Fuzzy Clustering Algorithm Based on the Time and Surplus Energy Constrain for Ad Hoc Network

Dao-quan Li, Huai-cai Wang, Jun-hu Zhang, Qi-guang Cao
College of Computer Engineering, Qingdao Technological University, Qingdao, China
Email: lidaoquan@sina.com.cn, whckaikai@163.com, caoqiguang163@163.com

Abstract—Mobile ad hoc networks may be composed of a large number of nodes and hence a hierarchical clustered-based structure can be employed to address the scalability issues of the large network. This paper proposes a fuzzy Clustering algorithm based on the time and surplus energy constrain in Ad Hoc network. It defines the fuzzy rules of node’s TTL(time-to-live) and the node surplus energy, it also carries on the fuzzy reasoning, the clustering question is generalized to the question of finding an maximum Independent Set. The simulation reveals that the reconstruction number of cluster per unit time reduces obviously without increasing the calculation and communication cost, meanwhile the performance is improved greatly.

Index Terms—Ad Hoc network, fuzzy reasoning, clustering, On-demand weighted algorithm

I. INTRODUCTION

Wireless ad hoc networks are composed of mobile wireless points, which are used widely to recover in the disaster, war, abrupt affairs etc. Each node in the networks must act as the function of host and router at the same time, because there are no fixed network backbone and topology. There are many special protocols in the networks which are used to find and maintain the proper routers, but their expense can’t be neglected. DSR (Dynamic Source Routing), for example, it works well in the relatively small network, but the utilization of the bandwidth becomes very slow along with the increment of the node number because of the heavy communication cost. Under this case, cluster structure is used in the AD HOC network, which can lower the cost of router, shorten the delay of network. The algorithm of dividing AD HOC network, which can lower the cost of router, shorten the delay of network. The algorithm of dividing AD HOC network, which can lower the cost of router, shorten the delay of network. The algorithm of dividing AD HOC network, which can lower the cost of router, shorten the delay of network.

In lowest-ID algorithm, each node in network obtains a unique ID, each node broadcasts its own ID and exchanges with its adjacent nodes, the node whose ID is the lowest will be selected as th header of the cluster. Highest Degree algorithm draws on the method of selecting router in the internet, the algorithm’s goal is to minimize the number of cluster. Node can obtains the number of its adjacent node, which is called the node’s degree, by exchanging control information. The node which has highest degree among its adjacent nodes will be selected as the header of cluster, the lowest ID node will be selected as the header of the cluster, the lowest ID node will be selected as the header of the cluster, the lowest ID node will be selected as the header of the cluster, the lowest ID node will be selected as the header of the cluster, the lowest ID node will be selected as the header of the cluster.

With the continuous research in AD HOC networks, many clustering algorithm are proposed to construct and maintain the structure of AD HOC. Network structure based on clustering is be beneficial to decrease the cost of route algorithm, flooding and route maintenance, is easy to make control nodes and mobile nodes connect into the wireless channel, to widen network bandwidth, to improve the utility of sharing channel, is helpful for improving the Scalability of network and Qos service support, which is widely used in route, security, network management and Service Discovery. On the other hand, the cost of computing, communication and maintenance caused by clustering algorithm can not be neglected, so we must design reasonable algorithm. The selection of clustering algorithm depends on application requirement, network structure and node’s feature.

The earlier clustering algorithm, such as Lowest-ID algorithm [1], Highest-Degree algorithm [2] (HAD for short), and Adaptive On-demand weighted algorithm (AOW for short) [3], [4], have different research angles of problem, but each takes into account only a certain and particular factor, so there are many restriction for their application.

This work is supported by Qingdao Municipal Science and Technology Commission (NO. 09-1-3-56-jch).
the weighting factors to change the importance of different parameters to adapt to different system environments and application requirements, thereby to improve the adaptability, flexibility and versatility of algorithm. But On-demand-weighted algorithm is higher in computation and implementation complexity than other clustering algorithms which consider only a single factor, also, owing to the node’s mobility, some factors can not be described with accurate values.

This paper gives a fuzzy clustering strategy based on the time and surplus energy constrain in Ad Hoc network with the foundation of earlier clustering algorithm [5]–[8], which apply fuzzy control technology into clustering algorithm, and calculate the link’s TTL(time-to-live) and node’s energy.

