ATLANTIS PRESS
Advances in Intelligent Systems Research (AISR), volume 151
2018 International Conference on Computer Modeling, Simulation and Algorithm (CMSA 2018)

Ad-hoc Network Clustering Algorithm Based on Node Data Value

Jingmin Tang and Zhangbao Gao*
Kunming University of Science and Technology , Faculty of Information Engineering and Automation/ China
*Corresponding author

***Abstract*—Ad-hoc network is a repeatedly jumping, nothing central, and self-organizing network. the entire network has not stable infrastructure, consisting of mobile nodes. According to the problem that cluster Ad-hoc network cluster head selection mechanism is not reasonable and the cluster heads, and cluster members in sending data process transmit the data that is too large or too small to node failure. So, we proposed clustering algorithm of Ad-hoc network based on node data value, The algorithm chooses cluster heads in different clusters according to the data value, node energy value, optimal node degree and node average distance, the cluster head collects the date of cluster members and integration it to go to the next step Communication and other work. In contrast to simulation and comparison experiment of the other three traditional algorithms, the results indicated that applying the algorithm of Ad-hoc network in cluster number are optimized, the property in repeatedly into the cluster and rule setup date has improved, and in a certain extent it prolongs the network life cycle, in the meantime it improves the stability of the cluster structure.**

Keywords—Ad-hoc network; cluster; node data value; stability

I. INTRODUCTION

Ad-hoc network is a repeatedly jumping, nothing central, and self-organizing network [1-2]. the entire network has not stable infrastructure, and each node is mobile. It brings a series of inconveniences in Scalability and management because of its characteristics. In order to improving the scalability of Ad-hoc networks and achieving the network loading with balance, we often use the clustering algorithm for the network, by making use of the cooperative work of cluster heads and cluster members to improve the overall network property. At present, the most common clustering algorithm has the smallest ID algorithm, the highest node degree algorithm and the WCA algorithm. The minimum ID algorithm (LID)was introduced in literature [3], which is easy to come true. The replacement of cluster heads is slower and maintenance the cluster of spending is smaller. However, nodes with smaller ID consume a great deal of energy and shorten the cluster heads Network life cycle. Literature [4] introduced the highest node degree algorithm (HD), which has fewer clusters and reduced grouping delay. However, fewer clusters lead to reduce channel utilization factor. Literature [5] introduced the WCA algorithm, which takes into account various factors and calculates the corresponding weights, selects the cluster heads and divides the cluster structure of the network nodes by weights. In WCA algorithm, the clustering process needs to know the weight of the nodes in advance and extends Network cycle, making the transmission of information go higher, while increasing the node's energy consumption.

This paper presents a novel node data-based clustering algorithm (NDVA), based on the existing problems of traditional clustering algorithms. Its lead to cluster head better stability, lower update frequency, load balancing is high, the network life cycle is longer, while effectively avoiding the node transmitted excessive data to result in consuming excessive energy and data is too small lead to the unrealistic problem of transferring data, basing on node data values, node energy values, optimal node degree and node average distance Cluster head.

II. NDVA ALGORITHM DESIGN

The algorithm is an improved algorithm based on the traditional algorithm. The algorithm considers four factors that the node data value, the node energy value, the optimal node degree and the average node distance in the selection of cluster heads. and each node gives its own weight [6], by comparison the weights to select the cluster head, and then generate the entire cluster.

The total of flow chart of NDVA algorithm design is shown as in Figure. 1:

NDVA clustering algorithm specific process is as follows:

Step1 During the network initialization, each node in the undecided state obtains the number of neighbor nodes' information by periodically sending the probe packet information as its node degree d_i;

Step2 Each node calculates the difference value D_i between its degree and optimal node degree α;

Step3 Each node calculates its node data value V_i;

Step4 Each node calculates its energy value E_i;

Step5 Each node calculates the average distance $\overline{Si}$ from all the neighboring nodes to this node.

Step6 Calculate the combined weight of each node $W_i=\omega_1 V_i+\omega_2 D_i+\omega_3 E_i+\omega_4 \overline{Si}$, the values of ω_1, ω_2, ω_3 and ω_4 are constant in a given environment, and $\omega_1+\omega_2+\omega_3+\omega_4=1$;

Step7 Node are divided into clusters by comparing with the weight of the neighboring nodes, the smallest weight node as cluster head, it broadcast messages to the neighbor nodes to declare themselves as cluster heads. In the same weight under the condition, The node, having the smallest ID, will become

the cluster head first priority, at the same time, it accept the neighbor nodes to become the cluster member of the cluster head, the cluster members can modify their own list, no longer participate in the election of cluster head; if the node is an isolated node, the node will be clustered alone;

Step8 Repeat the above steps until all the nodes are in the range of the dominant.

