Energy Aware Cluster based load Balancing Technique for Cellular Networks
Sumit B. Parmar, Dr. J. W. Bakal
University of Mumbai, Maharashtra, India

Abstract—It is expected that Future Generation of mobile communication networks are heterogeneous, where several Radio Access Technologies will bound with each other to provide services to the users. These networks will be required to serve the different taste of users with different types of accommodation with different Quality of services (QoS) requirements. Due to Limited number of resources in Mobile cellular network provides constraint on many users which are tried to connect calls simultaneously using the similar resources. This results in degradation of Quality of service.

The main aim of system is to propose the technique of reducing the number of calls which is blocked due to various reasons per unit time, with that they have to provide superior quality of services (QoS) with minimum delay. Hence, for better quality of user services and system utilization, different resources sharing schemes have been proposed. Optimal Routing Path (ORP) method has been introduce in collaboration with Cluster-based approach for mobile cellular networks. Cluster head is responsible for computing all possible paths and it has to select least congested path. Major decreasing in the backtracking has been achieved resulting in optimization of the time delay. Here cluster head selection procedure is static. However, if the cluster head is heavily loaded, then there is a delay in the least congested path selection procedure. The proposed work considers dynamic selection procedure of cluster head with the energy as another parameter to select the cluster head for cluster based approach.

Keywords—Cluster, Load Balance, Optimum routing path, Quality of service.

I. INTRODUCTION

Cellular communication is an emerging and upcoming technology that allows different types of users to access information and services electronically, regardless of their geographical location. Because of their flexibility, independence of network, self-motivated and changeable nature of these networks demand new set of networking policies to be applied in order to provide effective node to node communication within the network [9].

In the present situation, the people with the cellular phones will increase day by day. Many users tries to connect calls concurrently, using the same set of resources. Because of limited number of resources, the calls can get rejected or blocked which result in degradation of Quality of Services (QoS). Therefore, reducing the number of calls which is blocked per unit time, with better quality of services (QoS) with least delay, is taken as an one of the major focus of research.

Cells provide radio coverage over a wide range of geographical area. Which are having fixed transceivers and telephones due to which large number of portable devices (e.g., Laptop, Cell phones, pagers, etc.) can communicate with each other anywhere in the network, via base stations.

II. RELATED WORK

A Mobile network is a most emerging and rapidly increasing technology which is based on a independence of network and fastest deployed network. Because of its advance Characteristics, the use of mobile networks is increasing day by day in various application areas towards itself. To increase quality of service in Mobile network different author come up with different techniques for reducing the number of calls which are blocked per unit time, with that they have to deliver improved quality of services (QoS) with smallest delay, which is one of the major focus of research. So to provide better user services and effective system utilization number of resource sharing schemes have been provided.

For the users of mobile networks, the concept of dynamic pricing has evolved to address two conflicting issues simultaneously-handling of congestion and providing improved QoS. In this scheme, the price of calls varies with the constantly fluctuating demand of users. During network congestion, a user who pays much more incentives due to his heavy demand is also responsible for the reduction in access capabilities of the other users as he holds the network resources for long periods. Consequently, user satisfaction cannot be provided.

Another dynamic pricing scheme allows choices for two types of mobile users- priority users and conventional users. The priority users take advantage of getting superior quality of service by paying much more price, whereas the other types of users (called the conventional users) obtain degraded performance service with low and fixed price. So, there is some possibility of ‘starvation’ for the conventional users.
In this paper the author has suggested a new dynamic pricing scheme which is called as Priority based tree generation (PTGM) for mobile networks. Where incoming calls from the various users of different locations are considered to be admitted under a Call Admission Control (CAC) mechanism of a static mobile terminal (MT). These call requests are prioritized and they are placed into appropriate priority levels. Thereafter, a scheduling method is invoked for these call requests and a tree is generated. Then corresponding codes for different possible paths are obtained, resulting in unique path sequences for each call requesting cell with respect to MT. The low-priority call requests can increment their priority values along with increasing their waiting time. This result has drawn a close relationship between the height of the tree and CBR and CDR these two rates of the call requesting cells are varied with the height of the binary tree. As an example, when the height of the binary tree increases, then CBR is also increased and CDR is decreased for that particular radius [1].