II. CONCEPT AND MAIN CLUSTERING ALGORITHM OF AD HOC NETWORK

A. Concept and structure of the Ad hoc network

Mobile Ad hoc network is the connection among the nodes which do not rely on any infrastructure. Wireless Ad Hoc network is composed of a group of mobile terminal nodes with a wireless communication transceiver, these nodes can quickly build up a mobile communication network at any time, anywhere, without the support of current information network infrastructure, each terminal node in network can move freely, and is equal. So Ad Hoc network is a temporary, multi-hop no-center network.

Wireless Ad hoc network is divided into two Categories according to topology: planar structure and hierarchical structure. In planar structure network, the status of all nodes are equal without centralized control, so management of network is simple, in theory, there is no bottleneck problem , it is a relatively robust (robust) network. However, when more nodes each node needs to maintain a large number of routing information, and management overhead, thus contributing less scalability, suitable for small and medium-scale networks. In hierarchical structure network, network is divided into several clusters, each cluster consists of a cluster header and several members nodes, the high-level networks formed by cluster header can also be clustered deeply. Each node in hierarchical network can be selected as a cluster header node using the algorithm based on the network topology changes, cluster header is responsible for forwarding data between groups through the gateway node (gateway node simultaneously belongs to two different cluster), cluster members can not communicate directly with the members of other cluster, they must communicate with cluster headers or gateway nodes. Hierarchical structure can reduce the cost of route discovery caused by topology changes; it has good scalability and strong survivability. Meanwhile, a single point failure problem is the performance bottlenecks. Therefore, it is one of the key issues to how to design a good clustering strategy.

B. Main factors of Ad hoc network design

Multi-hop wireless ad hoc networks have many advantages in application, such as rapid deployment, with self-organizing functions, which are suitable for military tactical and civilian emergency communications environment. But the advantages and disadvantages always complementary, and the advantages in the design become difficult questions or even shortcomings. Being different with wireless networks based on base stations, multi-hop wireless network, each node in multi-hop wireless network can be a combination of mobile devices, is not only a router, but also a terminal, as well as has wireless transceiver, each node communicate with other node through the air link. Many factors, such as the dynamic topology, limited bandwidth, link capacity of instability, limited-energy, wireless spectrum radiation, etc., must be considered when designing network. Generally speaking, the following factors must be considered when Ad hoc network is designed:

(1) Wireless communication technology. Ad hoc network data link layer transmission technology mainly uses the following ready-made digital wireless communication technology, so its performance will be largely limited to the performance of the underlying wireless communication technology, including a possible offer maximum data transfer rate, the transmission delay of end-to-end , network throughput, and so on.

(2) Node density. The node density in the Ad hoc network has a direct impact on the complexity of network topology, the node density, the higher end of the transmission path of the article the more the number of data packets throughout the transmission process by the network topology changes more obvious impact of the road by the control mechanisms will also be more complex.

(3) Node speed. Node speed will largely determine relative changes of topology structure between adjacent nodes, the more higher it is, the more frequent handoff between adjacent nodes, the worse the stability of the logical topology, the heavy calculation of the route and switching loads.

(4) Communication load and communication patterns. Node characteristics and flow distribution will directly affect the network throughput performance, and thus affect the mobile Ad hoc network equipment design, especially for the critical resources needed reserve and control by routing and caching mechanism.

C. Main clustering algorithm of Ad hoc network

The clustering algorithm widely used currently in Ad hoc network mainly includes the lowest ID algorithm, Highest-Degree algorithm and On-demand weighted algorithm, which we mention them earlier. Based on the algorithm above, aiming at their disadvantages, a number of improved algorithms are proposed, which are the Least
Clusterhead Change (LCC) algorithm [9], the algorithm based on node mobility [10], clustering algorithm based on channel access [11], etc.

1) LCC (Least Clusterhead Change) algorithm: With the foundation of lowest ID algorithm, or Highest-Degree algorithm, LCC algorithm mainly considers a maintenance strategy of dynamic link. Algorithm gives network initialization using lowest ID algorithm or Highest-Degree algorithm. When a non-cluster header node moves away from the original cluster, then the node re-finds its neighbor node table, if it finds a unique cluster header, then join the cluster. If there are multiple cluster headers, then it selects a cluster which its ID is the lowest to join. If there is no cluster header, then establish a new cluster and selects him as cluster header. When a cluster moves to other cluster, the two headers completed for being header based on the minimum ID or the Highest-Degree strategy. LCC is a clustering algorithm with a good stability. However, the factors used in this algorithm are relatively too small, only the lowest ID or the highest degree, so the algorithm has a great unfairness. Meanwhile it is a passive clustering algorithm, that is, the algorithm will only be active only when a node joins in a cluster or goes out of the cluster.