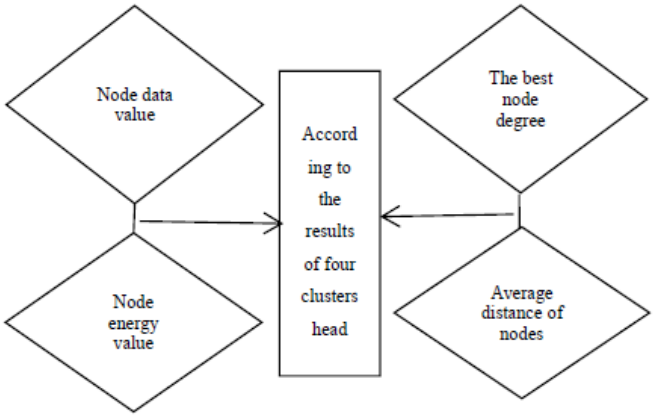

FIGURE I. NDVA ALGORITHM DESIGN PROCESS

A. Calculation of Node Data Values

Node data value is composed of two parts: data size value and data consistency value [7]. When node transmits data, the node should give full consideration to the size of the transmitted data value. When the data value is too large, the energy consumption of the node will increase, which may lead to invalid for node. When the transmitted data value is too small, node will choose to drop a packet to save itself energy [8], which eventually it leads to a shortage or untruth of transmitted data. Therefore, we not only should take full account of energy consumption, but also ensure that the size of the transmission data does not exceed a reasonable limit, which is the node data value. Supposing the gate threshold of node in the transmission of data is V_s, through the node M_i to M_j data D, the data size value N_{ji} is:

$$N_{ji} = \begin{cases} D/V_s, & D \le V_s, \\ 1 - \dfrac{D - \mathrm{Vs}}{\mathrm{Vs}}, & D \ge V_s, \end{cases} \tag{1}$$

In a cluster, it calculates the data size value that node M_i transmits from to all its neighbor nodes and determines the average value, and finally get the data size value of node M_i:

$$N_{ji} \frac{\sum_{i=p} N_{ji}}{k} \tag{2}$$

P is the set of nodes in the entire clusters: $P=\{M_1,M_2,M_3,\ldots,M_i,\ldots,M_j\}$.

Suppose node M_j collecting the number of consistent data is X_j and collecting the number of inconsistent data is Y_j, then the data consistency value Q_j of node M_j is:

$$Q_j = \frac{X_j - Y_j}{X_j + Y_j}, \tag{3}$$

According to the node data size value and data consistency value, then t data value V_j of he node M_j is:

$$V_j = N_j + Q_j, \tag{4}$$

B. Calculation of the Optimal Node Degree

In Ad-hoc networks, the size of clusters is constrained by many conditions. Cluster head node coverage is too large to lead to the network occur congestion, while the coverage is too small, the network bandwidth utilization is low. In order to obtain the appropriate throughput, the conclusion in reference [9]: if K_1 and K_2 represent intra-cluster communication bandwidth and inter-cluster communication bandwidth respectively and N represents the number of nodes, the optimal node size is:

$$\alpha = \frac{K_2}{K_1}\sqrt{N} \tag{5}$$

The nodes studied in this paper are equal quality nodes. The intra-cluster communication bandwidth and the inter-cluster communication bandwidth are the same, so the optimal node degree is $\sqrt{N}$.

C. Calculation of the Average Distance Between Nodes

Calculate the average distance from a node to all its neighbor nodes. The design of Reference [10] uses the relationship between the strength of the transmitted signal and the strength of the signal received by the MAC layer:

$$P_i = P_j \left(\frac{h_i h_j \sqrt{G}}{s} \right)^2, \tag{6}$$

P_i represents the strength of the transmitted signal, P_j represents the strength of the signal received by the medium access control layer, s represents the distance between the sending node and the receiving node, G represents the gain between the two nodes, h_i represents the height of sending antenna, h_j represents the height of the receiving node, the distance between sending node and receiving node is:

$$s = \sqrt[4]{\frac{P_j G (h_i h_j)^2}{P_i}}, \tag{7}$$

The average distance from the last available node to all its

neighbors:

$$\overline{Si}\ \frac{\sum_{i=1}^{\alpha_i} s_i}{\alpha_i}, \tag{8}$$

D. Calculation of Node Energy Value

We suppose the initial energy value of node M is E_0, the existing energy value is E_1, and the remaining energy value E_2 size is:

$$\frac{E_1}{E_0}, \tag{9}$$

Using the energy of node in the process of communication data transmission, receiving data and processing data β ,H= β / E_1 , so,

$$Q = \begin{cases} 1 - \beta / E_1, & \text{H>1} \\ 0, & \text{H} \le 1, \end{cases} \tag{10}$$

The node energy value is E_i=E_2+ Q, the size of the node energy can be determined by the residual energy of the node and Q.

E. Cluster Update and Maintenance

After the completion of cluster initialization, due to the influence of node movement and energy consumption, the link between nodes will change and the topology of the network will change accordingly. It requires updating and maintaining the network structure in real time, and is divided into the following two situations:

(1) Link failure. If a member of a node fails to link with its cluster head, the cluster head should update its member list and neighbor list in time. The member node that lost the connection re-searches for a new cluster head and joins. If there is no suitable cluster around to join, the node becomes Orphaned nodes, declaring themselves to be a single cluster head.

(2) Link reconstruction. If both ends of the link are member nodes or both ends of the link are cluster head nodes, one only need to update their neighbor list information, and the others states remain unchanged.

F. Determination of the Weight Value

The algorithm studied in this paper is determine the cluster head by calculating the weight value, and the node with the smallest weight can be as the cluster head. Under the same weight condition, the node with the smallest ID priority first becomes the cluster head. Therefore, it is very important to determine the weight. In this paper, the analytic hierarchy process (AHP) is used to determine the weight value. Analytic hierarchy process [11] is proposed by American Operational Researcher Pittsburgh University professor Satie, and it decomposes the elements to Goals, criteria, programs and other levels which are always related to the decision. On the basis of this, it is hierarchical weight decision-making methods that it carries through qualitative and quantitative analysis.

An important characteristic of AHP can express the corresponding degree of importance of two schemes by using the ratio of two importance, and rank them according to the degree of importance. The following table is the ratio-scale.

TABLE I. RATIO-SCALE TABLE

Scaling	meanings
1	Representing two elements is equal of importance
3	Said the two elements compared to the former slightly more important than the latter
5	Compared to the two elements, the former is obviously more important than the latter
7	Compared to the two elements, the former is more important than the latter
9	Compared to the two elements, the former is more important than the latter
2,4,6,8	Indicates the middle value of the above adjacent judgment
reciprocal	If the ratio of importance of the former to the latter is b, the ratio of the latter to the former is $1/b$

TABLE II. THE RESULTS OF RELATIVE RATIO OF TWO FACTORS

Factor	Node data value	Node energy value	The best node degree	Average distance of nodes
Node data value	1	b	b	b
Node energy value	$1/b$	1	1	1
The best node degree	$1/b$	1	1	1
Average distance of nodes	$1/b$	1	1	1

Since the algorithm studied in this paper is aimed at the problem of data transmission in Ad-hoc network that node transmission data is too large to excessive energy consumption and too small data to the unrealistic problem. Therefore, we focus on the study of the node data value, which is more important than other factors. In order to study the calculation conveniently, assuming that the node energy value, the optimal node degree and the average node distance are equally important. According to Table 1, node data value factor is more important than the other three factors. The scale is b, other factors comparing with the scale of node data value

is $1/b$. The energy scale of nodes, the optimal node degree and the average distance of nodes comparing with the latter scale is 1, and the relative ratio of two factors is shown in Table 2.

Because the node data value is obviously important compared with the node energy value, the optimal node degree and the node average distance. The value of b is 5, and the final judgment matrix is:

$$B=\begin{pmatrix} 1 & 5 & 5 & 5 \\ 1/5 & 1 & 1 & 1 \\ 1/5 & 1 & 1 & 1 \\ 1/5 & 1 & 1 & 1 \end{pmatrix} \tag{11}$$

By calculating,the judgment matrix go through the consistency test, the final weights are obtained as ω_1=0.625, ω_2=0.125, ω_3=0.125, ω_4=0.125.

