CLUSTERING TECHNIQUES IN WIRELESS SENSOR NETWORK:

¹Asst.Prof. Amaresh Choudhury, Dept.of Electrical Engineering ²Asst.Prof. Bishnu Prasad Panda, Dept.of Electrical Engineering ³Asst.Prof. Priteesa Mahapatra, Dept. Of Electronics and Telecommunication

Abstract:

The rapid advancement in technology, particularly in Micro-Electro-Mechanical Systems has facilitated the development of wireless networks of smart sensors [1]. Smart sensor nodes are low power devices subject to tight communication, storage and computation constraint. A variety of sensor nodes can be deployed in huge numbers in order to monitor, detect and report time-critical events such that the urgency of the situation can be evaluated, and efforts are coordinated in a timely manner. Wireless sensor networks (WSNs) have the potential to enable a substantial class of applications like ecological habitat monitoring, structure health monitoring, environment contaminant detection, industrial process control, and military target tracking.

In sensor networks the individual sensor nodes are generally assumed to be static. However, some recent applications of sensor networks (e.g. in medical care and disaster response) make use of mobile sensor nodes where different nodes often have different mobility patterns [2]. Some nodes are highly mobile, while others are primarily stationary. The network topology changes randomly since sensor nodes are free to move arbitrarily with different speeds. The ability of delivering the data to the sink node or the base station decreases under this scenario, meaning that mobility significantly increases data loss [3]. In addition, node mobility has an adverse effect on the network lifetime since the lost data may need to be re-delivered to the sink. This motivates to explore an energy and mobility aware fault-tolerant clustering protocol for sensor networks. Clustering is an important topic for wireless mobile sensor networks because clustering guarantee basic levels of system performance such as throughput and delay. There are many clustering algorithms, which are classified based on their objectives. In a clustering scheme, the mobile nodes in a wireless sensor network are divided into different virtual groups. They are allocated geographically adjacent into the same cluster according to some rules with different behaviors. Mobile nodes may be assigned different functions such as cluster head which is the local coordinator that performs intra-cluster transmission arrangement and data forwarding, cluster gateway which is a non-clusterhead node with inter-cluster links, and Cluster member which is an ordinary non-clusterhead node without any inter-cluster links. Clustering is crucial because of the following reasons:

- The spatial reuse of resources to increase the system capacity,
- Second benefit is in routing where cluster head and cluster gateways form a virtual backbone for inter-cluster rouging, and
- A cluster structure makes a mobile network more stable in the view of each mobile node.

In an efficient localized approach, each node make forwarding decisions based on a neighbourhood local view constructed simply by collecting messages. In a network with mobile sensor nodes, that kind of neighbourhood local view can become outdated and inconsistent. This in turn induces a low coverage problem for efficient broadcasting tasks and a low-delivery ratio problem for efficient routing tasks [4]. This work aims a neighbourhood tracking scheme to guarantee the accuracy of forwarding decisions. Based on historical local information, nodes predict the positions of neighbours when making a forwarding decision, and then construct an updated and consistent neighbourhood local view to help to derive more precise forwarding decisions.

In this work, we consider data delivery ratio, message complexity, and latency (time complexity) as performance metrics to evaluate the performance of our work.



Wireless sensor network (WSNs) is a most popular research area in world now a daze. WSNs can be treated as a special family of wireless ad hoc networks. A WSN is a selforganized network that consists of a large number of low-cost and low powered sensor devices, called sensor nodes, which can be deployed on the ground, in the air, in vehicles, on bodies, under water, and inside buildings. Each sensor node is equipped with a sensing unit, which is used to capture events of interest, and a wireless transceiver, which is used to transform the captured events back to the base station, called sink node. Sensor nodes collaborate with each other to perform tasks of data sensing, data communication, and data processing.

The rapid advancement in technology, particularly in Micro-Electro-Mechanical systems has facilitated the development of smart sensors (e.g., Mica motes from Crossbow, Tmote Sky from Moteiv, the MKII nodes from UCLA, etc.). This made it possible to connect independent sensor nodes together to create a Wireless sensor networks (WSNs) with greater monitoring and target tracking [2–6]. Smart sensor nodes are low power devices subject to tight communication, storage and computation constraint. A variety of sensor nodes can be

deployed in huge numbers in order to monitor, detect and report time-critical events such that the urgency of the situation can be evaluated, and efforts are coordinated in a timely manner. The WSNs have the potential to enable a substantial class of applications [3,7–14]. In military target applications; a WSN can assist in intrusion detection and identification. Sensor nodes can sense and detect the environment to forecast disasters before they occur [7]. Surgical implants of sensors can help to monitor a patient's health in biomedical applications (body sensor networks). Deployment of sensors along the volcanic area can detect the development of earthquakes and eruptions. The development in wireless communication technologies, ad hoc wireless networks have gained worldwide attention in recent years. The great popularity of Internet services makes more people enjoy and depend on the networking applications. However, the Internet is not always available and reliable, and hence it cannot satisfy people's demand for networking communication at anytime and anywhere. MANETs, without any fixed infrastructures, allow mobile terminals to set up a temporary network for instant communication. Hence, MANETs bear great application potential in these scenarios, including disaster and emergency relief, mobile conferencing, sensor dust, battle field communication. Ad-hoc networks are Self configuring network of wireless links connecting mobile nodes. These mobile nodes may be routers or hosts. Its peer to peer or may be peer to remote networks. It has no access points and no fixed infrastructure. It has a greater ease and speed of deployment. It reduces the administrative cost. Sensor networks are different from ad-hoc networks. The number of nodes in a sensor network can be several orders of magnitude higher than the nodes in an ad-hoc network. Sensor nodes are densely deployed. Sensor nodes are limited in power, computational capacities and memory. Sensor nodes are prone to failures. The topology of a sensor network changes frequently. Sensor nodes mainly use broadcast, most ad hoc networks are based on peer to peer communication. Sensing the physical world by embedding large collection of selforganizing micro computers with appropriate sensors attached, forms the next revolutionary jump in information gathering and processing. These sensor nodes can then together form a Wireless Sensor Network (WSN). A WSN can monitor (sense) a region or phenomenon of interest and provide useful information about it by combining measurements (computing) taken by individual sensor nodes and then routed (communication) over the wireless interface to a base station. A base station provides a connection to the wired world where the collected data is processed, analysed and presented to useful applications. Thus by embedding processing and communication within the physical world, WSN can be used as a tool to bridge real and virtual environments.