2) Algorithm based on node mobility: As the node mobility directly affects the changes of Ad hoc network topology, so Prithwish proposed a clustering algorithm based on node mobility strength, using the signal strength received to measure the node’s local mobility. We can obtain the relative mobility between current node and its neighbor nodes by receiving the follow-up packets send by the neighbor nodes. Initially all nodes are in state with un-clustering, each node periodically broadcasts its mobility measure, the neighbor node will store them in its local neighbor table after receives the mobility measure, and compare with its own mobile measure. If it finds itself with minimal mobility measure, then set itself as a cluster header, otherwise as a cluster member. If a node is a neighbor of two cluster headers, then the node becomes the gateway nodes. The greatest contribution of the algorithm is to propose network node’s local mobility metrics, set the mobility of network nodes as main clustering factors, which create a precedent which set a variety of network elements as the clustering conditions.

3) Clustering algorithm based on channel access: Channel Access-based passive clustering algorithm does not require cluster to send explicit messages, but will attach the relevant state information to data packet, thus reducing control overhead, and has a more stable cluster structure. During process of forming a cluster, each node attempts to access the control channel to affirm that it is the cluster header, a node firstly successfully send a control message among its neighbor nodes, then the node becomes the cluster header, the ordinary node which receives the cluster header’s control message becomes the member of this cluster. After a node become a cluster header and at least has a member node, it will always act as a cluster header until it leave the network or is in failure. When a node wants to join in network, or want to change cluster, it can join a neighbor cluster according to broadcasting message received. This passive clustering algorithm has a large randomness to generate cluster, distribution of the cluster header is not reasonable and not easily be controlled.

III. AD HOC KEY CLUSTERING Technique BASED ON FUZZY CONTROL THEORIES

A. The principle of clustering algorithm

For the clustering ad hoc networks, key consideration is to search the better clustering algorithm. A better clustering algorithm should have certain stability for the node’s movement. That is, when only some nodes move or network topology changes occur slowly, the cluster structure don’t change dramatically, the entire network changes only according to parts of the adjustment in the cluster, so makes the node out of the cluster possibly join in the nearest cluster in the shortest time, while the rest remain unchanged. In particular, if there is already a header in the cluster, the frequent change of the cluster header will lead that cluster header can not effectively control its cluster, lost control on the region. Meanwhile, this will seriously impact on a number of other related protocols (such as resource allocation, routing control protocol), increase overhead of re-cluster, and reduce the channel utilization.

B. Fuzzy control theories

The fuzzy control is composed of fuzzy set and logic. Because the fuzzy concept can be described by fuzzy set, so the description of fuzzy concept can be inverted into the description of the fuzzy set membership function. So we can research them by using fuzzy set theories with regard to the problem lack of accurate mathematics model. According to the operation and the control experience, people summary the control rule using fuzzy condition sentence, and then give the reasoning rules by mathematics processing. So we can make fuzzy decision, complete a control action according to the input fuzzy information, control rule and reasoning rule.

The basic structure of the system are composed of three parts: (1) The rule database: include a series of fuzzy rule; (2) Database(or dictionary): Define a series of membership functions used by fuzzy rule; (3) Reasoning mechanism: obtain proper output or conclusion by reasoning using the fact and rules.

In Ad Hoc network, the node is moving continuously, the power energy of each node also changes continuously, the surplus power energy of the node has relation to the beginning power energy, consuming power energy (node acting as cluster header or cluster member). The link’s TTL of two nodes has relation to do with the distance of the node, the running direction of the nodes. It is very difficult to describe them by normal mathematics resolution expression, so we compute power value of each node, the link’s TTL between two nodes etc., using fuzzy set membership function of fuzzy control techniques [12].
C. Estimate stability model of link based on fuzzy reasoning

Estimate stability model of link based on fuzzy reasoning We will give the TTL model between nodes based on fuzzy reasoning. Figure 1 is the fuzzy membership function of node’s distance, \( \mu_{dist}(d) \), it’s domain is \([0,r]\), and \(r\) represents transmission scope, it includes 5 linguistic terms: Nea, NM, Med, MF, Far, their meanings are respectively near, medium short, medium, medium long, long. Figure 2 is the membership function of node’s speed, \( \mu_{mov}(v) \), it’s domain is \([-v_{max}, v_{max}]\), where \( -v_{max} \) represents the threshold value of speed between two nodes which run at the opposite direction, \( v_{max} \) represents the threshold value of speed between two nodes which run at the same direction, it includes 3 linguistic terms: AwSt, To, The meanings of them are respectively opposite, static, the same direction.