Based on GSM architecture a bearer services called GPRS is provided to mobile subscriber with “always on’’ Packet data service on the GSM channels. The problems in GPRS are associated with the assignment of available bandwidth to various services such as data service or voice service. Earlier in GPRS network, data traffic has a less priority over voice traffic. That is the reason GPRS data services are having scares of network resources. Within the congestion period, the problem gets more serious. So the scheme which is based on the dynamic pricing has been introduced by the author according to him the price of the call changes with the change in user demand. The pricing scheme is changed as available traffic channel falls down to specific level which shows congestion situation, because of this the performance model for GPRS network has been developed which is called as Markov Modulated Poisson which increases the performance and also generated the revenue but it does not include computational overhead problem [2].

Optimal QoS can be achieved by using Effective joint Call Admission Control (JCAC) algorithms. For this purpose JCAC uses Multi-dimensional Markov decision process. This results in Better quality of service, with that they can also achieve lower call blocking and call dropping rate. The major advantage to service provider is that it can generate higher revenue during peak hours. The major issue with this paper is reduction of delay was outside the scope of this paper. [3].

Priority based tree generation model (PTGM) [1] is another dynamic pricing scheme which suggesting the use of affixed MOBILE TERMINAL (MT) which can be substitution of mobile switching center (MSC). On basis of optimal path some enhancement is needed. So to minimize the routing path introduction of certain imperative rules based on the radii oriented topology of nodes and well-defined priority factors [4]. When the actual transmission is going on between the nodes, each cluster is busy in calculating least congested ORP in the direction of its upper radius, so here it is not define in advance that the computed path is really reaching to destination or not.

The chances the delay can be increased to a larger extent if the probable chances of backtracking can be increased. With the introduction to clustering [5] methodology. The earlier work in [1, 4] has been extended, among the nodes in network. Here the main goal of this paper is to find out the ORP through upper cluster to the destination with least backtracking [8]. In this case, the cluster head is whole and sole responsible for computation of ORP and the cluster head selection mechanism is static. Once the cluster head is selected, it remains same for all the communication to find their optimum routing path; which is the main job of cluster head. This happens even though, the cluster head is heavily loaded. Due to this time delay increases.

### III. PROPOSED SYSTEM

#### A. Cell and Sector

In Cellular network, by using simple hexagon we can represent the geographical area covered by cellular radio antennas. These areas are known as cells. Cell is a hexagonal shape on a map, it only approximates the covered area. If we want to show a cellular system we want to represent an area totally covered by radio signal, without any gaps.

A cell site gives radio coverage to a cell. The cell site is a position and the cell is a geographical area. Cells can be divided into sectors to make them more efficient and to let them handle more calls. The job of antennas is to transmit signal to each cell. They cover a portion of each cell, not a whole cell. Antennas from other cell sites cover the other portions. The covered area is closely looking like a rhomboid.

![Cell Site](image)

**Right**  
A cell site lies at the edge of several cells, not at the center.  
**Wrong**

Fig. 1. Position of Cell Site

Normally we assume that a cell as the blue hexagon, which is defined by the tower at the center, with the antennae directing in the directions shown by the arrows but in reality the cell is the red hexagon, with the towers at the corners and each can provide coverage in the 120 degree in the direction of arrow. so, to cover the entire cell we require three cell site. A ‘cell’ is the coverage area where the ‘cell site’ is the base station location but they are not the same [10].

© 2015, IJARCSSE All Rights Reserved
The proposed method is grounded on the cellular layout structure of mobile network shown in Fig. 3.

![Fig. 2. Cellular structure of mobile networks](image)

All the calls are going through the MT [1] to reach to the destination. The proposed system is grounded on the activities shown in the following.

