Survey on the Clustering Schemes of Mobile Ad-hoc Wireless Sensor Network (WSN)

Department: EEE of BRAC University

Thesis Supervisor:
Hasan Shahid Ferdous

Thesis Co-supervisor
Sanjida Hossain Sabah

Prepared by:

Mushrif Ahmed
09221095
Rakesh Sarker
09221071

Declaration:

We hereby declare that this thesis is our original work. It has been prepared only for the purpose of meeting the requirement of the undergraduate program of Electrical and Electronic Engineering (EEE) of BRAC University. This has not been submitted anywhere else. And wherever another person's work has been used, it has been duly accredited.

Signature of Supervisor:

4/09/12

....

Hasan Shahid Ferdous

Signature of Co-Supervisor:

....

Sanjida Hossain Sabah

Signature of Authors:

Rakesh Sarker
04/09/2012

....

Rakesh Sarker
[Student Id: 09221071]

Mushrif 4-9-2012

....

Mushrif Ahmed
[Student Id: 09221095]

Acknowledgement:

First of all we would like to thank Almighty God for giving us the ability, knowledge and good health to complete this thesis.

Then we would like to thank our respected thesis supervisor, Hasan Shahid Ferdous, for all his guidance and help. Without his help this thesis could not have carried out at all.

Next we would like to thank our co-supervisor, Sanjida Hossain Sabah for her help and support.

We would also like to thank our friends and family for their unending support and belief in us, which gave us strength to complete our thesis

Abstract:

This report aims to do a complete study on the different clustering algorithms of an ad-hoc wireless sensor network and compare them. In this paper we have understood, what are the basics of the 802.11 protocol, what is a wireless ad-hoc sensor network, what is a Mobile Ad-hoc Network (MANET), and details of Clustering. We have further understood what is clustering, why it is necessary, and the different clustering schemes. Among these we have understood their clustering process, the cluster maintenance, benefits and drawbacks, and finally compared them with each other. In our thesis we work with four clustering scheme and try to find out their formation and cluster head selection process time. We have understood these various concepts by extensive study of IEEE papers and acquiring information from the internet. We hope that our report will aid anyone seeking to understand information on this topic and be a concise compilation of information for their research or study.

Motivation

Wireless sensor network has drawn the interest of many researchers over the past decade and many works have been directed to find a suitable clustering structure for a large number of sensor nodes deployed over a large field operating in ad hoc manner. In this research, we are focusing on these clustering schemes and identify the merits and drawbacks for each of them. We investigate clustering schemes based on dominating sets, energy efficiency, load balancing and low maintenance based clustering. We find out the cluster formation and cluster head selection process time so that they can be used based on application scenario.

ABBREVIATION

ACK	Acknowledgement
AP	Access Point
BSS	Basic Service Set
CTS	Clear To Send
DIFS	Distributed Inter Frame Space
DS	Distribution System
DS	Direct Sequence
ESS	Extended Service Set
FH	Frequency Hopping
IR	Infrared
LAN	Local Area Network
MAC	Medium Access Control
MANET	Mobile Ad-hoc Network
RA	Receiver Address
RTS	Request To Send
TA	Transmitter Address
WEP	Wired Equivalent Privacy
WSN	Wireless Sensor Network

DS	Dominating Set
CDS	Connected Dominating Set
WCDS	Weakly Connected Dominating Set
VID	Virtual ID
IDBLC	ID Load Balancing Clustering
AMC	Adaptive Multi-Hop Clustering
DLBC	Degree Load Balancing Clustering

Contents

List of Figures

List of Tables

1. Background Knowledge:

1.0 802.11 Protocol Basics:

1.1.1 Architecture component:

802.11 is a cellular based architecture where the total system is subdivided into some cell. Each cell is called Basic Service Set or BSS and each BSS is controlled by a base station called Access Point or AP. The APs of different cells are connected through a connection (Typically Ethernet or wireless) called Distribution System or DS. The whole system including Distribution System is called Extended Service Set (ESS).

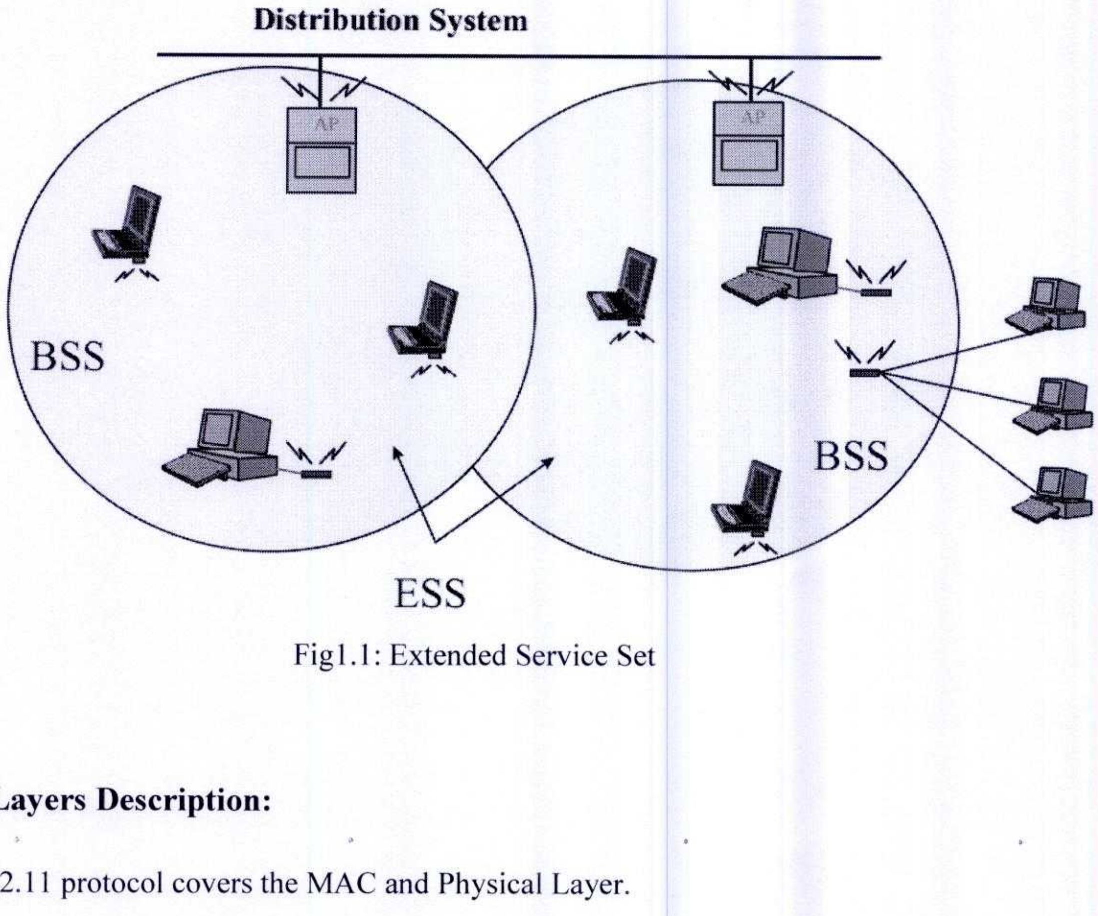

Fig1.1: Extended Service Set

1.1.2 Layers Description:

The 802.11 protocol covers the MAC and Physical Layer.

- Frequency Hopping Spread Spectrum in the 2.4 GHZ band
- Direct Sequence Spread Spectrum in the 2.4 GHz Band
- Infrared

<table>
<tr><td colspan="3">802.2</td><td rowspan="2">Data Link Layer</td></tr>
<tr><td colspan="3">802.11 MAC</td></tr>
<tr><td>FH</td><td>DS</td><td>IR</td><td>PHY Layer</td></tr>
</table>

Chart 1.1- Layer Description

1.1.3 Fragmentation and reassembly:

Typically the whole data are send in small packets .The data can be send whole at a time but there is some reason to divide data in small range.

- Because of higher bit error rate of a radio link, the probability of a packet to get corrupted increase with the packet size.
- Smaller packet contains small overhead. So, in case of re transmission it is very easy.
- On a frequency hopping system, the medium is interrupted periodically for hopping, so the smaller the packet, the smaller the chance that the transmission will be postponed to after the dwell time.

The mechanism is a simple Send and wait algorithm, where the transmitting station is not allowed to transmit the next fragmentation until one of the following happens:

- If the station get an acknowledgement from the receiver.
- Or the station decides it retransmit the frame so many times

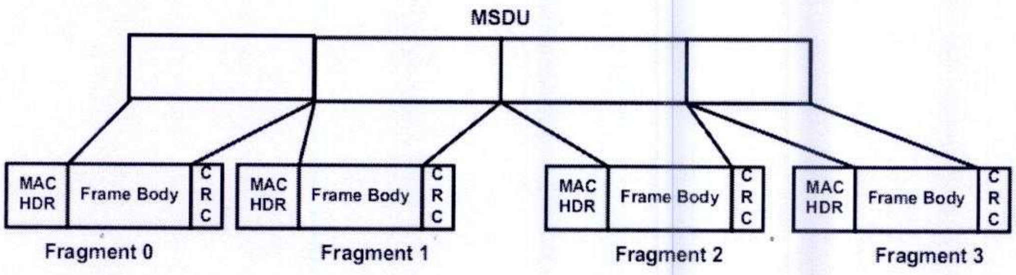

Fig 1.2: Fragmentation of a big data

1.1.4 Joining an existing BSS cell:

When a station wants to access an existing BSS cell it has to make a connection to the access point of that BSS through synchronization information. The station can get this information in two ways:

- Passive Scanning: In this case the station just wait to receive a Beacon Frame from the AP
- Active Scanning: In this case the station tries to find an Access Point by transmitting probe request frames and wait for the response from the AP.

After getting the AP and taking decision to join its BSS, the station and AP will interchange information, where each side proves the knowledge of a given password. This process is called **Authentication Process**. After Authentication process **Association process** starts. It is also an information exchange process where both station and BSS exchange information about their capabilities. When both Authentication and Association Process done, a station is capable of transmitting and receiving data frame.