The input data of system is accurate, regarding as a single fuzzy point, from which accurate output data can be deduced. Sugeno fuzzy model is used by system because it need not consume time to un-fuzz up. The classic fuzzy rule in Sugeno fuzzy model is:

\[
\text{If } x \text{ is } A \text{ and } y \text{ is } B \text{ then } L_i = f(x, y)
\]

Where \( A \) and \( B \) is the fuzzy sets of antecedent, \( L_i = f(x, y) \) is the accurate function of consequent. Zero degree Sugeno fuzzy model is used in the system, table 1 shows the rules, where \( L_i \) represents TTL of link \( i \), it’s value is LLMM MSS, represent respectively long, long-middle, middle, middle-short and short, and they all are accurate value, and represented as 1, 0.75, 0.5, 0.25, and 0.

<table>
<thead>
<tr>
<th>( i )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A )</td>
<td>Nea</td>
<td>NM</td>
<td>Med</td>
<td>MF</td>
<td>Far</td>
<td>Nea</td>
<td>NM</td>
<td>Med</td>
<td>MF</td>
<td>Far</td>
<td>Nea</td>
<td>NM</td>
<td>Med</td>
<td>MF</td>
<td>Far</td>
</tr>
<tr>
<td>( B )</td>
<td>Aw</td>
<td>Aw</td>
<td>Aw</td>
<td>Aw</td>
<td>Aw</td>
<td>St</td>
<td>St</td>
<td>St</td>
<td>St</td>
<td>St</td>
<td>To</td>
<td>To</td>
<td>To</td>
<td>To</td>
<td>To</td>
</tr>
<tr>
<td>( L_i )</td>
<td>L</td>
<td>LL</td>
<td>M</td>
<td>MS</td>
<td>S</td>
<td>L</td>
<td>LL</td>
<td>M</td>
<td>MS</td>
<td>S</td>
<td>L</td>
<td>LL</td>
<td>M</td>
<td>MS</td>
<td>S</td>
</tr>
</tbody>
</table>

D. surplus power energy estimate stability model based on fuzzy reasoning

The surplus power energy of the node has relation to the beginning power energy, consuming power energy (node acting as cluster header or cluster member), and it is very difficult to describe this relationship by normal mathematics resolution expression. The following is the surplus power energy model of node based on fuzzy reasoning.

\( \mu_{pow}(p) \), fuzzy degree membership function of node’s surplus power energy, is shown in figure 3, it’s domain is \([0,1]\), where \( t \) is time, represents supporting time of surplus power energy. It includes 5 linguistic terms: SSMMMLL their meanings are respectively short, short-middle, middle, middle-long, long. cost(v), fuzzy degree membership function of power energy consumption of node, is shown in figure 4, it’s domain is \([0,1]\), where 0 represents that the node don’t consume any energy(such as cluster’s member), and 1 represents that the node...
consume the biggest energy (such as cluster’s header). It includes 3linguistic terms: CH, CG, CM, their meanings are respectively energy consumption speed of cluster’s header, energy consumption speed of gateway, energy consumption speed of cluster’s member. Sugeno fuzzy model is used in system, the classic fuzzy rule is:

If \( x \) is A and \( y \) is B then \( T_i = f(x, y) \)

Where A and B is the fuzzy sets of antecedent, \( T_i \) represents supporting time of surplus power energy of node i, it’s value is LMS, represent respectively long, middle and short, and they all are accurate value, are respectively 1, 0.5, 0.