**B. Clustering Model**

The nodes are gathered into several levels, which start from level 1 where, the MT [1] is located. Then the cluster is formed by grouping the nodes. The details of cluster construction are given later. To reached to destination the transmission call forward via MT which can travel from one cluster to another cluster.

a. **Configuration of Index numbering**

In a graph for the entire network each node are having left child and right child. We have to give index number to each and every node within the network in such a way that the index number of left child is less than the parent node and index number of the right child is greater than the parent. That is the reason the intermediate nodes assign index number which is in between the left extreme node and right extreme node. If the destination node index is present in between the left extreme node index and right extreme node index then only the call is forwarded by the cluster head otherwise the call is blocked by the same cluster head. Within the index numbering it is possible that a same index number can be assigned to two different nodes in a two different radii. To eliminate such ambiguity we have to combine the radius number with node index and this combination is called as unique identification (uid) for each and every node within the network [8].

The algorithm for above procedure is as follows:

```plaintext
Set (MT, n, r) { /* Pre_Calculation */
for (i=1 to n)
{
Si   Um; // Um stands for status “unmarked”
Ai      ; // Acceptance Rate of (i);
Ri      ; // Rejection Rate of (i);
Li      ; // Load on the node i;
}
IMT= n/2; Ident_no(MT,n/2,1);
/*allocating unique identification to every node*/
Ident_no(node,m,r) {
  if (node==NULL) Return;
  if (lc_node!=NULL)
    Ilc_node=m-1; radlc_node<-r+1; Ident_no (lc_node, m-1,r+1);
  if (rc_node!=NULL)
    Irc_node=m+1; radlc_node<-r+1; Ident_no (rc_node, m+1, r+1);
} /*combination of radi, li can be used as the uid */
calculation of Extreme index node (n,r,k) { /* Extreme node computation */
for (i=1 to k)
  E_Cal(ni,ni,ni->lc,ni->rc); E_Cal(node,temp,temp1,temp2) 
```

© 2015, IJARCSSE All Rights Reserved
If (temp! =NULL) {
    temp1  node->lc; temp2 node->rc;
    Call E_Call(node,node->lc,temp1,temp2);
    node->LeftExtreme<-Temp1;
    Call E_Call(node,node->rc,temp1,temp2);
    node->RightExtreme<-Temp2;
}

Assume that, we have to forward the call to the node duid. The call is transferred if $5: N/2-4<=duid<=5:N/2+4$ is fulfilled otherwise it is blocked [8].

![Fig. 3. uid assignment to nodes](image)

**b. Creation of cluster:**

The number of node is directly proportional to the radii. As a radii increases, number of node increases. At each level the size of the cluster is depend upon the number of node present at particular level. The balance can be maintained between number of cluster and cardinality of the cluster we have to consider the size of cluster is variable [8]. Depending upon number of nodes at level $L$ we have to select size of the cluster which is defined as $(2L+1)$ or $(2L-1)$. The author has chosen such cluster size because in the complete binary tree if the height is $h$ then $2h-1$ [6].

The algorithm based on the above Cluster creation procedure is as follows:

```
Algorithm: /* Cluster Creation */
Clustering (MT,n,r) {
    Set k =1;
    For (i=1 to r-1) {
        Size=2i+1;
        {((size/2) number of unmarked nodes consecutively from left in level [j] + [size - (size/2)] number of unmarked nodes consecutively from left in level [j+1]) € Ck;
            Set statues of all these nodes = Mn (marked);
            Min (index of (size/2) nodes from level [j] existing in cluster Ck) = Ck-head;  k++;
        }
    }
}
```

**c. Formation of various clusters by grouping different nodes:**

i. **Static cluster head Selection:**

The node having lowest index from the lowest radius within the particular cluster is selected as a cluster head for that particular cluster. For example we consider a cluster of five nodes with uid 2:2, 2:3 then the node which is having id 2 is selected as a cluster head [8].

![Fig. 4. cluster Formation](image)

ii. **Selection of the cluster head depend upon its load:**

Each head within its cluster is responsible to take decision regarding Selection of ORP as well as it take active part in call transmission. Sometime it may happen that cluster head may be too busy in transmissions of call, if its lies on very busy route, in this scenario the call transmission decision may get pending, because of which delay can occur.
To solve this problem we have to make cluster head selection dynamic as and when require instead of keeping them static. Any of the nodes can take this responsibility if it lies within the same cluster and that node is ideal or minimum loaded.