Nodes in WSNs are prone to failure due to energy depletion, hardware failure, communication link errors, malicious attack, and so on. Unlike the cellular networks and ad hoc networks where energy has no limits in base stations or batteries can be replaced as needed, nodes in sensor networks have very limited energy and their batteries cannot usually be recharged or replaced due to hostile or hazardous environments [2]. So, one important characteristic of sensor networks is the stringent power budget of wireless sensor nodes. Two components of a sensor node, sensing unit and wireless transceiver, usually directly interact with the environment, which is subject to variety of physical, chemical, and biological factors. It results in low reliability of performance of sensor nodes. Even if condition of the hardware is good, the communication between sensor nodes is affected by many factors, such as signal strength, antenna angle, obstacles, weather conditions, Wireless sensor network swears an and interference. exceptional fine-grained interface between the virtual and physical worlds. The clustering algorithm is a kind of key technique used to reduce energy consumption. Many clustering, power management, and data dissemination protocols have been specifically designed for wireless sensor network (WSN) where energy awareness is an essential design issue. Each clustering algorithm is composed of three phases cluster head (CH) selection, the setup phase, and steady state phase. The hot point in these algorithms is the cluster head selection. The focus, however, has been given to the residual energy based clustering protocols which might differ depending on the application and network architecture.



In a clustering scheme the mobile nodes in a MANET are divided into different virtual groups, and they are allocated geographically adjacent into the same cluster according to some rules with different behaviours for nodes included in a cluster from those excluded from the cluster. It can be seen that the nodes are divided into a number of virtual groups (with the dotted lines) based on certain rules. Under a cluster structure, mobile nodes may be assigned a different status or function, such as cluster head, cluster gateway, or cluster member. A cluster head normally serves as a local coordinator for its cluster, performing intra-cluster transmission arrangement, data forwarding, and so on. A cluster gateway is a non-cluster head node with inter-cluster links, so it can access neighbouring clusters and forward information between clusters. A cluster member is usually called an ordinary node, which is a non-cluster head node without any inter-cluster links.



First, a cluster structure facilitates the spatial reuse of resources to increase the system capacity. With the non-

overlapping multicluster structure, two clusters may deploy the same frequency or code set if they are not neighbouring clusters. Also, a cluster can better coordinate its transmission events with the help of a special mobile node, such as a cluster head, residing in it. This can save much resources used for retransmission resulting from reduced transmission collision. The second benefit is in routing, because the set of cluster heads and cluster gateways can normally form a virtual backbone for inter-cluster routing, and thus the generation and spreading of routing information can be restricted in this set of nodes. Last, a cluster structure makes an ad hoc network appear smaller and more stable in the view of each mobile terminal. When a mobile node changes its attaching cluster, only mobile nodes residing in the corresponding clusters need to update the information. Thus, local changes need not be seen and updated by the entire network, and information processed and stored by each mobile node is greatly reduced.

The process of clustering can be visualised as a combination of two phases, i.e., cluster formation and cluster maintenance. The cluster formation phase deals with the logical partition of the mobile nodes into several groups and selection of a set of suitable nodes to act as heads in every group. In mobile ad hoc network, where the topology changes frequently, selection of optimum number of cluster heads is a NP-hard problem [17]. There exists some representative algorithms that use the parameters like node identity number, mobility, battery power, degree of connectivity etc. as the factors to decide its suitability for cluster head [18]. Even some researchers combine multiple node parameters to select these set of routers in an efficient manner. These selected nodes are responsible for routing as well as node management in the mobile network and collectively called as the dominant set in graph theory terminology [19]. The objective of cluster maintenance is to preserve the existing clustering structure as much as possible. In one hop clustering, since every node is directly connected to a cluster head, the mobility of either the member node or the cluster head may drive them away from each other. There exists a bidirectional link between these two nodes till both of them are within their transmission range. When any of them moves away from the other, there occurs a link failure and the member node searches for another new head within its transmission range to get affiliated. This kind of situation is called as re-affiliation to a new head node.





Nodes in Hierarchical Structure

- Factors influencing design of clustering algorithms :
 <u>Fault Tolerance</u> : Fault tolerance is the ability to sustain sensor network functionalities without any interruption due to sensor node failures. The fault tolerance level depends on the application of the sensor networks.
- ✓ <u>Hardware Constrains</u>: The cost of a single node is very important to justify the overall cost of the networks. 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.
- ✓ <u>Sensor Network Topology</u>: It uses Pre-deployment and deployment phase , Post-deployment phase , Re-deployment of additional nodes phase.
- <u>Environment</u>: Busy intersections, Interior of a large machinery, Bottom of an ocean Surface of an ocean during a tornado, Biologically or chemically contaminated field, Battlefield beyond the enemy lines, Home or a large building, Large warehouse, Animals affects clustering schemes.
- Transmission Media : In a multi hop sensor network, communicating nodes are linked by a wireless medium. To enable global operation, the chosen transmission medium must be available worldwide.
 Power Consumption : More power is consumed

during Sensing , Communication , Data processing process .

2.1 Literature Survey and Related Research work

The wireless sensor network (WSN) technology is a key component for ubiquitous computing. A WSN consists of a large number of sensor nodes. Each sensor node senses environmental conditions such as temperature, pressure and light and sends the sensed data to a base station (BS), which is a long way off in general. Since the sensor nodes are powered by limited power batteries, in order to prolong the life time of the network, low energy consumption is important for sensor nodes. In general, radio communication consumes the most amount of energy, which is proportional to the data size and proportional to the square or the fourth power of the distance. In order to reduce the energy consumption, a clustering and data aggregation approach has been extensively studied [7].