**TABLE II.**

**THE RULES OF SURPLUS POWER ENERGY**

<table>
<thead>
<tr>
<th>rule</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>S</td>
<td>SM</td>
<td>M</td>
<td>ML</td>
<td>L</td>
<td>S</td>
<td>SM</td>
<td>M</td>
<td>ML</td>
<td>L</td>
<td>S</td>
<td>SM</td>
<td>M</td>
<td>ML</td>
<td>L</td>
</tr>
<tr>
<td>B</td>
<td>CH</td>
<td>CH</td>
<td>CH</td>
<td>CH</td>
<td>CH</td>
<td>CG</td>
<td>CG</td>
<td>CG</td>
<td>CG</td>
<td>CG</td>
<td>CM</td>
<td>CM</td>
<td>CM</td>
<td>CM</td>
<td>CM</td>
</tr>
<tr>
<td>( T_i )</td>
<td>S</td>
<td>S</td>
<td>M</td>
<td>M</td>
<td>s</td>
<td>s</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td></td>
</tr>
</tbody>
</table>

**IV. FUZZY CLUSTERING ALGORITHM BASED ON TIME AND SURPLUS POWER ENERGY CONSTRAINT**

Weighted heuristic algorithm determines each node’s corresponding weight based on the degree which node is suitable for a cluster header, the adjacent node which has the highest weight becomes a cluster header, the node which has the lowest ID will be the cluster header when all the nodes have the same weights. For example, we can give weights to the node according to the node’s moving speed, the faster, its weight is higher, the lower, its weight is lower. Therefore, this algorithm can be regarded as the lowest mobile clustering algorithm. This algorithm can significantly reduce the number of cluster when a node has a strong mobility, but its disadvantage is that the calculation of cost of cluster header is heavy due to frequent updates of node’s weight, meanwhile, some parameters, such as throughput, latency, can not be determined by an effective optimization method. It also does not take into account the system load balance and node energy depletion and other issues, which is a big problem for the limited computing power and energy for wireless ad hoc networks. On this basis, on-demand-weighted algorithm was proposed [5], the algorithm takes into account a variety of factors, such as node degree, node transmission power, node mobility and node surplus energy, whatever, they can be adjusted according to specific Environment. And selection of the cluster header only depends on the exchange of neighbor message, in other words, selection of the cluster header is local. On-demand weighted clustering algorithm takes comprehensively into account a variety of factors, overcome the other drawbacks of other algorithm which only consider a single factor. In practice, we can increase or decrease the impact of factors to improve the algorithm, we can also adjust the importance of various factors, so the algorithm has a strong versatility and flexibility.

But in this algorithm, each node must know the information of all the other nodes, each node broadcasts circularly itself information, one cluster is only grown each time, and then each node broadcasts again itself information. There are many calculation work and communication cost in this algorithm, although threshold value is taken for load balance, but experiment reveals that it is not ideal. In this paper, fuzzy control technique is applied into Weighted Algorithm on Demand, all of the weighted factor, TTL of links, and surplus power energy of nodes will be fuzzed, this makes clustering algorithm in AD HOC having a better performance.

**A. Algorithm principle**

Definition 1 For a given graph \( G(V, E) \), \( V, E \) represent respectively vertex and edge, \( S \subseteq V \), if \( \forall v_i \in V - S, \exists v_i \in S \) then \( (v_i, v_j) \in E \), we call that \( S \) is a dominating set. When whichever node is deleted from \( S \), \( S \) is no longer a dominating set, then we call that \( S \) is a minimum connected dominating set. If added some infinite nodes into \( S \), and the nodes in \( S \) form a connected sub-graph, then \( S \) is called as a connected dominating set. Given that \( U \) is the connected dominating set of graph \( G(V, E) \), if \( V_i \in U \), \( U - \{V_i\} \) is no longer a connected dominating set, then We call \( U \) is minimum connected dominating set of \( G(V, E) \).

Definition 2 Set \( M \) is a sub-set of set \( G \), \( \forall x \in M, \forall y \in M, x \) and \( y \) is not adjacent, then \( M \) is a Independent Set of \( G \). If we added a node in \( G-M \) into set \( M \), then \( M \) is not a Independent Set, and then \( M \) is a maximum Independent Set. In graph theory, clustering question can come down to the question of finding a maximum Independent Set.

Definition 3 For \( \forall v_i \in V \), the number of edge which connect to the vertex \( v_i \) is degree of \( v_i \), and written as \( sd(v_i) \).

Definition 4 If \( V = \{v_1, v_2, \ldots, v_n\} \), then the sequence \( \{sd(v_1), sd(v_2), \ldots, sd(v_n)\} \) is called as node degree sequence of \( G \).
Definition 5 if $\forall v \in V$, for a given time $\delta$, the number of edge which connect to vertex and time $t \geq \delta$, is called as degree based on time $\delta$ constrained by $v$.