To check this we have to find out Threshold load (Lt) value for each node which is defined as summation of load of each node in the cluster to the size of that cluster. It can be calculated as,

\[
Lt = \text{load of node } i / \text{ size of the cluster } k
\]

When the load on the cluster head increase then the threshold load value that is the time to change the cluster head. Load on cluster head in cluster k(Lk)>Threshold load (Lt). Then new cluster head is selected which is less loaded. All the data set which is having information about cluster is moved from old cluster head to cluster head which we have selected currently and the priority of the both the cluster head is also swapped.

So initially the static method is used to select the cluster head as soon as the load on cluster head increase

We used dynamic selection procedure to select the cluster head [13].

iii. **Function of Cluster Head**

The head of cluster is a node which is responsible for maintaining a detail information of every node within its cluster. The data structure is maintained by every node where it stores its detailed information.

1. Rate of Call Acceptance (Ai): Number of call accepted by particular node Ni in a particular time period.
2. Rate of Call Rejection (Ri): Number of call rejected by particular node Ni in a particular time period.
3. Left Extreme Node (LeNi): Extreme node reachable form node i at left side.
4. Right Extreme Node (ReNi): Extreme node reachable form node i at right side.

Unique identification (uid), Radius (radi) and Load on particular node i (Li) =Ai -Ri.

The information about connectivity of the clustered network is maintained by the head of the cluster. The cluster head is not only responsible for calculation of ORPs but also it select one of the optimum path.

C. **Dynamic cluster head selection with energy as another parameter**

Cluster head is responsible for computing all possible paths and it has to select least congested path. Minimization of incurred time delay is possible due to major reduction in the backtracking. Here cluster head selection procedure is static. So if the cluster head is heavily loaded then it occurs delay in the least congested path selection procedure. The proposed system uses dynamic selection procedure of cluster head [13] with the energy as another parameter to select the cluster head for cluster based approach. So if one of the nodes who is cluster head which is on heavily loaded or the load on the node is extended beyond the threshold value then the other least loaded node can be elected as a cluster head with least load one more parameter we have taken in to consideration that is energy.

Energy is also considered, due to the computation burden of the head of the cluster. This is because head of the maintains information about the nodes in cluster. The details of every node in cluster is maintained by cluster head. The changes in details of nodes are always transmitted to head of cluster which then update the details locally in between the nodes in that cluster. Many operations performed by the Cluster head, due to these reasons cluster head must have high energy compared to all other nodes in the cluster.

In wireless network simulation energy is one of the major parameter. The consumption of energy at a particular node, or at a particular component of a node, occurs in various simulation scenarios and of particular interest.

Further, to evaluate the performance of wireless network protocol energy consumption is an important metric.

If a protocol is energy-aware, it means to calculate ORP it depends on the energy level. Without an energy model it is very difficult to simulate the protocol. In ns-2 energy model can be easily implemented to provide the structure of energy source and energy consumption of various device or various events that occur in simulation.

The energy source class is used to represent batteries or power source etc. while the energy consumption model are energy consumed by various component or various events.

Energy source model.
- Power source, Battery etc.
- Device energy consumption model.
- Radio, different events like call transmission, sensing the path etc.

On the each node only one energy source will exist which is used to represent the total energy available at the node whereas multiple device energy model are present on the node, representing different devices or events that consumes the energy and remaining energy is updated to the source [14].

According to this Procedure any node can become cluster head. To keep track of all node within the cluster, an array data structure Cluster_Head[i] is used to store the cluster head index. The Cluster_Head[i], will store the index of cluster head of the i'th cluster. Thus the shifting of cluster head will require modifying this data also, with this we also maintain an Array data structure node[p] for all nodes within cluster, according to their load sort them by using bubble sort algorithm. To transmute data we check the load of first two nodes and compare them for selection of cluster head, if the load of first and second node is same then we have to check the energy of the node and swapping the cluster head if required will take place in each and every cluster from time to time after a certain timer interrupt T (constant).

So initially the cluster head will be statically selected and later during runtime the cluster head may change depending upon network traffic (load).