In clustering approach, sensor nodes are divided into clusters, and for each cluster, one representative node, which called cluster head (CH), aggregates all the data within the cluster and sends the data to BS. Since only CH nodes need long distance transmission, the other nodes save the energy consumption. In order to manage effectively clusters and CHs, distributed clustering methods have been proposed such as probabilistic selection clustering algorithms and nonprobabilistic selection clustering algorithms. In the category of probabilistic selection clustering algorithms [11-24], a priori probability assigned to each sensor node is used to determine the initial CHs (or some other type random election procedure is scheduled). The probabilities initially assigned to each node often serve as the primary (random) criterion in order for the nodes to decide individually on their election as CHs (in a flexible, uniform, fast and completely distributed way); however other secondary criteria may also be considered either during CH election process (i.e., the residual energy) or during the cluster formation process (i.e., the proximity or the communication cost) in order achieve better energy consumption and network lifetime. Beyond the high energy efficiency (which is facilitated also from the periodic CH reelection scheme usually adopted), the clustering algorithms of this category usually achieve faster execution/convergence times and reduced volume of exchanged messages. (e.g. HEED , LEACH , EEHC). In the category of non-probabilistic clustering algorithms, more specific (deterministic) criteria for CH election and cluster formation are primarily considered, which are mainly based on the nodes' proximity (connectivity, degree, etc.) and on the information received from other closely located nodes. The cluster formation procedure here is mainly based on the communication of nodes with their neighbours (one or multi-hop neighbours) and generally requires more intensive exchange of messages and probably graph traversing in some extent, thus leading time complexity sometimes to worse than probabilistic/random clustering algorithms. On the contrary these algorithms are usually more reliable toward the direction of extracting robust and well-balanced clusters. In addition to node proximity, some algorithms also use a combination of metrics such as the remaining energy, transmission power, mobility, etc. (forming corresponding combined weights) to achieve more generalized goals than single-criterion protocols. In the same category we also address a relatively new and quite challenging class of clustering algorithms for WSNs, namely, the biologically inspired protocols (based on swarm intelligence) which are probably the most promising alternative approaches for clustering in WSNs nowadays.(e.g. Node proximity and Graph based, Weight based algorithms). LEACH, which is the most popular method, guarantees that every nodes evenly become CHs but does not take into account battery level and the interrelationship among nodes[2]. HEED, ACE and ANTCLUST achieve better performance than LEACH by taking into account battery level, communication cost, node density, etc. However, they need additional inter-node communications for determining clusters and CHs.

The past few years have witnessed the rapid advancement in technology that made it possible the potential use of Ad hoc wireless sensor networks in applications such as disaster management, border protection and security surveillance. Ad hoc WSNs are suited for use in situations where an infrastructure is unavailable or to deploy one is not cost effective. Irrespective of their purpose all ad hoc sensor networks are characterized by the requirement for energy efficiency, scalability and fault tolerance. To support scalability, nodes are often grouped into clusters. An important issue in Ad hoc WSNs is that different nodes often have different mobility patterns. Some nodes are highly mobile, while others are primarily stationary. The network topology changes randomly since mobile sensor nodes are free to move arbitrarily with different speeds. To address the scalability issue a hierarchical architecture is constructed, which groups geographically close nodes into 1-hop clusters. One representative node (cluster head) is selected based on relative mobility and residual energy. A node may be attached to different clusters at different times while moving in a hierarchical network which results in frequent path rediscovery each time it changes the point of attachment. Specifically this paper investigates the maximum of cluster lifetime in given mobility environment through both analysis and simulation. Low-Energy Adaptive Clustering Hierarchy (LEACH) [10] forms clusters by using a distributed algorithm, where nodes make autonomous decisions without any centralized control. Initially a node decides to be a cluster head with a probability p and broadcasts its decision. Each non-cluster head node determines its cluster by choosing its cluster head that can be reached using the least communication energy. Two-Level LEACH (TL-LEACH) [3] is a proposed extension to the LEACH algorithm which utilizes two levels of cluster head (primary and secondary) in addition to the other simple sensing nodes. Here the primary cluster head in each cluster communicates with the secondary, and the corresponding secondary communicate with the nodes in their sub-cluster. TL-LEACH might not be effective if the Cluster head is far from the base station. In Energy Efficient Clustering Scheme (EECS) [4], each cluster head candidates broadcast their residual energy to neighbouring candidates. EECS extends LEACH by dynamically resizing clusters based on cluster distance from the base station. In this approach clusters closer to the base station may become congested which may result in early cluster head death. Hybrid Energy Efficient Distributed Clustering (HEED) [6] is a multi-hop clustering algorithm for Wireless Sensor Networks. Cluster heads are chosen based on two important parameters: residual energy and intra-cluster communication cost. Residual energy of each node is used to probabilistically choose the initial set of cluster heads, as commonly done in other clustering schemes. In HEED, Intra cluster communication cost reflects the node degree or nodes proximity to the neighbour and is used by the nodes in deciding to join the cluster. Drawbacks of HEED are not well distributed clusters, no inter-cluster communication, applied to static networks only. In Powerefficient and adaptive clustering hierarchy (PEACH) [7] cluster is formed by using overhearing characteristics of wireless communication to support adaptive multilevel clustering. PEACH is applicable in both location unaware and location-aware sensor networks.

The lowest ID algorithm assigns the cluster head role to nodes with the lowest IDs in their respective neighbourhood. The highest connectivity algorithm selects the nodes with the highest Neighbour set in their respective neighbourhood to become the cluster heads. In mobility inspired algorithms clusters are constructed by selecting low-mobility nodes to serve as cluster heads, because low-mobility nodes are expected to stay in their clusters longer than high-mobility nodes. Chiang et al. [13] proposed the least cluster change concept where re-clustering takes place only when either two cluster heads move into contact or a cluster member has lost contact to its cluster head. There are three metrics for quantifying network stability in a hierarchical architecture like the cluster lifetime, the inter cluster link lifetime and the endto-end path lifetime. In neighbourhood tracking scheme of wireless sensor networks, the problems of low broadcast coverage and low routing delivery ratio caused by outdated and inconsistent local views in existing broadcast protocols. A neighbourhood tracking scheme aimed at attaining high broadcast coverage or delivery ratio by achieving an accurate local view of the neighbourhood of a node. After analysis of all the network mobility models and derived a general prediction scheme to be used under any mobility model, the prediction model may still not be practical in the real world.