1.1.5 Distributed Coordination Function (DCF):

This is the basic access mechanism for the 802.11 standard. This uses a CSMA/CA with Binary Exponential Back off algorithm. CSMA/CA stands for Carrier Sense Multiple Access with Collision Avoidance mechanism.

Normally a CSMA protocol works in the following manner. First a station willing to transmit senses the medium, if the busy then the station will transmit later, if the medium is free, then the station can transmit. This is a very effective way of transmission unless the medium is highly loaded, but the only problem is that there is a chance of collision (which occurs when two stations transmit at the same time). A collision might occur when two stations sensed the medium free at the same time and decided to transmit simultaneously.

These collision situations must be identified, so that the MAC layer can retransmit the packet by itself and not by upper layers. If upper layers had to retransmit, it would cause significant delay. In the case of Ethernet, the collision is recognized by the transmitting stations which carry out a retransmission phase based on an exponential random back off algorithm. These collision detection mechanisms are good for wired LANs, but they cannot be used for a wireless LAN because of two main reasons. The first reason is that, implementing a Collision Detection Mechanism means that a Full Duplex radio is required. This means that the radio must be capable of transmitting and receiving at the same time. This would increase the price of the system significantly. The second reason is that, in a wireless environment, it cannot be assumed that all stations can hear each other. This is the basic assumption of the Collision Scheme. And if a station senses the area free around the transmitting area, it does not mean that the medium is free near the receiver area.

So to overcome these problems, 802.11 standards use a Collision Avoidance mechanism together with a Positive Acknowledgement scheme. The scheme is as follows: A station willing to transmit senses the medium. If the medium is free, then it delays the sending. If the medium is free for a specific amount of time (called DIFS, Distributed Inter Frame Space), then the station can transmit. The receiving station will check the CRC of the received packet and send an Acknowledgement packet (ACK). Receiving the Acknowledgement means that no collision has occurred. If the sender does not receive acknowledgement then it will retransmit the packet until it gets acknowledged or it will be thrown away after a given number of retransmissions.

1.1.6 Virtual Carrier Sense

The standard defines a Virtual Carrier Sense mechanism to avoid two stations colliding. A station which is about to transmit will first transmit a short control packet called **RTS** (Request To Send), which includes the source, destination and duration of the data transfer. The destination will respond (if the medium is free) with a response packet called **CTS** (Clear to Send), which includes the same information.

All stations that receive the RTS and/or CTS will set their Virtual Carrier Sense indicator (called NAV, Network Allocation Vector), for the given duration, and will use this information together with the Physical Carrier Sense when sensing the medium.

This mechanism reduces the chances of collision on the receiver area by a station that cannot hear the short RTS from the transmitting station. This is because the station will hear the CTS and reserve the medium as busy until the end of the transmission. The duration of the data transmission on the RTS also avoids collisions during the ACK near the transmitter area. The fact that RTS and CTS are short frames also reduces the number of collisions.

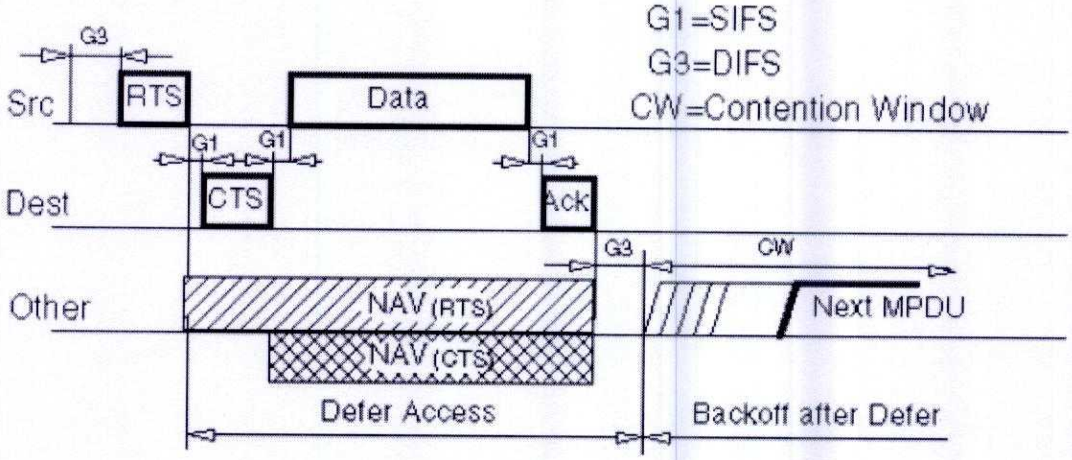

Fig 1.3: Transaction between two stations A and B and the NAV setting of neighboring stations

1.1.7 Frame Types and Frame Format

There are three main types of frames

1. Data Frames (those used for data transmission)
2. Control Frames (those used to control access to the medium; [e.g. RTS, CTS and ACK])
3. Management Frames (these are exchanged similarly to data frames, but contain management information and are not forwarded to upper layers)

1.1.8 Most Common Frames Format

1.1.8.1 RTS Frame Format

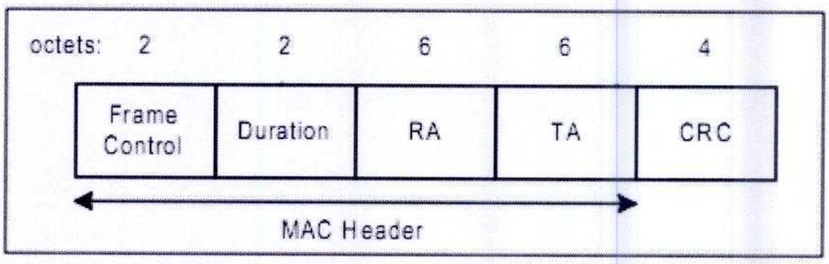

Fig 1.4: RTS Frame Format

The Receiver Address (RA) of the RTS frame is the address of the Station on the wireless medium that is intended to be the immediate recipient of the next Data or Management frame

The Transmitter Address (TA) is the address of the station transmitting the RTS frame.

The duration value is the time in microseconds required to transmit the next Data or Management frame, plus one CTS frame, plus one ACK frame, plus three SIFS intervals.

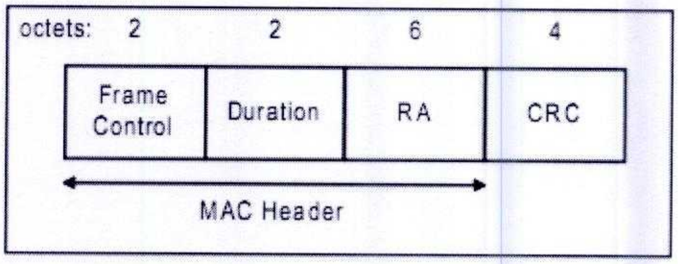

Fig1.5: CTS Frame Format

The Receiver Address (RA) of the CTS frame is copied from the Transmitter Address (TA) field of the immediately previous RTS frame to which the CTS is a response.

The duration value is the value obtained from the Duration field of the immediately previous RTS frame minus the time in microseconds required to transmit the CTS frame and its SIFS interval.

1.1.8.3 ACK Frame Format

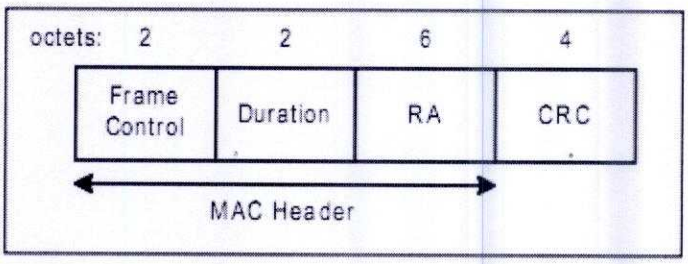

Fig 1.6: ACK Frame Format

The Receiver Address (RA) of the ACK frame is copied from the Address 2 Field of the immediately previous frame.

If the More Fragment bit was set to 0 in the Frame Control field of the previous frame, the Duration value is set to 0, otherwise the Duration value is obtained from the Duration field of the previous frame minus the time in microseconds, required to transmit the ACK frame and its SIFS interval.

1.1.9 Exponential Back off

Back off is a method which solves the problem of different stations access a medium. A station following the method must choose a random number 'n' between 0 and a given maximum number 'M'. Then it must wait for this number (n) of slots before accessing the medium.

Exponential back off means that each time the station chooses a slot and collides; it will increase the random number M exponentially

Exponential Back off Algorithm must be executed in the following cases

a) When the station senses the medium before transmission, and it is busy
b) After each retransmission
c) After a successful transmission

1.1.10 Roaming

Roaming is the process of moving from one cell (or BSS) to another without losing connection. The 802.11 protocol defines the basic tools for roaming. This protocol includes active / passive scanning and a re-association process (where a station roaming from 1 AP to another will be associated with the new one)

1.1.11 Keeping Synchronization

On an infrastructure BSS synchronization is done by all the stations updating their clocks according to the Access Points clock using the following mechanism:

The AP transmits periodic frames called Beacon Frames; these frames contain the value of the clock on the moment of transmission.

The receiving stations check the value of their clock at the receiving moment, and correct it to keep it synchronized with AP's clock.

1.1.12 Security:

Security always play an important role in a wireless lan.802.11 committee ensure the security by providing WEP (Wired Equivalent Privacy).

The main concern of the security system is preventing an intruder from:

- Access the network resources by using similar wireless LAN equipment.
- Be able to capture the wireless LAN traffic (Eavesdropping)

To connect a workstation a user have to enter the premises (by using a physical key).an intruder must not know the premises. So it can be easily said that he can't connect to the workstation. Thus access to the network resources can be prevented.

Eavesdropping is prevented by the use of the WEP algorithm which is a Pseudo random number generator initialized by a shared secret key. This PRNG outputs a key sequence of pseudo random bits equal in length to the largest possible packet which is combining with the outgoing /incoming packet producing the packet transmitted in the air.