III. Algorithm Simulation and Performance Analysis

A. Algorithm Simulation Environment and Performance Index

In this paper, we do a contrast experiment by using the Matlab simulation software to compare respectively with three kinds of traditional clustering algorithm containing minimum ID algorithm, the highest node degree algorithm and WCA algorithm and NDVA algorithm proposed in this paper. After a contrast experiment, we will get some conclusions of reference value and research value. The parameter of simulation experiment is shown in Table 3.

The advantages and disadvantages of Ad-hoc network clustering algorithm can be judged according to different performance indexes. In this paper, the performance of NDVA algorithm is evaluated from four aspects.

(1) The number of cluster heads. The number of cluster heads should be kept within a certain range. If the number of cluster heads is too large, the expenditure of network of communication will be increased. When the number of cluster heads are too small, the sliced time of the network will be accelerated.

(2) Re-enter the number of clusters. There is cluster structure more stable, the nodes re-entry into the cluster of number more few.

(3) Ruling set update times. Clustering algorithm in the clustering process will have a lot of communication expenditure, so it should reduce the cluster rule set update times.

(4) The total number of remaining nodes. As the number of transmitted data increases, the node's energy consumption will also increase. When the energy is exhaust, the node becomes the failed node. The failed node appears more later and the smaller the number, the network's life cycle will be longer.

TABLE III. SIMULATION EXPERIMENT PARAMETERS TABLE

Parameter	Parameter value
Simulation area/m	200*200
The total number of nodes/Number	100
Node initial energy/J	5
Node data length/bit	600
Node data control length/bit	200
Node transmission range/m	0~50
Node energy consumption value/(nJ.b-1)	0.005
Simulation time/s	300
Detection cycle/s	5

B. Simulation Results Analysis

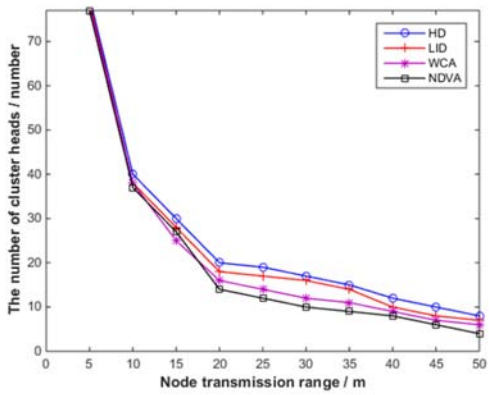

FIGURE II. THE NUMBER OF CLUSTER HEADS VARIES WITH THE TRANSMISSION RANGE OF THE NODE

The comparison of NDVA algorithm and three kinds of traditional algorithm simulation experiment of the number of cluster head are shown in Figure 2, we analyze the comparison results show that: as the node transmission range increases, the number of cluster heads gradually decreased, and when the transmission range get to10m, the trend of reduction gradually get slow down. The number of cluster heads in the NDVA algorithm is moderate and more suitable for the optimal node degree.

Figure. 3 and Figure. 4 are show that the changing of re-entrant number of clusters and ruling set update times with the moving speed of nodes, respectively. As can be seen from the comparison of Figure 3 and Figure 4, as the moving speed of nodes increase, the number of re-entry clusters and ruling set update times of the four algorithms are increase. Among them, the range of three traditional algorithms are increase more greatly. So NDVA algorithm take into account the data values of the nodes, that the number of re-entry clusters and ruling set update times are the slowest. Therefore, the clustering structure formed by this algorithm is more stable.