In CBR-Mobile, a cross-layer design between MAC and network layers to improve system performance and overcome challenges in packet delivery ratio and energy consumption in mobility environment. Cluster based routing collaborates with hybrid MAC protocol will be able to support sensor nodes mobility and provide better performance. In the proposed CBR-Mobile, each cluster head creates two schedules to achieve the mobility and traffic adaptive capabilities Sensor network-based adaptive navigation systems have enabled mobility in WSN's components including sensor nodes mobility, sink nodes mobility, event mobility, and user (observer) mobility. CBR-Mobile is traffic and mobility adaptive protocol whereby the assigned timeslots to the sensor nodes that have not data to send or have moved out of the cluster, can be reassigned to the new mobile sensor nodes. CBR-Mobile protocol decreases the packet loss in mobility environment by providing fast registration to the disconnected mobile sensor nodes. CBR-mobile performance is evaluated by packet delivery ratio, Energy consumption, Packet delivery delay, Fairness in sharing system bandwidth .CBRmobile is more stable than clustering schedule based(LEACH - mobile) and non-clustering contention based (AODV) protocols when no. of sensor nodes is varied in the network. CBR-mobile achieve higher and more stable packet delivery ratios than LEACH-mobile and AODV protocols when mobile nodes are increased. Mobility becomes a critical issue that must be considered in the design of WSNs with mobile nodes. Previous proposed protocols (i.e. CBR-mobile, LEACH-mobile)have drawbacks regarding successful packet delivery rate which decreases as no. of mobile nodes increases and increases control overhead, consume more power during data transmission. In MBC clustering protocol, a sensor node elects itself as a cluster-head based on its residual energy and mobility. A non-cluster-head node aims at its link stability with a cluster head during clustering according to the estimated connection time. Each non-cluster-head node is allocated a timeslot for data transmission in ascending order in a time division multiple address (TDMA) schedule based on the estimated connection time. It outperforms both the CBR protocol and the LEACHmobile protocol in terms of average energy consumption and average control overhead, and can better adapt to a highly mobile environment. In the steady-state phase, a sensor node transmits its sensed data in its timeslot and broadcasts a joint request message to join in a new cluster and avoid more packet loss when it has lost or is going to lose its connection

with its cluster head.

Real challenge in wireless sensor networks is tied up to the routing of a massive amount of collected data. Extensive research has been dedicated to the energy-efficient routing algorithms for wireless sensor networks, namely the works proposed in [1], [2]. It is largely accepted that the hierarchical routing provides better performance in terms of the lifetime in such kind of networks. In its conception, the hierarchical routing is mainly based on the clustering algorithms. These algorithms have several advantages which includes decreasing significantly the overhead of communication which reduces in turn the energy consumption and the interference between several sensor transmissions [3], grouping of sensors that are close in space allowing to exploit the correlation and to eliminate the redundancy in data reading [4], reducing the routing table of each sensor [5], and increasing the availability of bandwidth [6].

Another important challenge in wireless sensor networks lies in the limitation of residual energy in each sensor. This constraint can be strongly critical in a hostile environment like a toxic area due to the fact that battery cannot be easily replaceable. In this case, the death of one or more nodes may cause partial or complete interruption of the communication. For these reasons, energy efficient algorithms are necessary to improve the network lifetime. On the other hand, given the impact of the network connectivity especially in a mobile environment, it becomes critical to consider this constraint carefully, because it allows to any node to reach other nodes in the network via the multi-hop transmission technique.

Many of such techniques care mostly about node reach ability and route stability, without much concern about critical goals of WSNs such as network longevity and coverage. In proposed work, the issues like the node mobility , hot spots problems by creating clusters of unequal size where cluster closer to the sink node are smaller. Clustering is completely distributed and cluster heads have relatively high average residual energy compared to regular nodes. In this work, a node is assigned to the cluster head role based on Neighbour time and residual energy where Neighbour time is the duration during which two nodes remain in transmission range of each other. Neighbour time is proportional to node's relative velocity to its Neighbour: a larger value means higher stability. The main contribution of this paper is proposing a distributed clustering algorithm for mobile wireless sensor networks. Regardless of the topology changes, our algorithm aims to improve the network stability and to minimize the consumed energy for each mobile sensor. In this context, the CHs are dynamically elected, i.e., according to both the topology change and the remaining energy. Hence, the CH is the sensor having the best capability in terms of energy, connectivity, distance from the BS and speed. This leads to an energy saving at the level of each sensor and at the level of the whole network, and consequently an increase in network lifetime.

III. SYSTEM MODEL

3.1 Mobility Model

To demonstrate the effectiveness of the proposed protocol this

work adopts linear (Random Waypoint). The Random Waypoint Mobility Model includes pause times between changes in direction and/or speed. In this model, a node alternates between the moving and the pausing phases where the pause time is exponentially distributed with mean Tp. A node moves from its current location to a new location by randomly choosing a direction and speed in which it will travel. The new speed and direction are both chosen from [vmax, vmin] and $[0, 2\pi]$ respectively. Where, vmax and vmin are the maximum and minimum node speed of a node. Upon arriving at the destination; the node pauses for an exponentially distributed random time *Ts* before starting another movement.

3.2 Energy Consumption Model

Similar to (*Heinzelman et al.*, 2002), this work assumes a simple model for the radio hardware energy dissipation where the transmitter dissipates energy to run the radio electronics and the power amplifier. The receiver dissipates energy to run the radio electronics. Both the free space (d^2 power loss) and the multipath fading (d^4 power loss) channel models are used, depending on the distance between the transmitter and receiver. The energy spent for transmission of an *r*-bit packet over distance *d* is: $E_{Tx}(r, d) = rEelec + red^{\alpha}$

 $= rEelec + reampd^{4} \text{ iff } d \ge d_{0}$ $= rEelec + refsd^{2} \text{ iff } d < d_{0}$ Where $d_{0} = (\epsilon fs/\epsilon amp)^{V_{2}}$

The electronics energy, *Eelec*, depends on factors such as the digital coding and modulation. The amplifier energy, εfsd^2 or $\varepsilon ampd^4$, depends on the transmission distance and the acceptable bit error rate. To receive this message, the radio expends energy:

 $E_{Rx}(r, d) = rEelec$

Cluster head consumes $E_{DA}(nJ/bit/signal)$ amount of energy for both routine and diagnostic data aggregation.