1.1.13 Power saving:

Wireless LANs are typically use for mobile device, and battery power is a scare type of resource for this type of device.802.11 take this matter very carefully and allow the station to go sleep mode without losing any information.

The mechanism is AP of a station always maintains an update about the devices those are in sleeping mode in his station. AP buffer the messages those are comes to that sleeping devices and don't pass them until the devices send a poling request or change their operation mode.

AP also transmit information about buffered frames hopping that when a device change his sleeping mode will get this information and understand that a frame is waiting for him in AP and send a pole message to get the message.

1.1.14 Ad-hoc Networks

In certain circumstances the user will desire to build up Wireless LAN networks without an infrastructure (more specifically without an Access Point),this may include file transfer between two notebooks user, a coworkers meeting outside the office, etc.

The 802.11 Standard addresses this need by the definition of an "ad-hoc" mode of operation, in this case there is no access point and part of its functionality is performed by the end user stations, and other functions are not supported.

1.1.15 Int Coordination Function (PCF)

Aside from the Distributed Coordination Function, there is an optional Point Coordination Function (PCF). PCF uses the higher priority the Access Point may get by using a smaller Inter Frame Space. By using higher priority access, the Access Point issues polling requests to the stations for data transmission, hence controlling medium access.

2. Wireless Sensor Network (WSN)

Wireless Sensor Networks consists of spatially distributed autonomous sensors to monitor physical or environmental conditions, such as temperature, sound, vibration, pressure, motion or pollutants and to cooperatively pass their data through the network to a main location. The WSN is consists of nodes, which can number from a few to several hundreds or even thousands, where each node is connected to one (or sometimes several) sensors. Each such sensor network node has typically several parts: a radio transceiver with an internal antenna or connection to an external antenna, a microcontroller, an electronic circuit for interfacing with the sensors and an energy source, usually a battery or an embedded form of energy harvesting. The cost and size of the sensor nodes are variable. Due to the size and cost restrictions, there are other restrictions such as energy, memory, computational speed and communications bandwidth. The topology of the WSNs can vary from a simple star network to an advanced multi-hop wireless mesh network. The propagation technique between the hops of the network can be routing or flooding.

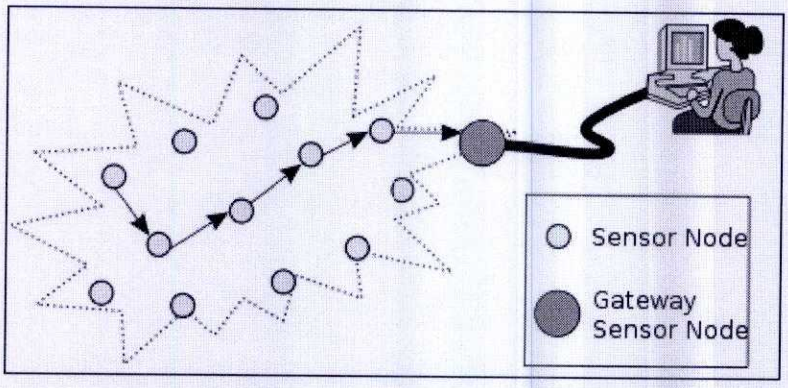

Fig 2.1: Typical multi-hop wireless sensor network architecture

The main characteristics of a WSN include

- Power consumption constrains for nodes using batteries or energy harvesting
- Ability to cope with node failures
- Mobility of nodes
- Dynamic network topology
- Communication failures
- Heterogeneity of nodes
- Scalability to large scale of deployment
- Ability to withstand harsh environmental conditions
- Ease of use
- Unattended operation.

3. Mobile Ad Hoc Networks (MANET)

Wireless ad hoc networks are a method by which data and Internet access may be shared directly between one or more computers or other devices. This is different from a traditional network where a router is used to connect the computers to each other and the Internet. With a wireless ad hoc network, one computer can act as wireless access point, or a number of computers may act as nodes in the network. In this case, the routing of data is usually dynamic.

Mobile ad hoc networks are systems that allow a variety of computers and other devices to interconnect wirelessly. They operate in a similar manner to other wireless ad hoc networks, though there may be special adjustments in place to deal with each part of the network moving freely in relationship to all of the other parts. In a mobile ad hoc network, each computer or other device typically operates as a node to facilitate communications within the system. This may allow every computer in the network to communicate with every other computer, though the path that the data travels may vary depending on factors such as wireless signal strength and the physical location of the components.

MANETs tend to be self-configuring, in that components can be moved around, added, or removed dynamically. The network will typically be able to continue using networked computers as routers regardless of their position, unless they are moved outside the wireless range of other components. Other computers that are correctly configured can typically join the network upon entering wireless range.

4. Problems in WSN

Here are some of the problems of a WSN:

1. Low battery power of nodes
2. The fact that nodes are mobile
3. The fact that nodes must be maintained
4. The fact that new nodes are entering and leaving a network
5. Nodes must be made for the specific purpose for which they must be used

5. Clustering:

5.1 What is Clustering?

Clustering is an important topic for Mobile Ad-hoc Network (MANET) because clustering makes it possible to guarantee basic labels of system performance. This performance is for both mobility and a large number of mobile terminals. With increase of using internet: ad hoc network gained worldwide attention in recent years. Though internet is popular among people but it cannot satisfy the demand of using internet at any time any place. MANET allows mobile terminals to set up a temporary network for instant communication. For this specific application it is very helpful in some place like: battle field communication, conferencing, sensor dust and so on.

In a clustering scheme the total nodes in a MANET are divided into some virtual groups and the groups are connected to each other. The nodes on a cluster have to maintain some rules with different behavior then the nodes outside of the cluster. On a cluster the nodes can have different roles like: 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 neighboring clusters and forward information between clusters. A cluster member is an ordinary node which has no outer cluster links, only have interred cluster links.

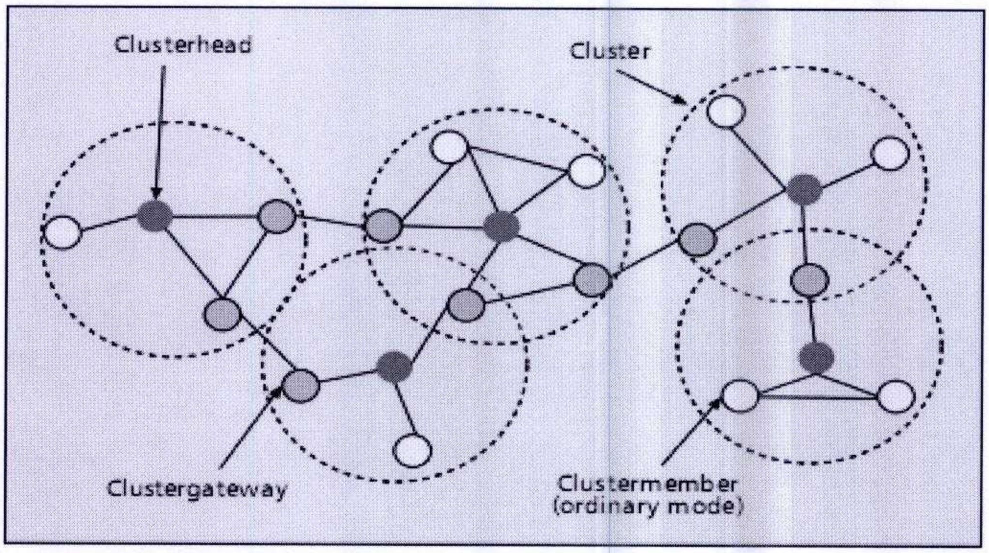

Fig 5.1: A cluster structure illustration

5.2 Necessity of clustering:

For a large number of mobile networks it has been shown that cluster structure guarantees basic performance of a MANET. It provides at least three benefits. First, a cluster structure facilitates the special reuse of resources to increase the system capacity. In non overlapping multi cluster structure, two clusters can share same frequency if they are not neighboring cluster. Even cluster uses a special nod called AP residing on it maintain re transmission of information saves munch resources. Secondly, in a cluster cluster-head and cluster gateway play as a role like route. Thus the generation and spreading of routing information can be restricted in this set of nods. Finally, it makes ad hoc network appear smaller and more stable. When a nod changes its cluster then it makes a connection to the new cluster. Only the mobile nodes of the corresponding cluster have to updates its 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.

6. Classify clustering schemes:

Clustering schemes of MANET can be classified into different criteria, like: clustering protocol can be classified with the presence of cluster head –cluster head based clustering and non cluster head based clustering. Hop based clustering like: 1 hop clustering and multi hop clustering.

In this paper we will classify the clustering protocol based on their objectives. According to the objective schemes can be grouped into four categories-

- DS-based clustering
- Energy Efficient clustering
- Load Balancing clustering
- Combined Metrics clustering

Dominating –Set based clustering tries to find a DS for a MANET so that the number of mobile nodes that participate in routing table maintenance can be reduced. Energy efficient clustering manages to use the battery energy of mobile nodes more wisely in a MANET. Load balancing clustering schemes attempt to limit the number of mobile nodes in each cluster to a specified range so that cluster are similar size. Combined matrices based clustering usually consider multiple matrices, such as node degree, cluster size, mobility speed and battery energy, in cluster configuration, especially in cluster head decision.

6.1 Objectives of clustering scheme:

Clustering Schemes	Objectives
DS-based Clustering	Finding a (weakly) connected dominating set to reduce the number of nodes participating in route search or routing table maintenance.
Energy Efficient Clustering	Avoiding unnecessary energy consumption or balancing energy consumption for mobile nodes in order to prolong the lifetime of mobile terminals and a network.
Load Balancing Clustering	Distributing the workload of a network more evenly into clusters by limiting the number of mobile nodes in each cluster in a defined range.
Combined Metrics Clustering	Considering multiple metrics in cluster configuration including node degree, mobility, battery energy, cluster size, etc. And adjusting their weighting factors for different application scenario.