This method is described by the following Pseudo Code:
Dynamic_Cluster_Selection with energy (Cluster i)
{
    set L = 0;
    for (j=1 to last node in ith cluster)
    {
        Lj = (Aj - Rj);  //cluster load summation of cluster i
        Costj = Lj * Pj;  //Cost Calculation of each node in cluster i
        Costj = 1 / Costj;
    }
    Ci = Cluster head[i];
    When new call comes it calls the following procedure.
    //Take all the node in of cluster i in the data structure array node[]
    //Arrange them in sorting order of load
    for (p=0 ; p<total node in ith cluster-1; p++)
    {
        for (q=0; q<total node in ith cluster -(p+1); q++)
        {
            if (Cost (node[q]) > Cost (node [q+1]))
            {
                temp=node[q];
                node[q] = node[q+1];
                node [q+1] = temp;
            }
        }
        if (Cost(node [0]) = = Cost (node [1]))
        {
            if (energyConsumption (node[0]) > energyConsumption (node[1])
            {
                temp= node[0];
                node[0]=node[1];
                node[1]=temp;
            }
        }
        Swap_Priority (Ci ,node[0])
    }
    Transfer of cluster data (Ci ->node[0])
    Cluster_head[i]=node[0];
}
Return;
}//end of Dynamic_Cluster_Selection with energy

D. Node Priority :
The major function of selecting ORP is performed by the head of cluster so it is necessary that it should be less burdened. To reduce this burden we have to check that less number of calls physically passes via the cluster head. Now to implement this a new concept is developed which is called Node Priority. In this concept nodes other than the cluster head are having low priority and they are responsible for transmitting physical calls whereas the cluster head is having high priority and it is responsible only for computation.

- Node priority is given as follows :
  - Set priority = 1/ \( \alpha \) for nodes other than cluster head
  - Set priority = \( \alpha \) for cluster head, where \( 0 \leq \alpha \leq 1 \) ( \( \alpha \) is constant value)
  - To achieve the required performance the appropriate value is assigned to \( \alpha \).
  - For the maximum traffic value of \( \alpha \) is near to 1 , for steady performance the value of \( \alpha \) near to 0.5 and for limited use its value near to 0 [8].

E. Balancing of load by head of cluster
A request of call is reached at a particular node in the cluster then it is whole and sole responsibility of respective cluster head to calculate cost of congestion for each and every path which leads to the destination and then select one of the path which is less congested.

- For the routing path (\( \mathcal{E} \)) the cost of congestion (Coc) can be computed as follows :
  \( \mathcal{E} = 1/ (\mathcal{O}) \)
  where, \( \mathcal{O} = (\text{Load of node}) \times (\text{Priority of node}) = L \times P \)
  \( \mathcal{E} = (A-R) \times P \)
  Therefore, \( \text{ORP} = \min \{ \mathcal{E}_i \} \) for all \( i \) [8].
F. Procedure of the Call transmission:
At first, request for call from each and every node arrives at the MT. As a request arrives at the MT the corresponding cluster head of MT read the destination node uid. Then it find out ReNi and LeNi of MT node and if LeNi≤uid≤ReNi then calculation of ORP towards the destination is being carried out otherwise the call has been blocked.

The first step is to calculate all possible path towards the destination from lower radius (radii) to its upper radius (radii + 1).

Then the cost of each path is calculated as discussed above.

\[ e_i = 1/ [(Ai-Ri) * i] \]

Above i indicates the node in the radius (radii+i) neighboring to MT node which is available in cluster Ci.

The cluster head Chi pick up the low cost \( e_i \) and transfer the received request which has come from MT. the same set of steps are followed until the intended destination node is reached or the call gets blocked because of the earlier reason that is if LeNi≤uid≤ReNi then Computation of ORP procedure take place otherwise cluster head drops the call [8].

The algorithm for selection of ORP is given as follows.