3.3 Network Model and Assumptions

The system under consideration accommodates n number of nodes out of which some are highly mobile and some are primarily stationary. Each node occupies a position (x_t, y_t) inside of a fixed geographic area $(1 \times 1 \text{ m}^2)$ at time t and are initially uniformly distributed. Every node independently moves and obeys the aforementioned mobility model. The mobility model maintains uniform node spatial distribution over time. Each node can adjust its transmission radius (power level) to any value up to a given maximum level such that any two nodes can communicate directly with each other. Two nodes n $_i$ and n $_j$ are within transmission range r_{tx}, if the Euclidean distance d(n_i , n_j) between n_i and n_j is less than r_{tx} . The topology graph G(t)=(V,E(t)) consists of a set of vertices V representing the nodes of the network and the set E(t) of undirected edges corresponding to communication links between nodes at time t. As suggested by the authors in [17], to have a connected graph with high probability assuming uniform node distribution in a unit square area, it is necessary and sufficient that r_{tx} be $\Theta((\ln n/n)^{\frac{1}{2}})$. Thus, this work assumes that $R = \Theta((\ln n/n)^{\frac{1}{2}})$.



(a)Analysis model for neighbourhood interval



(b) Relative Velocity V_R of nodes S and U

It is important to note that in our model, no assumptions are made about

- 1. homogeneity of node dispersion in the field,
- 2. network density or diameter,
- 3. distribution of energy consumption among sensor nodes,
- 4. proximity of querying observers, or
- 5. node synchronization.

The essential operation in sensor node clustering is to select a set of cluster heads from the set of nodes in the network, and then cluster the remaining nodes with these heads. Cluster heads are responsible for coordination among the nodes within their clusters and aggregation of their data (intra cluster coordination), and communication with each other and/or with external observers on behalf of their clusters (inter cluster communication). Figure depicts an application where sensors periodically transmit information to a remote observer (e.g., a base station). The figure illustrates that clustering can reduce the communication overhead for both single-hop and multi hop networks.

Periodic re clustering can select nodes with higher residual energy to act as cluster heads. Network lifetime is prolonged through :

1. reducing the number of nodes contending for channel access,

2. summarizing information and updates at the cluster heads, and

3. routing through an overlay among cluster heads, which has a relatively small network diameter.

Clustering protocols have been investigated in the context of routing protocols [3], [14], or independent of routing [16], . In this work, we present a general distributed clustering approach that considers a hybrid of energy and communication cost. Based on this approach, we present the Distributed clustering protocol. It has four primary objectives:

1. prolonging network lifetime by distributing energy consumption,

2. terminating the clustering process within a constant number of iterations,

3. minimizing control overhead (to be linear in the number of nodes), and

4. producing well-distributed cluster heads.

This clustering approach does not make assumptions about the distribution of nodes or about node capabilities, e.g.,

location-awareness. The approach only assumes that sensor nodes can control their transmission power level. Finally, a node may fail if its energy resource is depleted, which motivates the need for rotating the server role among all nodes for load balancing.

Consider a set of sensors dispersed in a field. We assume the following properties about the sensor network:

- . The sensor nodes are quasi-stationary. This is typical for sensor network applications.
- . Links are symmetric, i.e., two nodes v1 and v2 can communicate using the same transmission power level.
- . The network serves multiple mobile/stationary observers, which implies that energy consumption is not uniform for all nodes.
- . Nodes are location-unaware, i.e., not equipped with GPS-capable antennae. This justifies why some techniques, such as [10], [23] are inapplicable.
- All nodes have similar capabilities (processing/ communication), and equal significance. This motivates the need for extending the lifetime of every sensor.
- . Nodes are left unattended after deployment. Therefore, battery recharge is not possible. Efficient energy-aware sensor network protocols are thus required for energy conservation.
- . Each node has a fixed number of transmission power levels.

Assume that n nodes are dispersed in a field and the above assumptions hold. Our goal is to identify a set of cluster heads which cover the entire field. Each node vi, where $1 \le i$

<= n, must be mapped to exactly one cluster cj, where $1 \le j \le n$, and nc is the number of clusters (nc <= n). A

node must be able to directly communicate with its cluster head (via a single hop). Cluster heads can use a routing protocol to compute inter cluster paths for multi hop communication to the observer(s),. The following requirements must be met:

1. Clustering is completely distributed. Each node independently makes its decisions based only on local information.

2. Clustering terminates within a fixed number of iterations (regardless of network diameter).

3. At the end of each clustering process duration TCP, each node is either a cluster head, or not a cluster head (which we refer to as a regular node) that belongs to exactly one cluster.

4. Clustering should be efficient in terms of processing complexity and message exchange.

5. Cluster heads are well-distributed over the sensor field and have relatively high average residual energy compared to regular nodes.



Sensor information forwarding with and without clustering and aggregation:

(a) Single hop without clustering. (b) Multi hop without clustering. (c) Single hop with clustering. (d) Multi hop with clustering.

(b)

3.4 Network Radio Model



A typical sensor node consists of four major components: a data processor unit, a micro-sensor unit, a radio communication subsystem that consisting of transmitter/ receiver electronics, antenna, and amplifier; and a power supply unit. Although energy is dissipated by the first three components of a sensor node, we mainly consider the energy dissipation associated with the radio component. First Order Radio Model

We consider the first order radio model as discussed in with identical parameter values. The energy per bit spent in transmission is given by

$$\mathbf{e}_{\mathrm{tx}}(\mathrm{d}) = \mathbf{e}_{\mathrm{t}} + \mathbf{e}_{\mathrm{d}}^{*} \mathrm{d}^{\mathrm{n}}$$

where e_t is the energy dissipated per bit in the transmitter circuitry and $e_d * d^n$ is the energy dissipated for transmission of a single bit over a distance d, n being the path loss exponent (usually 2.0<= n<= 4.0). For a first order model we assume n=2 for simulation purposes.

Thus the total energy dissipated for transmitting a K-bit packet is

$$E_{tx}(K,d) = (e_t + e_d * d^2) * K$$

If e_r be the energy required per bit for successful reception then the energy dissipated for receiving a K-bit packet is

 $\mathbf{E}_{\mathrm{rx}}(\mathbf{K}) = \mathbf{e}_{\mathrm{r}} * \mathbf{K}$

It is assumed that the channel is symmetric so that the energy spent in transmitting from node i to j is the same as that of transmitting from node j to i.



In order to transmit a package of n bits at a distance of r, the radio transmitter will consume the following amount of energy:

$$E_{Tx}(n,r) = E_{tc}(n) + E_{amp}(n,r)$$

where $E_{tc}(n)$ is the energy that the radio circuitry needs to expend in order to process *n* bits, and $E_{amp}(n,r)$ is the energy needed by the radio amplifier circuit to send *n* bits at *r* meters. We can further refine by elaborating on the formula for $E_{amp}(n,r)$:

$$E_{Tx}(n,r) = E_{tc}(n) + E_{amp}(n,r) = n \cdot E_{trans} + n \cdot \varepsilon_{amp} \cdot r^{\gamma}$$

where E_{trans} is the energy needed to process a single bit by the radio transmission circuits, ε_{amp} is the transceiver's energy dissipation and γ represents the path loss exponent.