Table 6.1: Summary of six clustering schemes

7. Clustering schemes:

7.1 DS Based Clustering:

This is the algorithm in which routing is done based on a set of dominating nodes, which function as cluster heads to relay routing information and data packets. A node belongs to a cluster head subset of a cluster head DS if the node belongs to or is adjacent to the cluster head. This is shown in the figure. A DS is called a connected DS (CDS) if all the dominating nodes are directly connected with each other. When table-driven routing is applied, only nodes in the CDS are required to construct and maintain the routing tables. When on-demand routing is adopted, the route search space is limited to the CDS.

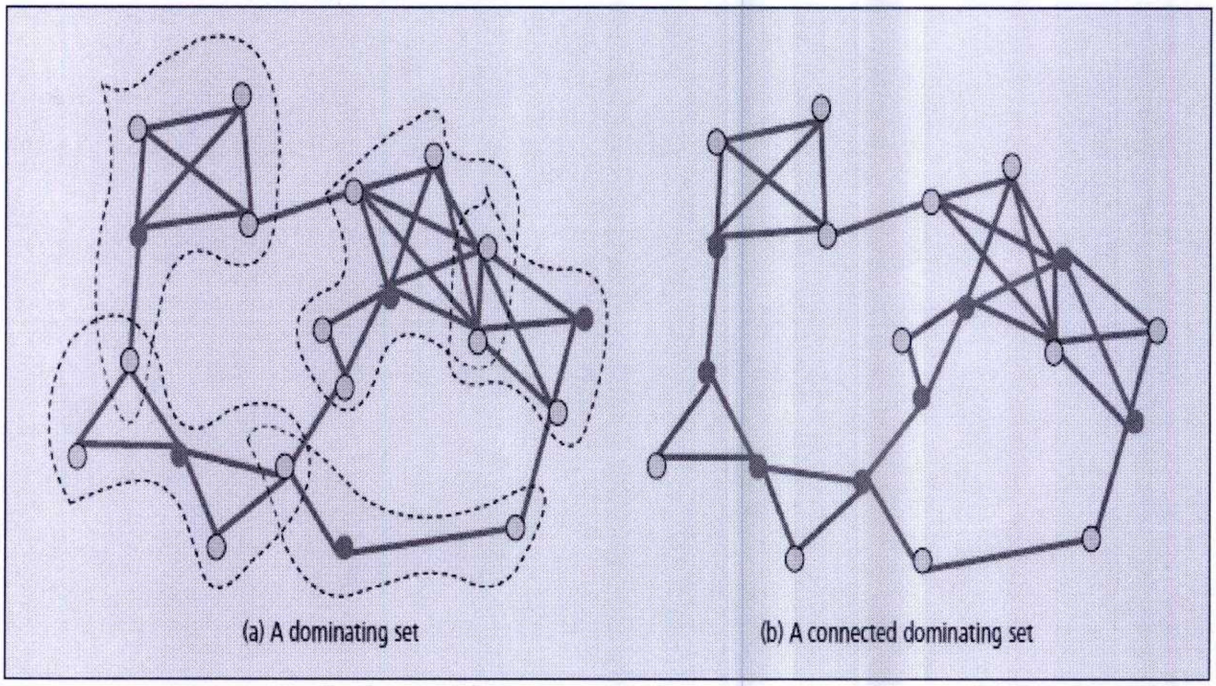

Fig 7.1: A Dominating Set

7.1.1 Connected Dominating Set (CDS) Algorithm:

Wu proposed a distributed algorithm to find a CDS in order to design efficient routing schemes for a MANET. Initially, every node exchanges its neighbor list with all its neighbors. A mobile node sets itself as a dominating node if it has at least two unconnected neighbors. This is called a *marking* process. Then some extension rules are implemented to reduce the size of a CDS generated from the marking process. A node deletes itself from the CDS when its close (open) neighbor set is completely included in the neighbor set(s) of a (two connected) neighboring dominating node(s) and it has smaller ID than the neighboring dominating node(s). The close neighbor set of a mobile node includes all its direct neighbors as well as itself, whereas the open neighbor set only includes a mobile node's direct neighbors.

7.1.2. Merits of CDS Clustering

- The extension rules proposed in Wu's algorithm are effective for reducing the size of the DS
- Wu's algorithm promises that the cluster construction can be completed in just two rounds, one round for the marking procedure and the other round for extension rules. So this means that this algorithm requires very little computation time.

7.1.3. Demerits of CDS Clustering

- For CDS maintenance, any moving node needs to continue to send out a beacon message every second during its movement. Other related mobile nodes keep monitoring the messages from that moving node. Therefore, a single mobile node's movement may suppress many mobile nodes from transmitting or receiving their own packets.
- Also, if many mobile nodes in the network are in movement complete recalculation of a CDS is required which requires a large amount of message exchange.

7.2 Energy Efficient Clustering

The main aim to this algorithm is to reduce the battery usage of the mobile nodes as this is a challenge for network performance. We know that a cluster head has more work than an ordinary member, so it is most likely to die earlier. Hence, it is also important to balance the energy consumption among mobile nodes to avoid node failure, especially when some mobile nodes bear special tasks or the network density is comparatively sparse.

7.2.1 IDBLC (ID Load Balancing Clustering:

Each mobile node has a variable, virtual ID (VID), and the value of VID is set as its ID number first. The mobile node with the highest VID in their local area wins the cluster head role. After the cluster head exhausts its duration (Max Count) it resets its VID to 0 and becomes a non cluster head node. When two cluster heads move into the reach range, the one with higher VID wins the cluster head role. Each non-cluster head node keeps a circular queue for its VID and shifts the VID value by one every time unit in one direction. The newly chosen mobile node is the one whose previous total cluster head serving time is the shortest in its neighborhood, and this guarantees good energy level for being a new cluster head.

7.2.2 Merits of IDBLC

- This algorithm ensures that a node remains cluster head for a long time and there is no excessive load on a node that it may deplete its energy

7.2.3 Demerits of IDBLC

- New cluster head selection may introduce ripple effect of re-clustering over the whole network without considering the network topology
- The cluster head re-election may require time synchronization of the VID value shift among different mobile nodes
- In addition, the cluster head serving time alone may not be able to promise a good indication of energy consumption of a mobile node.

7.3 Load-Balancing Clustering

Load-balancing clustering algorithms believe that there are an optimum number of mobile nodes that a cluster can handle. A too-large cluster may put too heavy of a load on the cluster heads, causing cluster heads to become the bottleneck of a MANET and reduce system throughput. A too-small cluster, however, may produce a large number of clusters and thus increase the length of hierarchical routes, resulting in longer end-to-end delay.

7.3.1 AMC (Adaptive Multi-Hop Clustering)

AMC maintains a multi hop cluster structure based on load-balancing clustering. AMC does not describe how the clusters are initially constructed. However, for cluster maintenance each mobile Node periodically broadcasts its information, including its ID, CID, and status (cluster head/member/gateway) to others within the same cluster. Each gateway also periodically exchanges information with neighboring gateways in different clusters and reports to its cluster head. Thus, a cluster head can recognize the number of mobile nodes of each neighboring cluster. AMC sets upper and lower bounds on the number of cluster members that a Cluster head can handle.

In AMC, when the number of members in a cluster is less than the lower limit, the merge mechanism is invoked. The cluster tries to find a neighboring cluster such that their combined sum is not greater than the upper limit. If the sum of the two clusters is greater than the upper limit then it tries to find another cluster so that their sum is lower than the upper limit. When two clusters merge, the cluster head with more member nodes wins to continue the cluster head role.

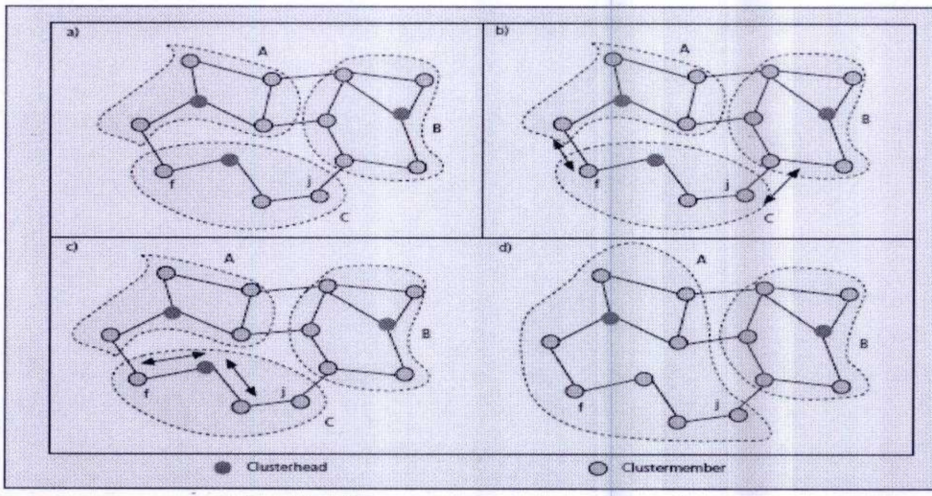

Fig 7.2: The process of cluster merging

A cluster performs the division mechanism if its number of nodes is greater than the upper limit. By choosing a suitable node as a new cluster head for the detached cluster, the initial cluster can be separated into two clusters with almost the same size.

7.3.2 Merits of AMC

- It keeps the number of nodes in a cluster within the optimum range

7.3.3 Demerits of AMC

- AMC does not address how to select a proper node to serve as the cluster head for the newly detached cluster.
- AMC mentions that the upper and lower bounds should be decided by network size, mobility, and so on, these values must be given in advance

7.1.1 DLBC (Degree Load Balancing Clustering)

DLBC periodically runs the clustering scheme in order to keep the number of mobile nodes in each cluster around a system parameter, which indicates the optimum number of mobile nodes that a cluster head can handle. A cluster head degrades to an ordinary member node if the difference between the system parameter and the number of mobile nodes that it currently serves exceeds a certain value. This mechanism tries to make all cluster heads almost serve the same and optimal number of member nodes.