Definition 6 Given $V=\{v_1, v_2, \ldots, v_n\}$, the corresponding sequence $\{sd(v_1), sd(v_2), \ldots, sd(v_n)\}$ is called as $\delta$ — degree sequence of graph $G$.

Definition 7 For $v_i \in V$, available energy of $v_i$ is called surplus energy of node $v_i$ and written as $ed(v_i)$.

Definition 8 Given $V=\{v_1, v_2, \ldots, v_n\}$, the corresponding sequence $\{ed(v_1), ed(v_2), \ldots, ed(v_n)\}$ is called as node’s available energy sequence of graph $G$.

By analyzing the principle of clustering algorithm and the experiment result, We discover that the overwhelming majority clustering algorithm [13]—[17] only use a certain parameter of a node, such as node’s connectivity, minimum ID or surplus energy etc., for example, the node whose connectivity degree is maximum will be chosen as the cluster header in maximum connectivity degree algorithm. The first, the sequence $\{sd(v_1), sd(v_2), \ldots, sd(v_n)\}$ changes continuously its order because of the highly dynamic topology in AD HOC, this makes relationship between header and members also is changed continuously. The second, if the energy of cluster header is too small, then the relationship between header and members will be changed continuously. This paper proposes a fuzzy clustering strategy based on the time and surplus energy constrain in Ad Hoc Network.

B. The basic idea and description of algorithm

The algorithm takes into account following factors: the frequent changes in the group arising from node mobility, the rate of change associated with the node, as well as the node energy and so on. The link’s TTL between two nodes is concerned not only with the distance between two nodes, also with two node’s speed and direction of movement-related; Node’s surplus energy is concerned not only with TTL of node, but also with the time as a Gateway, cluster header or cluster member, this relationship is very difficult to describe using conventional mathematical models. The fuzzy logic can give a good description for it by the reflection from a space to another space using a specific rule, so we can describe the link’s TTL and node’s surplus energy with fuzzy control model. Based on the traditional on-demand weighted algorithm, we take into account the link’s TTL and node’s surplus energy to select a cluster header, so the cluster header can survive for a longer time. The following is an example of cluster initialization.

The computing formula of node’s weight is:

$$w = a_1 \times d + a_2 \times l + a_3 \times m + a_4 \times t$$

where $d$ is the difference between its degree and its ideal degree for each node; $l$ is each node’s TTL based on the fuzzy rules above; $m$ is node’s mobility based on the moving speed; $t$ is the surplus energy based on the fuzzy rules above. $a_1, a_2, a_3, a_4$ are weight factors, which indicates the importance of the these parameters in network, in accordance with specific needs for a actual network, more important parameter is, the greater the weighting factors is, but it meets $a_1 + a_2 + a_3 + a_4 = 1$.

Firstly, algorithm selected the smallest weight node as a cluster header, without loss of generality, we set it as $i$, meanwhile all the neighbors of node $i$ have also been classified into his cluster. Secondly, , select a node from the the remaining un-clustered nodes , which its weight is the minimum, and it is adjacent with neighbor nodes of cluster $i$, in the same way, without loss of generality, assume that $j$ is adjacent nodes of $i, k$ belonging to a neighbor node set of $j$, but $k$ does not belong to neighbor set of node $i$, we choose the smallest weight node $k$ as the cluster header of next cluster, and so on , all the nodes are divided into a cluster. Thirdly, if these cluster’s header connect each other, then connect two cluster by cluster header; If the cluster header and cluster header has no connection between, that is, the two clusters can not communicate with each other after clustered, then we have another work to do. Because it is a connected graph, then the two clusters must have connection between two ordinary nodes, we can choose both nodes as cluster header respectively. And so on, we can have connected graph of a maximum independent set (also known as the minimum connected dominating set), then clustering has been completed.

The distributed clustering algorithm with pseudo-code C is as follows:

```
for (j ∈ N(i)) /* N(i) is 1-jump neighbor’s form of node i */
  if (i.td > j.td ∧ i.ed > ε) {
    i.role = Head;
    i.send clusterhead message to j; /*Cluster header i
        send message to node j*/
    return;
  }
  else if (i.td == j.td ∧ i.ed > ε ∧ j.ed > ε) {
    if (i.ed > j.ed)
      i.role = Head;
      j.role = Head;
  } else /* select a node whose surplus energy is the biggest so that it can be a cluster header */

When receiving clusterhead from j {
  j.role = Head; /* mark node j as a cluster header */
  i.send Join message to j; /* node i sends message to node j who is cluster header in a cluster*/
  return;
}
When receiving Join(j,k) message from j {
  j.role = member; /* mark node j as a member */
  if (m ∈ N(i) ∧ m ≠ i)
    if (m.td > i.td ∧ m.role == member)
      i.send clusterhead message to m; /* node i
        send cluster header information to node m */
  return;
```
Gateway node selection:

\[
\text{if (} i, j \in \text{MaximumIndependentSet}(N)\text{) /* node } i, j \text{ are the nodes of largest independent set */}
\]

\[
\text{if (} i \text{ to } j \in e\text{) /* nodes between } i \text{ and } j \text{ are not connected */}
\]

\[
\text{for (} k \in \text{clustering}(i)\text{)}
\]

\[
\text{for (} m \in \text{clustering}(j)\text{) /* nodes } k \text{ and } m \text{ respectively belong to the two clusters whose cluster header is } i \text{ and } j\text{ */}
\]

\[
\text{if (} k \text{ to } m \in e\text{) { /* nodes } k \text{ and } m \text{ connect */}
\]

\[
K.\text{role} = \text{Head};
\]

\[
m.\text{role} = \text{Head}; /* nodes } k \text{ and } m \text{ join in a great independent focus */}
\]

V. THE EXPERIMENT ANALYSIS

We compare three kinds of algorithm [18]–[20]: Adaptive On-demand weighted algorithm, Highest Degree Algorithm (HDA for short), Fuzzy clustering algorithm based on the time and surplus energy constrain (FCATE for short). We place 100 nodes randomly in a 100 x 100 scope, the node transmission distance is r, and \( r \in (0,100) \). The simulation experiment results are shown in figure 5 and figure 6.

![Figure 5. membership function of node's speed](image1)

Figure 5. membership function of node’s speed

Figure 5 reveals that the cluster produced by FCATE is the most stable, and in figure 6, accumulation times of clustering based on fuzzy control technique is the most, and its curve locates at top of the graph. This is the reason that we take into account the parameter of surplus energy for each node, although accumulation times of clustering is most, but network dominating set is the least, and is the most stable, and network has a better performance accordingly. In general, the changing times of cluster header in FCATE are fewer than AOW and HDCA, so the total structure stability is also obviously better compared with the other two kinds of algorithm.

VI. CONCLUSION

Fuzzy control technique is applied into FCATE in this paper, and its application is discussed. The corresponding fuzzy membership degree functions are defined for the link’s TTL and node’s surplus energy, so as to predicting a more stable network link. It completes a fuzzy rule database to produce the fuzzy cost for link and node’s surplus energy, also proposes a fuzzy clustering algorithm based on the TTL of link and node’s surplus energy, the changing times of cluster header in this algorithm is fewer than AOW and HDCA, the system using FCATE has more stability and fewer cost.

REFERENCES


Dao-quan Li received his MS degree in electronic engineering from x’ian Jiao tong university, x’ian, China, in 1994, his BS degree in electronic engineering from Shandong industry university, ji’nan, China, in 1991, now is studying for his doctor’s degree in computer science at Shandong university of science and technology, Qingdao, China. He is currently an associate professor of computer science at Qingdao Technological University, Qingdao, China. His current research interests includes computer network, E-commerce, enterprise information system etc.

Huai-cai Wang received her BS degree in computer science from Qufu Normal university, Qufu, China, in 2007, now is a postgraduate at Qingdao Technological University, Qingdao, China., her current research interests focuses on the ad hoc network.

Jun-hu Zhang received his PhD degree in computer science from Peking University, Beijing, China in 2006, his MS degree in computer science from Tongji University, Shanghai, China in 2001, and his BS degree in mathematics from Qingdao University, Qingdao, China in 1997. He is currently an Assistant Professor of Computer Science at Qingdao Technological University, Qingdao, China. He was a Post-doctor at LIAMA (Sino-French Laboratory for computer Science, Automation and Applied Mathematics) in the institute of Automation, Chinese Academy of Science from 2006 to 2008. His current research interests are on ad-hoc networks, wireless sensor networks, data grids, distributed database systems, peer-to-peer systems, embedded systems.

Qi-guang Cao received his BS degree in computer science from Qufu Normal university, Qufu, China, in 2007, now is a postgraduate at Qingdao Technological University, Qingdao, China, his current research interests focuses on the ad hoc network.