Alternatively, we can express in the same way the energy required for the transceiver to successfully receive and process n bits of data:

$$E_{R\mathfrak{X}}(n)=R_{\mathrm{FC}}(n)=n\cdot E_{\mathrm{FeCF}}$$

This model assumes that the communication through the radio channel is symmetric and that the energy to send a package from node A to B is the same as the one needed to send the same package from B to A, for a constant SNR. As can be seen in the above relations, any type of communication is not a low cost operation so the protocol stacks that run on the nodes should always try to minimize the number of transmit and receive operations in order to keep the energy budget of the network under a certain threshold.

3.4.2. Multi hop WSNs

So far, we have been focusing on modelling the communication between only two nodes but the same model can be scaled up to estimate the energy consumption at network level. For this, there are two cases worth taking into consideration: a network in which nodes talk to the gateway using a direct communication protocol, and the more general multi-hop network scenario, in which messages are passed from Neighbour to Neighbour until they reach the data sink.

Using the direct communication approach, each node has direct access to the gateway. As the distance between nodes and the gateway is not constant and can vary within radio connectivity range, some remote nodes will need greater amounts of transmit power to communicate with the data sink. In this case r is large, which leads to more energy spent and quicker battery drainage. On the other hand, there is no need for the nodes to receive any information from their neighbors, as the communication is done over a star topology network. This could prove advantageous or even optimal if nodes are in close proximity to the gateway or the cost of reception on the battery-powered nodes is sizeable.



First, for the single-hop case, the node is communicating directly to the gateway. For the N-th node, this would imply that it needs to increase its transmitter signal strength in order to cover the entire distance to the gateway, which would in turn lead to higher energy consumption. This can be expressed as:

$$E_b(N, n, r) = E_{TX}(n, N \cdot r) = n \cdot E_{trans} + n \cdot \varepsilon_{anp} \cdot (N \cdot r)^{\gamma}$$

For the multi-hop case, the N-th node needs to send data to his nearest neighbour, which would expend energy in receiving the package and retransmitting it to its nearest neighbour, and so on until it reaches the data sink. The total energy expenditure of the network can be calculated as a sum of N transmits and (N-1) receives:

$$\begin{split} E_{MH}\left(N,n,r\right) &= N \cdot E_{Tx}\left(n,r\right) + (N-1) \cdot E_{Rx}\left(n\right) \\ &= N \cdot n \cdot (E_{trans} + \varepsilon_{amp} \cdot r^{\gamma}) + (N-1) \cdot n \cdot E_{recv} \\ &= n \cdot (N \cdot (E_{trans} + E_{recv} + \varepsilon_{amp} \cdot r^{\gamma}) - E_{recv}) \end{split}$$

where n is the number of bits in a message. In most cases, however, all nodes in the network need to send packages to the base station. For the multi-hop case, we can generalize the above relation into N nodes:

$$\begin{split} \mathcal{E}_{MH}^{all}(n,r) &= \sum_{i=1}^{N} \mathcal{E}_{MH}(i,n,r) \\ &= N \cdot \mathcal{E}_{T\chi}(n,r) + (N-1) \cdot \mathcal{E}_{R\chi}(n) \\ &= \frac{N(N+1)}{2} \cdot n \cdot (\mathcal{E}_{trans} + \varepsilon_{anp} \cdot r^{\gamma}) + \frac{N(N-1)}{2} \cdot n \cdot \mathcal{E}_{recv} \end{split}$$

The same generalization can be made with the single-hop case given by:

$$\begin{split} E_{b}^{all}\left(n,r\right) &= \sum_{i=1}^{N} E_{Tx}\left(n,i\cdot r\right) \\ &= n \cdot N \cdot E_{trans} + n \cdot \varepsilon_{amp} \cdot r^{\gamma} \sum_{i=1}^{N} i^{j} \end{split}$$

Using the equations, we can derive the conditions for which direct communication to the gateway has a lower energy cost for the whole network, compared to the multi-hop scenario. This is equivalent to the following condition:

$$E_{b}^{all}\left(n,r\right) \leq E_{MH}^{all}\left(n,r\right)$$

Certain assumptions must be made in order to simplify the above relation. First, we can assume that the energy expended in processing one bit for transmission is roughly equal to the energy of processing a received bit, as most radio transceivers use the same electronics for both functions:

$$E_{trans} = E_{recv} = E_{circ}$$



5.1 Distributed Clustering Algorithm

In order to save the energy consumption of WSN, a clustering approach for WSN has been considered. In the approach, N sensor nodes are divided into clusters, and each cluster has a representative sensor node called cluster head (CH). Each non-CH sensor node sends the sensed data to the CH node in its own cluster, instead of to BS. Each CH node aggregates the received data into smaller size and sends it to BS. This approach has the following advantages: 1) non-CH sensor nodes can save the energy consumption because the nodes can avoid long-distance communication and have only to send data to its own CH being nearby and 2) the amount of data to be sent to BS can be reduced, which also saves the energy consumption.

The operating cycle of clustering methods, each round consists of consecutive frames. The first frame is for set-up, and the others are for steady state. In the set-up frame, CH nodes and clusters are determined based on the used clustering algorithm, and each CH assigns a non-CH node to a slot in order to create time-division multiple-access (TDMA) schedule. In the steady-state frames, each non-CH node sends data to CH at the assigned slot in TDMA fashion, and CHs fuse (compress) the received data and send it to BS. In order to decide CHs and clusters, clustering algorithms such as LEACH and HEED have been proposed [2, 3]. In this protocol, CHs are determined in a distributed autonomous fashion. At each round l, each node v independently decides to be a CH with probability Pv(t) . if the node v has not been a CH in the most recent (l mod (N/k)) rounds.





where k is the average number of CHs for each round. This means that each node becomes CH at least once every N/k rounds.