7.3.5 Merits of DLBC

- Compared with other algorithms DLBC can reduce the rate of cluster head change because a cluster head does not need to relinquish its cluster head status whenever it has a member node with a higher node degree

7.3.6 Demerits of DLBC

- Since the cluster head change is still based on node degree, DLBC likely will cause frequent re-clustering because the movement of mobile nodes and consequent link setup/break results in dynamic variation of mobile node degree.
- In addition, how to select a cluster head is not addressed in DLBC

7.4 Combined Metric Based Clustering

Combined Metrics based clustering takes into account a number of metrics such as node degree, residual energy capacity, moving speed, etc. This scheme aims to elect the most suitable cluster-head in a local area and does not give preference to a node with a specific attribute.

7.4.1 on Demand WCA (Weighted Clustering Algorithm)

On demand WCA [18] considers four parameters for each mobile node in the cluster-head election process. They are degree difference Dv , sum of the distance with all neighbors Pv, average moving speed Mv , and cluster-head serving time Tv. Here, Dv = jdv □ Mj where dv is the number of neighbors of a mobile node v, and M is the number of nodes a clusterhead can handle ideally. The combined weight factor Iv, is calculated as Iv = c1Dv +c2Pv +c3Mv +c4Tv, where c1, c2, c3, and c4 are the weighting factors. All the parameter values are normalized according to some pre-defined values. On-Demand WCA chooses mobile nodes with minimum Iv in the local area to be cluster-heads. All mobile nodes covered by the elected cluster-heads cannot participate in further cluster-head selection. This procedure is repeated until each mobile node is assigned to a cluster. The cluster-head election algorithm is invoked at the very beginning of cluster formation or when the current cluster-heads are not able to cover all mobile nodes. The node that has the minimum Iv will be the Cluster head WCA does not re-cluster when a member node changes its attaching cluster. Even though this mechanism can maximize the stability of cluster topology, this also indicates cluster heads do not consider minimum Iv for some situations in cluster maintenance later. When a mobile node goes into a region not covered by any cluster-head, cluster-head election will be performed again. The new cluster-heads are elected to ensure that the all new cluster-heads are selected based on the minimum Iv. However, this kind of re-clustering completely destroys the current cluster architecture and may require a significant number of messages to be exchanged to build a new one.

7.4.2 Merits of On Demand WCA

1) It gives the option of choosing between multiple metrics.
2) It can flexibly adjust the weighting factors for each metric to adjust to different scenarios.

7.4.3 Demerits of On Demand WCA

1) It is not easy to estimate the distance between two mobile nodes in a practical environment.
2) Cluster-head serving time Tv is not a good indicator of energy consumption.
3) As a node must obtain a lot of information to calculate its combined weight for cluster-head election, the cluster formation requires a longer frozen period of motion.

8. Results

8.1 Assumptions for simulation

For the sake of simplicity we have made many assumptions to make our simulation easier. Some of the assumptions we have made are as follows:

1. Contention time for the medium has not been considered
2. Collision time and retry time of data packets has not been considered
3. Each scenario has been run 10 times to get accurate results
4. The time taken for a data packet to travel from one node to another is 300ms
5. The time taken for a control packet to travel from one node to another is 60ms
6. No. of hops is predetermined by user

8.2 Graphs from simulation:

The following two bar charts have been produced under the following assumption that the cluster head formation process has already occurred and that time is excluded. We have found out that for the different clustering schemes, the bar charts for number of nodes vs. number of data packets and the graph of no. of nodes vs. total data packet sending time is very similar. Therefore we have taken an average of the readings and shown them in one graph.

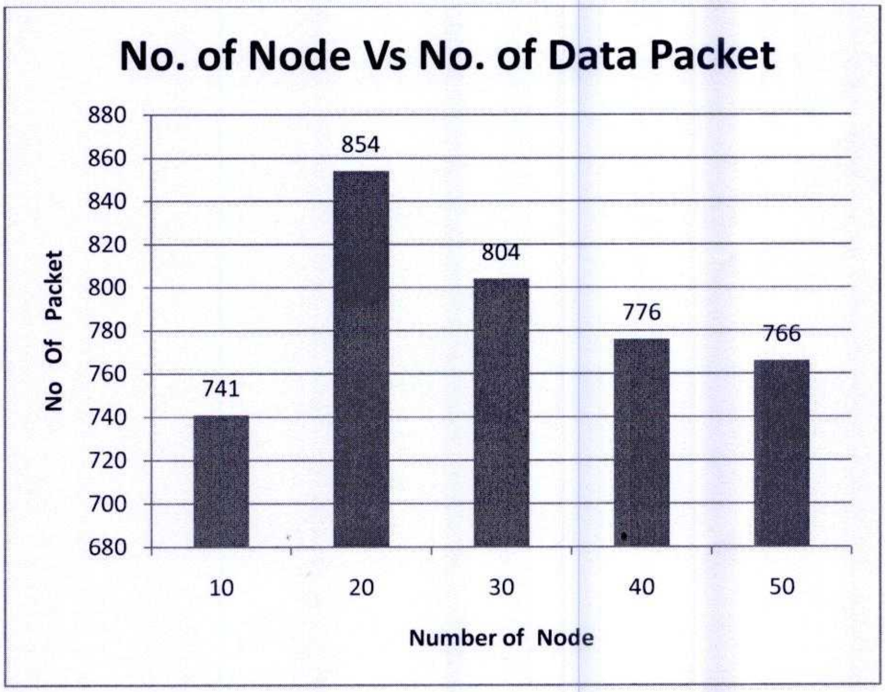

Fig 8.1: No. of Nodes vs. No. of packets

In Fig 8.1 we have a graph of no. of nodes vs. no. of data packets. Here we can say that as the number of nodes increases the number of packets decreases. This is so because, as the number of nodes increases, the number of hops in a cluster also increases. This means that each data packet has to travel a longer distance. And hence within the same amount of time fewer packets are being sent. But the first reading shows an exception to the trend. This is because in the beginning, there is less number of nodes and hence there are fewer packets sent within that time.

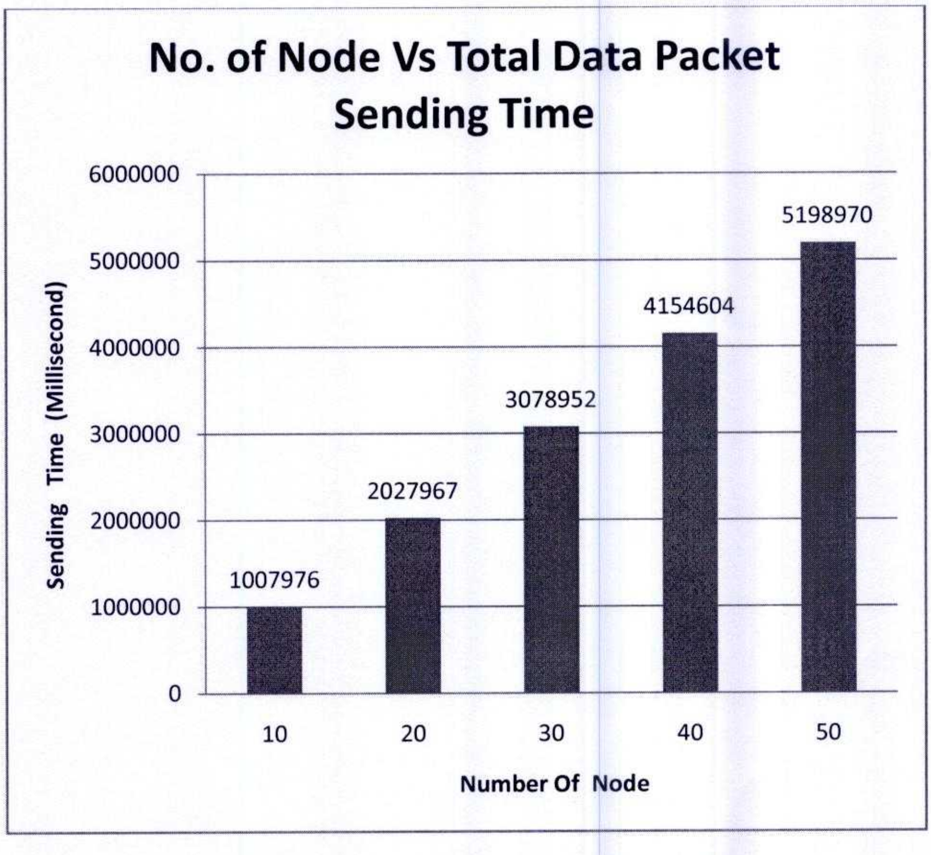

Fig 8.2: No. of Nodes vs. Total Data Packet Sending Time

This is the bar chart of no. of nodes vs. total data packet sending time. We can clearly see this follows a positive linear relation. Here as the number of nodes increases the packet sending time increases. This is because clearly as the number of nodes increases the size of the cluster increases. Hence data packets need more time to travel from sender to receiver.

8.2 Node Vs Cluster formation time

In the following pages we have the bar charts of No. of Nodes vs. Cluster Formation Time. We can see that all of them follow a linear relation. And for all of them as the number of nodes increases the formation time also increases. This is because as the number of nodes increases the computation time to see which nodes fulfill the cluster head criteria also increases.