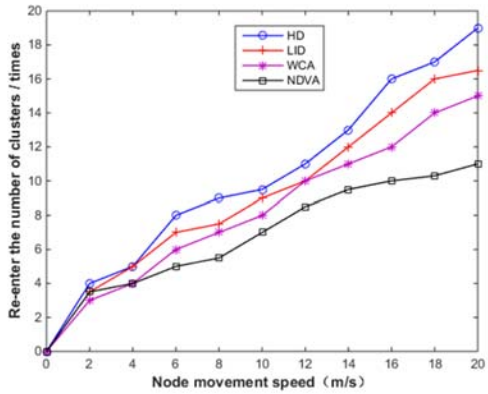

FIGURE III. CHANGE OF RE-ENTRANT NUMBER OF CLUSTERS

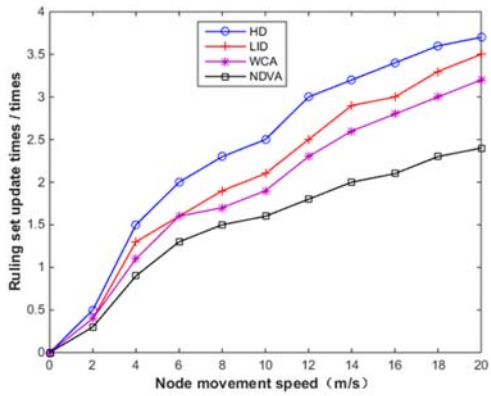

FIGURE IV. CHANGE OF RULING SET UPDATE TIMES WITH THE MOVING SPEED OF NODES WITH THE MOVING SPEED OF NODES

The comparison results that NDVA algorithm and three kinds of traditional algorithm simulation experiment remaining node are shown in Figure 5, we analyze the comparison results show that: when number of transmitted data is about 450 times, there occurs the first failure of the node by using the traditional three algorithms, when number of transmitted data is about 550 times, there occurs the first failure of the node by using the NDVA. algorithm. NDVA algorithm extends the life cycle of the network to a certain degree compared with the traditional algorithm.

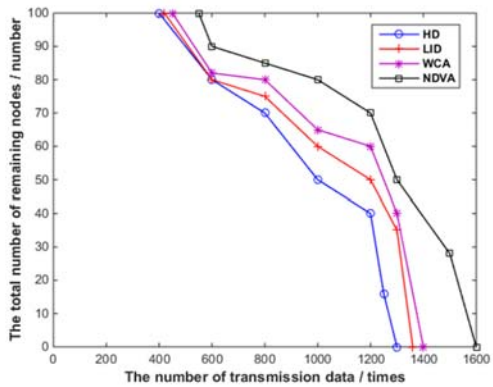

FIGURE V. THE REMAINING NODES VARY WITH THE NUMBER OF TRANSMITTED DATA

IV. CONCLUSION

In this paper, aiming at the irrational choice of cluster heads in clustered Ad-hoc networks and the existing problems in traditional clustering algorithms, a new clustering algorithm based on node data values is proposed. The algorithm nut only takes full consideration the node energy value, the optimal node degree and the average node distance, but also the node data value is added. It effectively avoids the problem of data transmission that node transmission data is too large to excessive energy consumption and too small data to the unrealistic problem. The simulation results show that the Ad-hoc network with NDVA algorithm is optimized in the number of cluster heads, and that it gets improved in reentrant clusters and ruling set update times. In a certain extent, it extends the life cycle of the network and balances the network loaded, while improving the stability of the cluster structure.

REFERENCES

[1] Young Jun Oh, Kang Whan L. A clustering algorithm based on mobility properties in mobile Adhoc network[J]. International Journal of Distributed Sensor Networks, 2015,56(4):15-29.

[2] Akbari Torkestanj J, Meybodi M R. A mobility based cluster formation algorithm for wireless mobile ad-hoc networks[J]. Cluster computing,2011, 14(4):312-313.

[3] Vodopiver S, Baster J. A survey on clustering algorithm for vehicular Ad hoc network[C]//2012 35th International Conference on Telecommunications and Signal Processing. Prague:IEEE, 2012: 52-56.

[4] Wang H T, Chen H, Song L H. Improved AOW clustering algorithm for wireless self-organized networks performance analysis[C]. 2012 cross strait quadd-regional radio science and wireless technology conference, 2012:221-228.

[5] Chatterjee M, Das S. A weighted clustering algorithm for mobile Ad hoc networks[J]. Cluster Computing, 2012,54(6):193-204.

[6] TANG C W. Stable and load balanced mobile Ad Hoc network weighted clustering algorithm[J]. Computer System Applications,2016,25(5):129-134.

[7] TANG J M, WANG C. Clustering data fusion algorithm based on relay node mechanism[J]. Journal of Yunnan University: Natural Science Edition, 2016, 38(5 703-707.

[8] MA Y Q, LI X Y. Ad Hoc network adaptive security weighted clustering algorithm[J]. Computer Engineering and Design,2014, 35(10): 3346-3350.

[9] ZHOU H. Research on dynamic entropy combination weighted clustering algorithm for Ad hoc networks[D]. Lanzhou: Lanzhou Jiaotong University, 2014.

[10] WANG C, LI C G. A relative mobility Ad-hoc network clustering algorithm[J]. Computer Applications and Software, 2016, 33(3):151-155.

[11] SUN L. The least-maximum optimization method used in AHP to determine the weight[J]. Journal of Southeast University: English version, 2012,28(2): 246-248.