Thus the Threshold T(n) for static wireless sensor networks can be calculated as:

T(n)=[p/(1-p(r *mod(i/p)))] [(E_Nmax- E_{res})/ (E_{avg}- E_{res})][1- (cs/n)] , n ϵ G

cs= no. of cluster head + no. of non cluster heads , or it can be calculated as

$$cs = (1/r_t)^* (3D/\alpha\pi)^{1/3}$$

where r_t is the transmission range, α is the wavelength, D is the distance between the CH and the base station.

 $E_c = E_{CH} + E_{NCH}$ $E_{CH} = Cluster head energy.$

$$E_{CH} = Cluster head energy.$$

 E_{res} is the energy difference between initial energy and cluster energy (E_c).

$$E_{\rm CH} = \left(\frac{n}{k} - 1\right) L E_{\rm elec} + \frac{n}{k} L E_{\rm DA} + L E_{\rm elec} + L E_{\rm amp} d^{\rm HS}$$
(*

$$E_{\text{nonCH}} = LE_{\text{elec}} + LE_{\text{fs}}d^{2\text{CH}}.$$

Equation (8) is the energy consumption of both, in a cluster head nodes and noncluster head nodes in cluster:

$$E_{\text{cluster}} = E_{\text{CH}} + \left(\frac{n}{k} - 1\right) \cdot E_{\text{nonCH}}.$$
 (;

Therefore, all the energy consumed in the network is as follows:

$$E_{\text{total}} = L \{ 2nE_{\text{elec}} + nE_{\text{DA}} + E_{\text{fs}} (kd^{2BS} + nd^{2CH}) \}.$$
 (9)

p is the rate between the number of heads and the no. of total sensor nodes , 1/p is the expected no. of nodes in one cluster , r is the index of current round and G is the set of nodes that have not been cluster heads in the most recent r mod(1/p) rounds.

Here cs is the cluster size , n are the no. of nodes in a network , E_{N_max} is the initial energy , E_{res} is the residual energy , E_{avg} is the average energy of the network .

The cluster head probability must not fall below T(n) in a static homogeneous wireless sensor networks. A node with high residual energy and low communication cost (minimum distance of one hop from its non-cluster heads members) is selected as cluster head.

The Threshold $T(n)_{new}$ for mobile wireless sensor networks can be calculated as:

$$\label{eq:transform} \begin{split} T(n)_{new} =& [p/(1\text{-}p(r \ *mod(i/p)))] \ [\ (E_{N_Residual}/ \ E_{max})((V_{max} - V_{N_current})/V_{max})] \end{split}$$

$$R_{i} = \left(1 - c \frac{D_{max} - D(v_{i}, BS)}{D_{max} - D_{min}}\right) R_{0}$$
$$E[T_{N}] = \frac{\pi A}{E[V_{R}]L}$$

 R_0 as the predefined maximum competition range. D_{max} and D_{min} denote the maximum and minimum distance between network boundary and the base station. A is the area of the transmission range and L is the perimeter of this area. c is the coefficient between 0 to 1.

 $E_{N_Residual}$ is the current residual energy , E_{max} is the initial energy of the node , V_{max} is the maximum speed and $V_{N_current}$ is the current speed of the node.

The cluster head probability must not fall below $T(n)_{new}$ in a mobile homogeneous wireless sensor networks. A node with high residual energy and low communication cost (minimum distance of one hop from its non cluster heads members) is selected as cluster head in inter cluster routing and minimum R_i (distance from base station to the node) in intra cluster multi hop routing and high neighbourhood $E[T_N]$ time (low relative velocity) is selected as cluster head. As a result balanced clusters with uniform distribution clusters are formed with a greater network lifetime.

In this section the clustering algorithm intended to achieve longest cluster lifetime is proposed. Before proceeding with the presentation of the various steps of the algorithm, the major features of the algorithm is presented: 1) It sets an upper bound (N) on the number of cluster members under a cluster head to balance energy consumption, 2) a new cluster head does not force an existing valid cluster to reconstruct, 3) the cluster lifetime lasts until all of its affiliated cluster members have moved away and 4) it attempts to maximize the cluster lifetime at cluster construction by choosing the most stable nodes in mobility perspective to become the cluster heads. The expected number of cluster heads in an area is larger than or equal to the minimum number of overlapping circles of radius r_{tx} that covers the entire area. The expected number of cluster heads E[c] is calculated as

$$E[C] \ge \frac{1}{3^{3/2}} \left(\frac{2l^2}{r_{tx}^2} + \frac{l}{r_{tx}} \right)$$

The upper bound (N) on the number of cluster members under a cluster head is given by

$$N = \frac{n - E[C]}{E[C]}$$

The proposed Energy Efficient and Mobility aware distributed clustering algorithm possess fault tolerance characteristics with double cluster heads selection data redundancy schemes.

VI. SIMULATION RESULTS

The performance of the proposed scheme via simulations is presented in this section. This work uses MATLAB as the simulation tool where all simulations are conducted on networks using the IEEE 802.15.4 at the MAC layer. The free space physical layer model is adopted where all nodes within the transmission range of a transmitting node receive a packet transmitted by the node after a very short propagation delay. Direct Sequence Spread Spectrum (DSSS) codes are used to minimize inter cluster interference. Thus, this work ignores collisions in simulation. The set of simulation parameters are summarized Simulation parameters for static nodes:

Parameter	value
Network Grid	(0,0) to (100,100)
Sink	(50,50)
No of Nodes	100
Cluster Radius	25m
Initial Energy	1 j/battery
Data Packet Size	4000 bits
Broadcast Packet Size	200 bits
Round	2000
Eelec	50 nj/bit
E _{DA}	5nj/bit/signal
Efs	10pj/bit/m ²
Eamp	0.0013pj/bit/m ⁴
Threshold Distance	75 m

Simulation parameters for mobile nodes:

Parameter	va lu e
Network Grid	(0,0) to
(9000,10000)	
Sink	(0,0)
No of Nodes	20
Cluster Radius	50m
Initial Energy	0.03
j/battery	
Propagation Scheme	Two Ray
Ground	
Antenna Scheme	Omni
Directional	
Eelec	50 nj/bit
E _{DA}	
5nj/bit/signal	
Eamp	
0.0013pj/bit/m ⁴	
Threshold Distance	87m

Mobility Pattern in mobile nodes:

Pattern	pause time (min)	[V _{min} ,V _{max}](m/sec)	
MP1	10	[1,2]	
MP2	8	[1,4]	
MP3	6	[1,6]	
MP4	4	[1,8]	
MP5	2	[1,10]	



Node ids in static nodes



Distributed inter and intra clustering in static nodes

In static nodes distributed clustering of double nodes are selected for data redundancy and fault tolerance. The yellow and green lines indicates the intra cluster communication and blue lines indicates the inter cluster communication and the red lines indicates the multi hop routing which decrease the communication cost and message overhead. Base station is at the centre.