8.2.1 Ds based clustering

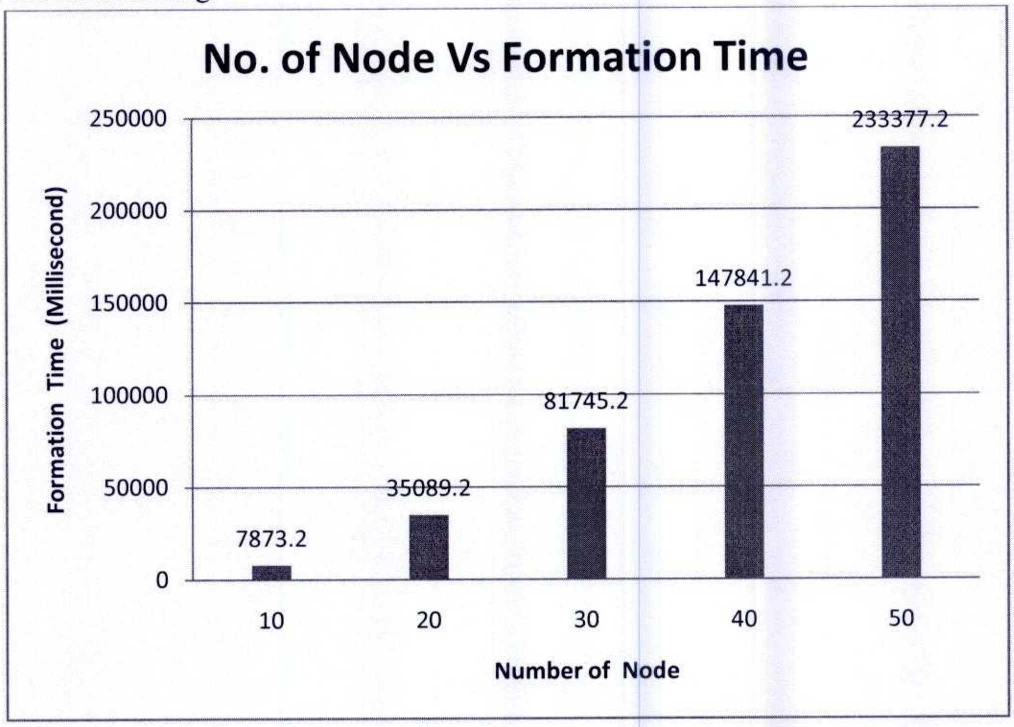

Fig8.3: No. of Nodes vs. Formation Time for DS based Clustering

8.2.2 Energy efficient clustering

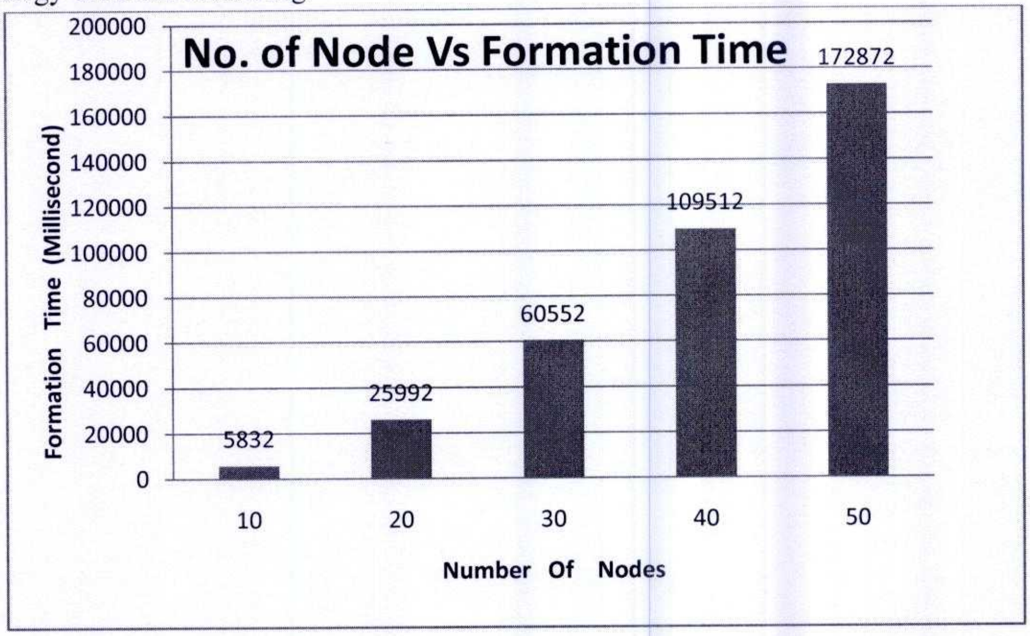

Fig8.4: No. of Nodes vs. Formation Time for Energy level based Clustering

8.2.3 Load balancing clustering

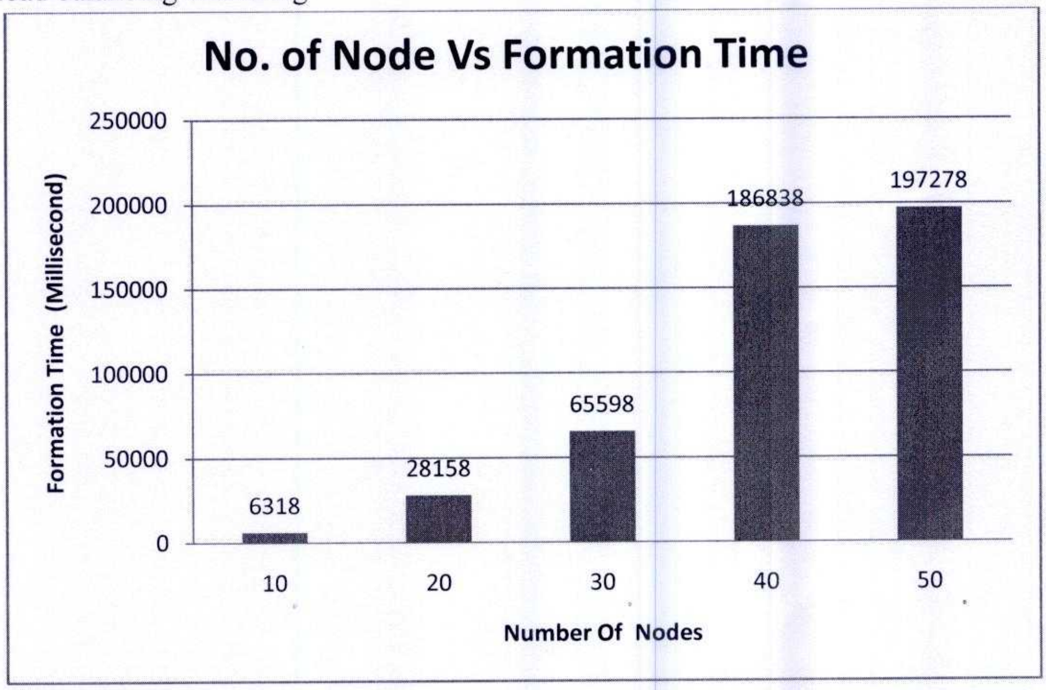

Fig 8.5: No. of Nodes vs. Formation Time for Load Balancing Clustering

8.2.4 Combined metrics clustering

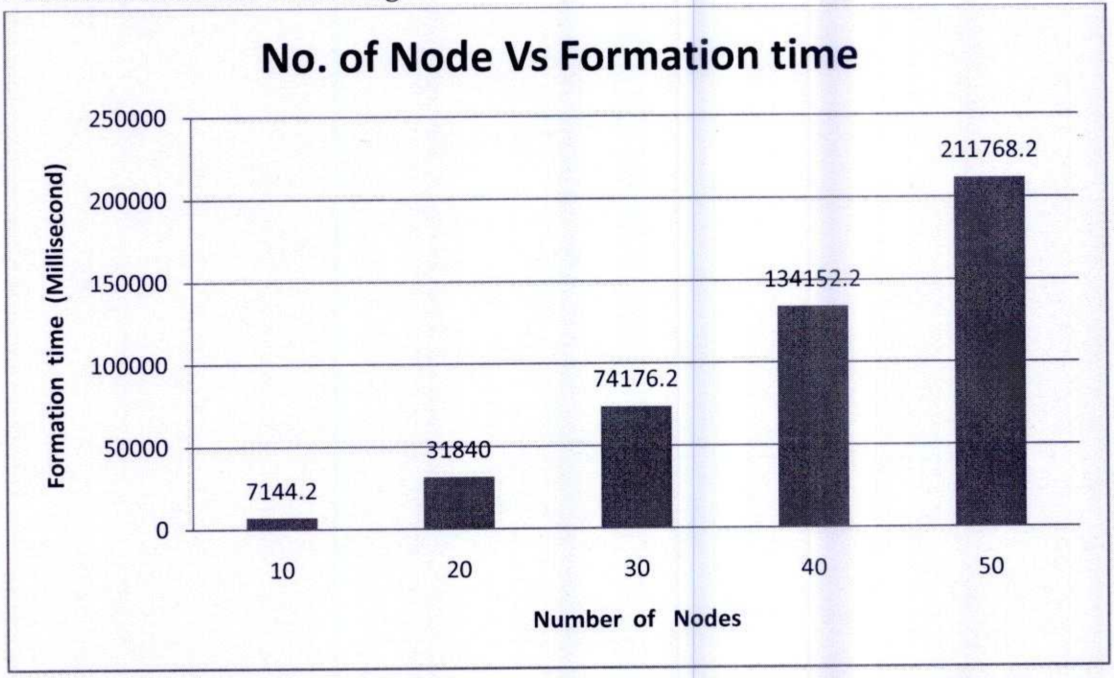

Fig 8.6: No. of Nodes vs. Formation Time for Combined Metrics Clustering

8.3 Node vs. total data sending time

In the following pages the bar charts of No. of Nodes vs. Total Time (cluster head selection + sending time). The trends in these bar graphs are also linear. And they also show that as no of nodes increases the total time increases. As mentioned before the sending time is constant so in each case the total time is just an increment of the formation time by the constant sending time.

