensor networks"
[4] Chalermek Intanagonwiwat, R.Govindan and D.Estrin, "Directed Diffusion: A scalable and robust communication paradigm for sensor networks", in proc. ACM MOBICOMM, Boston, MA, 2000
[5] C. Perkins, E.Royer, "Ad hoc on demand distance vector routing", Proc. 2nd IEEE Workshop on Mobile Computer System and Applications, February 1999, pp. 90-100.
[6] D .Estrin, R.Govindan, J.Heidemann, S.Kumar, "Next century challenges: Scalable co-ordination in sensor networks", Proceedings of MOBICOM, 1999, Seattle, 263-270.
[7] D. Johnson, D. Maltz, "Dynamic source routing in ad hoc wireless networks", Mobile computing, Chapter 5, pp. 153-181, Luwer Academic Publishers, 1996.
[8] Edoardo Biagioni, Kent Bridges and B.J.S Chee, "PODS: A remote ecological Micro sensor network ", http://www.botany.hawaii.edu/pods/overview.htm
[9] I. D.Aron., "On the scalability of on-demand routing protocols for mobile ad hoc networks: an analytical study"
[10] I.Stojmenovic and Xu Lin, "Power-aware localized routing in wireless networks", IEEE Int. Parallel and distributed processing symp., Cancun, Mexico, May 1-5, 2000, pp. 371-376.
[11] J.P. Macker and M.S. Corson, "Mobile ad-hoc networking and the IETF", Mobile computing and communications review, 2, 1, 1998, 9-14.
[12] L. Chang-RI, Z.. Yun, "A clustering algorithm based on cell combination for wireless sensor networks" In Second International Workshop on Education Technology and Computer Science, PP.74–77.
[13] R. Wang, L. Guozhi, and C. Zheng "A clustering algorithm based on virtual area partition for heterogeneous wireless sensor networks", In International Conference on Mechatronics and Automation, PP. 372 – 376.
[14] S. Zainalie and M. Yaghmaee "CFL: A clustering algorithm for localization in wireless sensor networks", In International Symposium on Telecommunications, PP. 435–439.
[15] M. Alaei, and J.M. Barcelo-ordinas "Node clustering based on overlapping FOVs for wireless multimedia sensor networks". In Proceedings of the IEEE wireless communication and networking, PP. 1–6.
[16] M. Youssef and M. Younis, "Overlapping Multi-Hop Clustering for Wireless Sensor Networks", IEEE Transactions on Parallel and Distributed Systems, PP. 1844 – 1856.
[17] F. Bai, H. Mu, and J. Sun " Power-efficient zoning clustering algorithm for wireless sensor networks", In Proceedings of the Information Engineering and computer science, PP. 1-4.
[18] W. Yang, and W.T. Zhu "Voting-on-grid clustering for secures localization in wireless sensor networks", In Proceedings of the IEEE International Conference on communication, PP. 1–5.
[19] S. Lindsey,. And C.S. Raghavendra, C.S. "PEGASIS: Power efficient gathering in sensor information systems", In IEEE proceeding on aerospace, PP. 1125-1130.
[20] W. Linping, B. Wu, C. Zhen, and W.Zufeng "Improved algorithm of PEGASIS protocol introducing double cluster heads in wireless sensor network". IEEE

International Conference on Computer, Mechatronics, Control and Electronic Engineering, PP. 148–151.