Distributed inter and inter clustering in mobile nodes

In mobile nodes distributed clustering takes place to bring stability to sensor network. The green lines indicates the intra cluster communication and blue lines indicates the inter cluster communication and the red lines indicates the multi hop routing which decrease the communication cost and message overhead. Base station is at the corner. When velocity of the mobile nodes increases frequent re clustering takes place as a result high neighbourhood time is taken as a secondary parameter to select cluster heads in mobile nodes clustering.



For a fixed value of c, the number of clusters decreases for an increase in R0 and for a fixed value of R0, the number of cluster increases with c. The reason is that the competition range (Ri) decreases either by increasing c while keeping R0constant or by decreasing R0 while keeping c constant.



This EEMADC protocol considers the Neighbour time as a primary parameter while constructing the clusters and considers both residual energy and Neighbour time while constructing the cluster head backbone used for data delivery to the sink. In contrary, nodes in LEACH-Mobile and CBR choose the cluster head according to received signal strength and do not consider the node mobility. It is observed that the packet delivery ratio in MAUCR protocol, and other Mobility protocols decreases for an increase in node velocity. The reason is that nodes will keep changing their affiliating clusters more frequently and faster with node velocity. This results in many disconnection periods which in turn causes high packet loss. The percentage of successfully received packets suffers more when the disconnection periods are more frequent and extended for long time.

However, it is observed that EEMADC protocol is less affected by an increase in node velocity. This is because of the stable link between the sensor nodes and their affiliating cluster heads and the stable links between relay cluster heads created by EEMADC protocol.



The key conclusion from these plots is that the performance of the clustering algorithm decreases with the increase of node speed. This is justified because the degree of network instability is proportional to network dynamics i.e. the average speed of the nodes in the network. In the normal human running speed scenario, a negligible difference in performance is observed for both mobility models. Moreover, performance parameters like the accuracy of throughput and end-to-end delay are not affected by different levels of randomness setting in the Gauss-Markov mobility model.



It is observed that energy consumption decreases as the rate increases ; however the curve flattens when the rate

is larger than 5% , and it is assumed that sensor nodes are uniformly deployed and the rate should be larger than 5% to satisfy net

CONCLUSION

- In our future work, we would like to consider more practical network mobility, and find out experimentally the accuracy of the neighbourhood tracking scheme.
- Being spread out over a possibly hostile area, the sensor nodes face an adversary model which varies depending on the attacker's capability and the hardware configuration. The attacker may be able to eavesdrop on communications, interfere with message transmissions, insert its own messages or even have physical access to the device and access to its stored data. Data encryption, authentication and integrity, faulty node detection are included in my future research area interests.
- My future work will deals with secure energy efficient and mobility aware distributed unequal clustering fault tolerance algorithms and faulty node diagnosis.
- My work addresses the fundamental problems under a varying degree of node mobility in wireless sensor network.
 - The protocol is simple and frequent re-clustering is avoided since the protocol ensures a longer cluster life time for a wide range of node mobility while maintaining low energy overhead.
- The re-clustering time is very less which ensures less network overhead.
- The performance of the clustering algorithm decreases with the increase of node speed. This is justified because the degree of network instability is proportional to the network dynamics i.e. the average speed of the nodes in the network

REFERENCES

- [1] Mahapatro, A., and Khilar, P.M., Fault Diagnosis in Wireless Sensor Networks: A Survey. IEEE Communications Surveys and Tutorials, in press.
- [2] Awwad, S., C. Ng, N. Noordin, and M. Rasid, Cluster based routing protocol for mobile nodes in wireless sensor network, Wireless Personal Communications, 61, pp. 251–281, 2011.
- [3] Mahapatro, A., and Khilar, P.M., Mobility Adaptive Unequal Cluster-based Routing Protocol in Wireless Sensor Networks. International Journal of Sensor, in press.
- [4] C. R. Lin and M. Gerla. Adaptive clustering for mobile wireless networks. IEEE Journal on Selected Areas in Communications, 15(7):1265–1275 sept 1997.
- [5] Alberto Sangiovanni-Vincentelli Farinaz Koushanfar, Miodrag Potkonja,"FaultTolerance in wireless sensor networks "A Survey Computer Networks, vol. 38, no.4, pp. 393-422, 2002.
 - Chen, G., C. Li, M. Ye, and J. Wu, An unequal

[6]

cluster-based routing protocol in wireless sensor networks, Wireless Networks, 15, pp. 193–207, 2009.

- [7] A. B. MaDonald and T. F. Znati, "A Mobility-based Frame Work for Adaptive Clustering in Wireless Ad Hoc Networks," IEEE JSAC, vol. 17, Aug. 1999., pp. 1466–87
- [8] W. Su, S.-J. Lee, M. Gerla, Mobility prediction and routing in ad hoc wireless networks, Int. J. Network Manage. 11 (1) (2001) 3–30.
- [9] G.J. Pottie and W.J. Kaiser, "Wireless Integrated Newtork Sensors," Comm. ACM, vol. 43, no. 5, pp. 51-58, May 2000.
- J. Agre and L. Clare, "An integrated architecture for cooperative sensing networks," IEEE Computer, vol. 33, pp. 106–108, May 2000work coverage.



Mr.Amaresh Choudhury, was born in India. He received M.Tech degree in Electrical Engineering from Biju Patnaik University of Technology, Orissa. He is working as an Assistant Professor in Gandhi School of Engineering, Berhampur . His major research interests include wide-area Monitoring system and its application in power system



Mr. Bishnu Prasad Panda, was born in India. He received M.Tech degree in Electrical Engineering from Biju Patnaik University of Technology, Orissa. He is working as an Assistant Professor in Gandhi School of Engineering, Berhampur. His major research interests include Photovoltaic system and its application in power system



Ms. Priteesa Mahapatra , was born in India. She received M.Tech degree in Electronics and Communication Engineering from Biju Patnaik University of Technology, Orissa. She is working as an Assistant Professor in Gandhi School of Engineering, Berhampur. Her major research interests include embedded system and its applications