8.3.1 DS based clustering

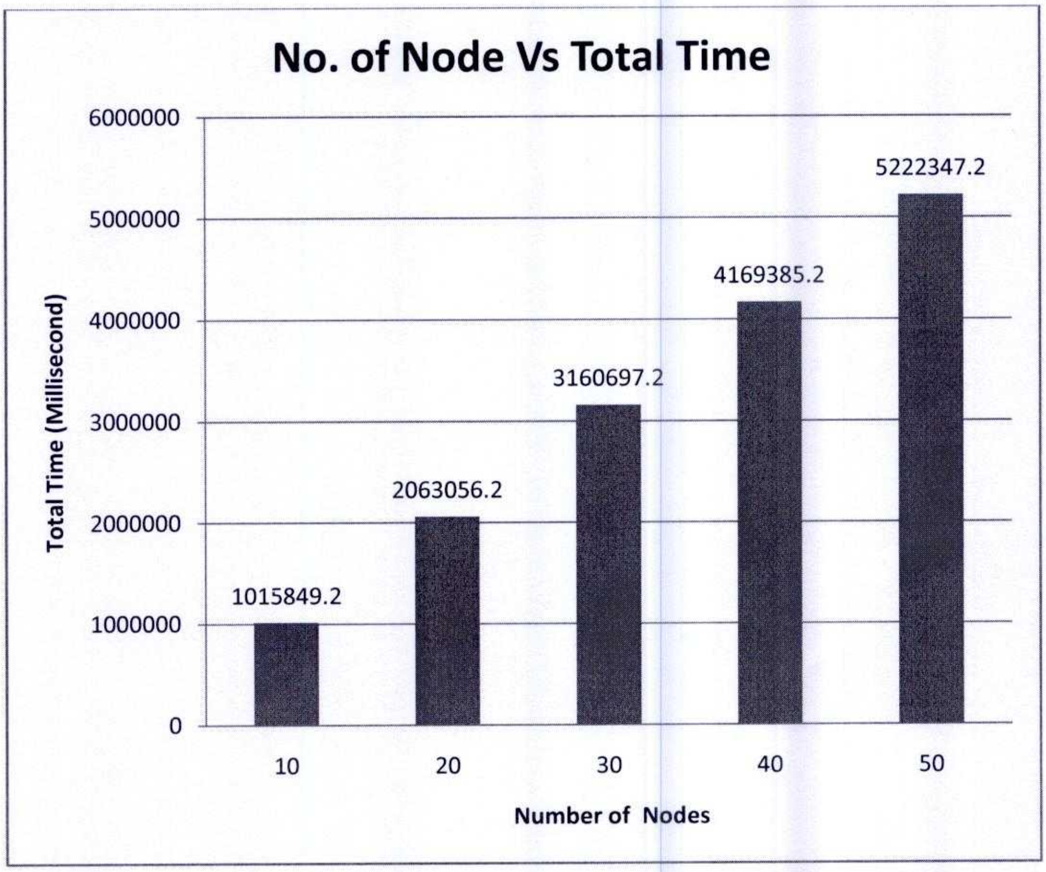

Fig 8.7: No. of Nodes vs. Total Time for DS based Clustering

8.3.2 Energy level based clustering

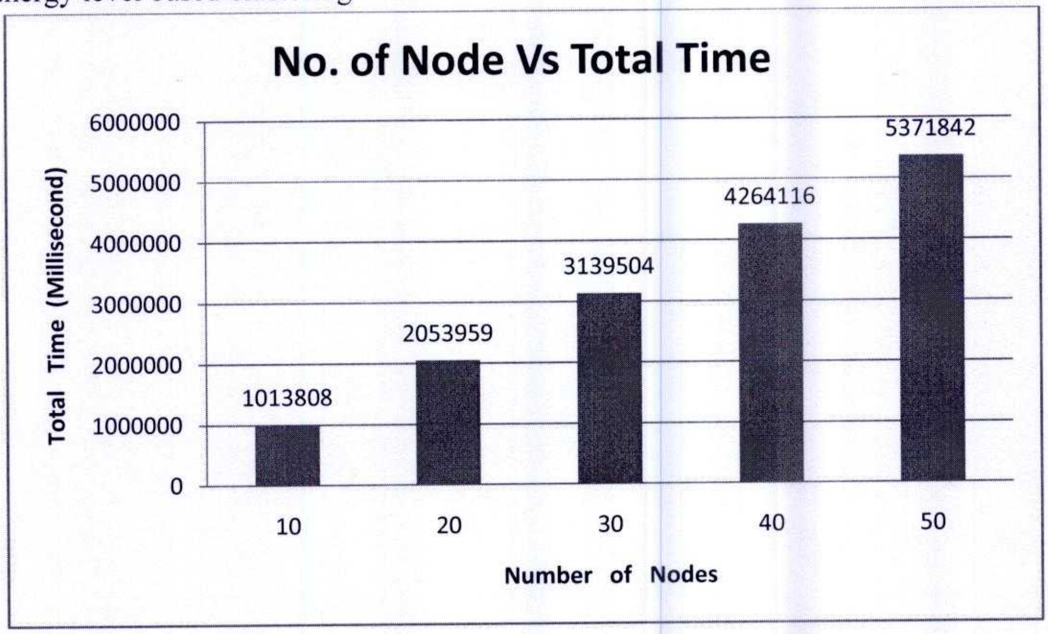

Fig 8.8: No. of Nodes vs. Total Time for Energy Level based Clustering

8.3.3 Load balancing clustering

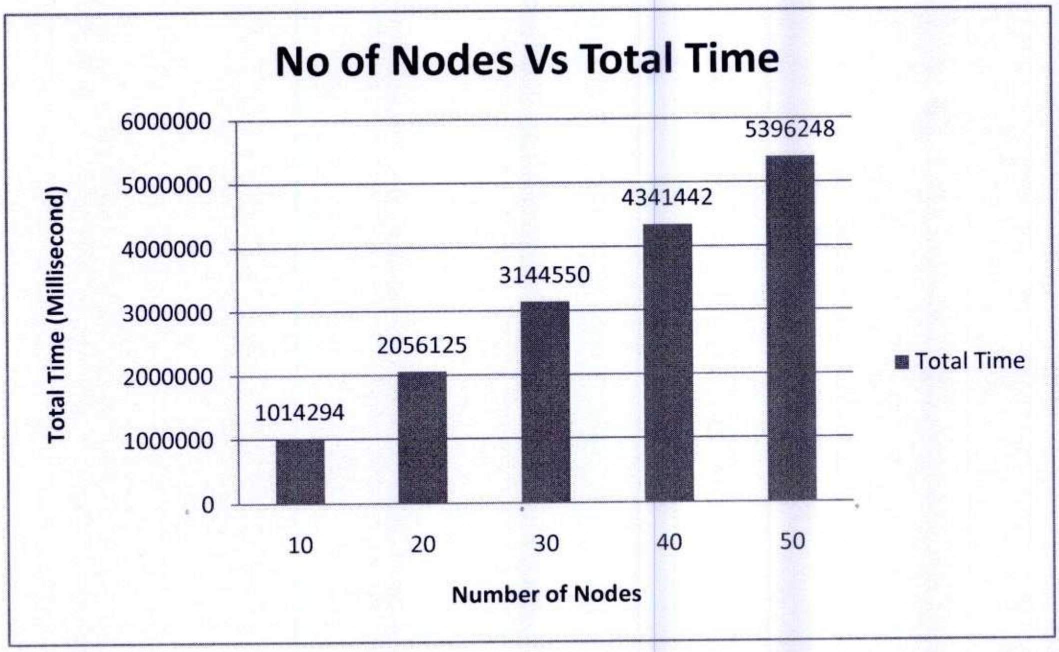

Fig 8.9: No. of Nodes vs. Total Time for Load Balancing Clustering

8.3.4 Combined metrics clustering

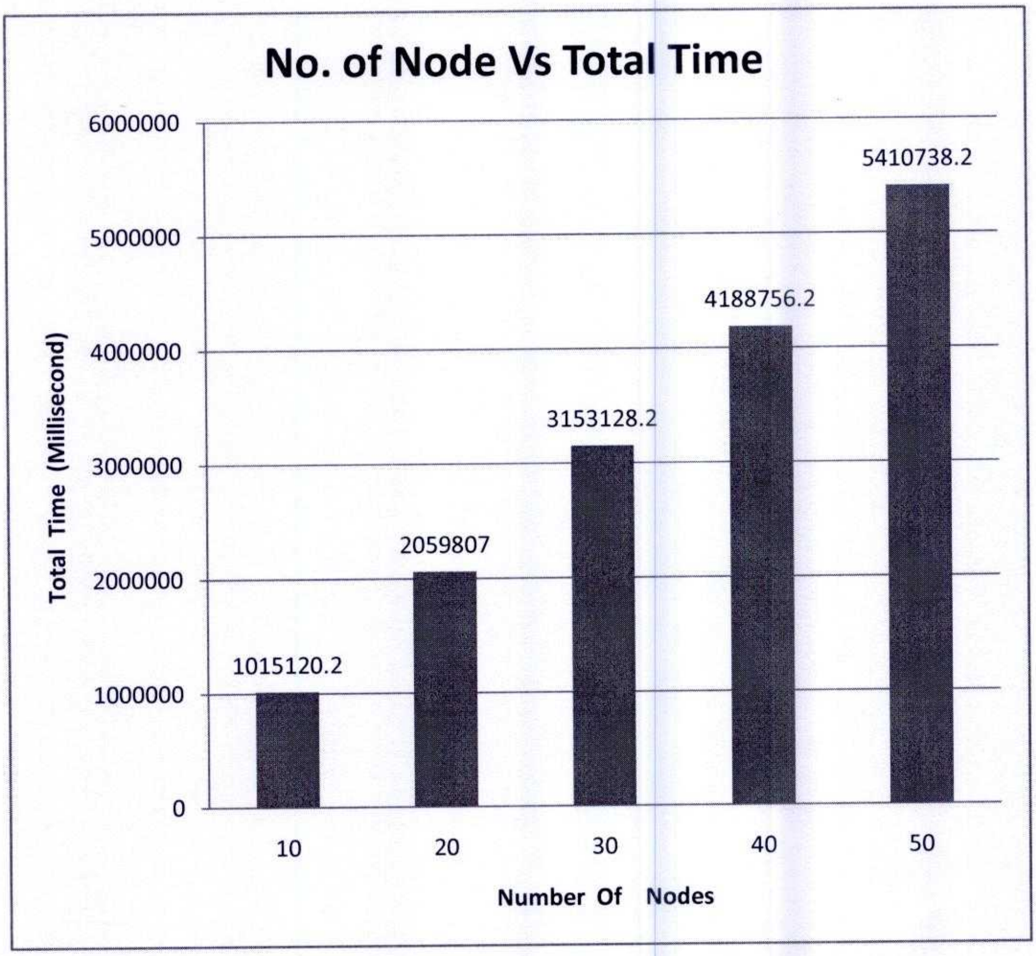

Fig 8.10: No. of Nodes vs. Total Time for Combined Metrics Clustering

9. Comparison

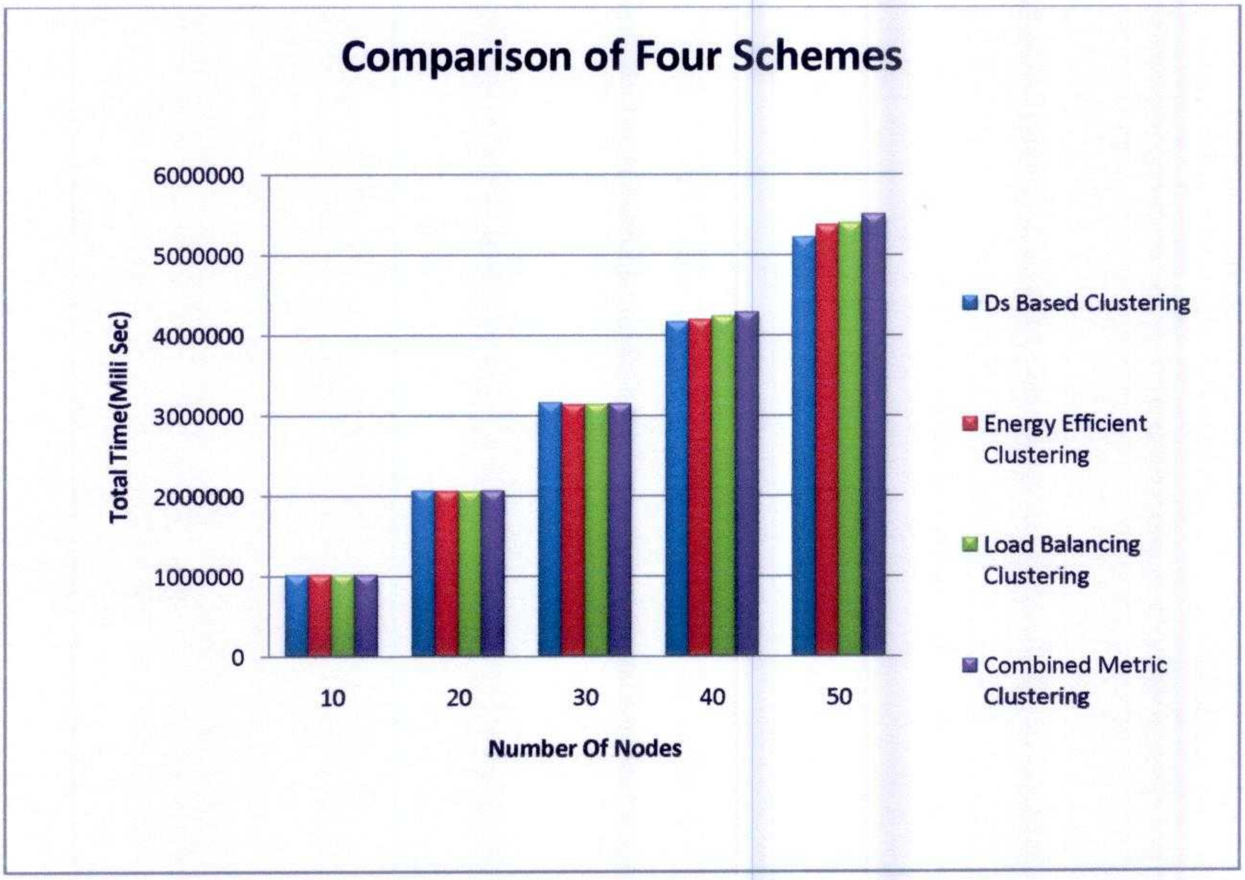

Fig 9.1: Comparison of total time of four clustering schemes

In the comparison between all four clustering schemes we can see that in the initial stages the total time is more or less the same. This is because even though there are differences, they are not that significantly visible in comparison. But when the numbers of nodes are increasing we can see visible differences between each of the clustering schemes. We can see that DS based clustering takes the least time for formation and sending. The values of Energy efficient and Load Balancing clustering are nearly the same. And we can see that the scheme that takes the most time for cluster formation and sending is Combined Metrics clustering. Combined Metrics Clustering takes the longest time because it takes a long time to find out all the parameters (Dv, Pv, Mv, and Tv). And after that, the computation of Iv it also takes quite some amount of time. And finally the transmission of data packets to each node to select the cluster head takes quite some time.

10. Difference between Theoretical Scenario and practical scenario:

Theoretical scenario (Best to worst)	Practical scenario (Best to worst)
DS Based Clustering.	Combined metrics Clustering
Energy Level Based Clustering	Energy Level Based Clustering
Load Balancing Clustering	Load Balancing Clustering
Combined Metrics Clustering	DS Based Clustering

Chart 10.1: Difference between theoretical scenario and practical scenario.

In our result we found that DS Based clustering scheme requires the lowest total (formation+data sending) time and combined metric needs the highest total time. But in real scenario the order is combined metrics clustering, Energy efficient clustering, Load balancing clustering, combined metrics clustering. Though in our result the position of energy efficient clustering and load balancing clustering is same as the real life scenario but Ds based and combined metrics clustering switch their position. This happened because in our simulation we worked only on cluster formation and message sending from one node to another whether all the nodes are situated in same cluster. The cluster formation and the cluster head selection process is very easy for Ds Based clustering. But it takes too much time for combined metric clustering. So, in our graph do based clustering scheme is in number first and combined metric clustering is in number four.

But in real life scenario the work of formation dominating set is not finished yet. In DS based clustering after finishing the cluster head selection process for all clusters the cluster heads will create a direct link to each other which is called CDS (Connected Dominating Set). So it actually requires higher formation time then combined metrics clustering. According to ds based clustering scheme if a node change its position it continuously send beacon message in every τ second. So there will be a huge amount of message transmission on Ds based clustering which we ignored. It requires a large amount of bandwidth. Another big problem of this scheme is if any node tries to entry into the MANET the reclustering process will be done. So we can see in real life scenario it's not a stable scheme where the combined metric clustering scheme is very stable.