Algorithm: /* ORP selection process */
Path_Selection (MT, destination_node)

{ Source_node MT;
  Count 0; Set int_node 0;
  While (destination_node!=int_node or Count<Max) { Chead Cluster_head of Source_node;
    Flag forwarding (Chead,Source_node,destination_node);
    If (Flag=1) {
      Block_Request(); break;
    } Int_node=Flag;
  }
  Forwarding (Chead, Source_node, destination_node) { If (Le Source_node <= destination_node <= Re Source_node) {
    Max=0;
    For (all nodes i in next radius in that cluster adjacent to Source_node)
      Cost = (Ai-Ri)* i; If (max<Cost) { Max=Cost; Node =i; }
    If (Max==0) Return -1;
    Ai++; Return (i);
  } Else
    Return (-1);
}

IV. EXPERIMENTAL RESULT
Here we have compare the proposed system with Existing system in terms of Call Delivery Ratio and End to End Delay.

A. Call Delivery ratio
Packet delivery ratio is the ratio of the number of call delivered to the destination. This illustrates the level of delivered data to the destination.

Call Delivery Ratio=Total Number of packet receive / Total Number of packet send.

The results of the proposed model have been plotted in a graph. Here, on the X-axis we have represented the Traffic load in terms of cumulative calls/ 2 second and the Y-axis represented by the corresponding call delivery ratio. We also compare the result of proposed system with the Static cluster head selection mechanism [8] and dynamic cluster head selection mechanism [13].

![Figure 5 Traffic Load vs Call Delivery Ratio](image.png)
Table 1: Analysis of existing and proposed system in terms of call Delivery Rate

<table>
<thead>
<tr>
<th>Sr. No</th>
<th>Data traffic</th>
<th>Dynamic with Energy</th>
<th>Dynamic</th>
<th>Static</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>98.8571</td>
<td>98.0952</td>
<td>99.2857</td>
</tr>
<tr>
<td>2</td>
<td>40</td>
<td>99.5714</td>
<td>99.4286</td>
<td>99.4286</td>
</tr>
<tr>
<td>3</td>
<td>60</td>
<td>99.5714</td>
<td>99.4286</td>
<td>99.4286</td>
</tr>
<tr>
<td>4</td>
<td>80</td>
<td>99.5714</td>
<td>99.5714</td>
<td>99.5714</td>
</tr>
</tbody>
</table>

B. End-to-End delay

End-to-end Delay means the average time taken by a data packet to arrive at the destination. It also includes the delay caused by route discovery process and the queue in data packet transmission. Only the data packets that successfully delivered to destinations that counted and drop packet are not counted. The lower value of end to end delay means the better performance of the protocol.

To compare proposed model with previous two models, we need to consider both the parameter performance and speed for that we have to compute the total delay while transmitting packet from MT to destination node.

Hence, overall delay: Propagation Delay + Computation Delay.

The results of the proposed model have been plotted in a graph. Here, on the X-axis we have represented the Traffic load in terms of cumulative calls/ 2 second and the Y-axis represented by the corresponding time delay (ms). We also compare the result of proposed system with the Static cluster head selection mechanism [8] and dynamic cluster head selection mechanism [13].

Figure 6 Traffic Load vs Delay

Table 2: Analysis of existing and proposed system in terms of Delay

<table>
<thead>
<tr>
<th>Sr. No</th>
<th>Data traffic</th>
<th>Overall Delay</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dynamic with Energy</td>
<td>Dynamic</td>
</tr>
<tr>
<td>1</td>
<td>20</td>
<td>0.117830</td>
</tr>
<tr>
<td>2</td>
<td>40</td>
<td>0.113053</td>
</tr>
<tr>
<td>3</td>
<td>60</td>
<td>0.190399</td>
</tr>
<tr>
<td>4</td>
<td>80</td>
<td>0.558353</td>
</tr>
</tbody>
</table>

V. CONCLUSION

It can be conclude that the Cluster based approach is used to divide all the nodes of cluster with cluster head in each cluster. Initially, the mechanism used for cluster head selection is static due to which delay can occur in the transmission procedure, and then we proposed dynamic cluster selection method, to carry out the smooth functioning of call transmission with ORP computation.

When the cluster head is busy in transmitting call through itself, we have to select least busy node with low energy consumption as a new cluster head to avoid the node overloading and access heating at that time. By using dynamic cluster head selection with least energy consumption, it can result in the increased efficiency of the network with improvement in routing by avoiding congested path selection. It provides significant reduction in backtracking and also reduces time delay.

REFERENCES