11. Conclusion:

In this paper, first the concepts of wireless sensor networks, clustering and the details of certain selected clustering schemes were provided. Then the clustering schemes were compared based on cluster structure stability, the control overhead of cluster construction and maintenance, the energy consumption of mobile nodes with different cluster-related status, the traffic load distribution in clusters, and the fairness of serving as cluster heads for a mobile node. The ripple effects of clustering and frozen period of motion were other important issue to be considered. After all the studies the conclusion that has been reached is that that no clustering scheme is best for a specific scenario. So there is no one scheme that performs well in all scenarios, and they all have their strengths and weaknesses. But, in our opinion On-Demand WCA, is the most effective because it is a very diverse scheme which offers the option of selecting between multiple metrics. So, we hope that this survey is of assistance to anyone who is looking to gain more knowledge on clustering, be that person a student or researcher in networking or an average person looking to study out of interest.

12. Future Work

- We find out the clustering formation time for all schemes. But in our graph we don't put extra time for unwanted data traffic which is generally occur in real time scenario. We will try to find out this extra time in our next simulation program.
- We schemes. Will work with more clustering schemes and compare them.
- Another aim is to propose another clustering scheme which will be more stable then these four.

13. Bibliography

1. A SURVEY OF CLUSTERING SCHEMES FORMOBILE AD HOC NETWORKS by JANE Y. YU AND PETER H. J. CHONG, NANYANG TECHNOLOGICAL UNIVERSITY.

2. A Technical Tutorial on the IEEE 802.11 Protocol by Pablo Brenner, Director of Engineering

3. C.E.Perkins, Ad hoc Networking, Addison Wesley ,2001

4. J. M. Kahn, R. H. Katz, and K. S. J. Pester, "Mobile Networking for 'Smart Dust'," *Proc. 5th IEEE/ACM MOBICOM*, Aug. 1999, pp. 271–78.

5. K. X. Xu, X. Y. Hong, and M. Gerla, "An Ad Hoc Network with Mobile Backbones," *Proc. IEEE ICC'2002*, vol. 5, Apr.–May 2002, pp. 3138–43.

6. R. Rajaraman, "Topology Control and Routing in Ad Hoc Networks: A Survey," *ACM SIGACT News*, vol. 33, no. 2, June2002, pp. 66–73.

7. C. R. Lin and M. Gerla, " A Distributed Control Scheme inMulti-hop Packet Radio Networks for Voice/Data Traffic Support," *Proc. IEEE ICC'95*, June 1995, pp. 1238–42

8. J. Y. Yu and P. H. J. Chong, "3hBAC (3-hop between AdjacentClusterheads): a Novel Non-overlapping Clustering Algorithm for Mobile Ad Hoc Networks," in *Proc. IEEE Pacrim'03*, vol. 1,Aug. 2003, pp. 318–21.

9. B. Das and V. Bharghavan, "Routing in Ad Hoc Networks Using Minimum Connected Dominating Sets," in *Proc. IEEE ICC'97*, June 1997, pp. 376–80.

10. J. Gomez *et al.*, "PARO: Supporting Dynamic Power Controlled Routing in Wireless Ad Hoc Networks," *Wireless Networks*, vol. 9, 2003, pp. 443–60.

11. http://www.wisegeek.com/what-is-a-wireless-ad-hoc-network.htm

12. http://www.wisegeek.com/what-is-a-mobile-ad-hoc-network.htm

13. http://en.wikipedia.org/wiki/MANET

14. http://en.wikipedia.org/wiki/Wireless_sensor_network