Routing based on Fuzzy Rule Base System and K-Means Clustering in Ad Hoc Wireless Network

P.S.Banerjee
Kalyani Govt. Engg. College
Kalyani, Nadia, West Bengal-741235, India

J. Pal Choudhury
Kalyani Govt. Engg. College
Kalyani, Nadia, West Bengal-741235, India

S. R. Bhadrachaudhuri
Bengal Enggg. & Science University (BESU), Shibpur, Howrah, West Bengal-71103, India

ABSTRACT
Mobile Ad Hoc network (MANET) is the need of the day. It is characterized by self-organized wireless interconnection of communication devices that would either extend or operate in concert with the wired networking infrastructure or, possibly, evolve to autonomous networks. Unlike traditional wireless networks, ad hoc networks do not rely on any fixed infrastructure. Instead, hosts rely on each other to keep the network connected. Efficient routing in *MANET*, is one of the most challenging task as any node may compromise the routing protocol functionality due to lack of bandwidth or less battery life or more power consumption and/or high mobility. There are various routing problems related to bandwidth, signal strength, mobility delay, throughput and power consumption. In this paper an effort has been made to select optimum route between source and destination. Fuzzy rule base has been proposed with respect to the values of the above mentioned parameters of MANET. Optimum route has been selected by the application of K-Means clustering technique over the fuzzy rule base. Despite the infrastructural constraint of the MANET, it is expected that this optimum route would provide efficient connection between the source and destination.

General Terms
In this paper Fuzzy Logic techniques along with K-means clustering have been used to find out the optimal path for packet delivery in MANET.

Keywords
Delay, Mobility, Power Consumption, Signal Strength, Throughput, Ad Hoc Network, K-means Clustering

1. INTRODUCTION
Ad Hoc networks are made up of a number of nodes, which are capable of performing routing without using a dedicated centralized controller as a base station. Ad Hoc networks can be represented as a connected graph *G (v, e)*, with a set of vertices *v* and a set of edges *e.* Each vertex of the set v represents a network node and each edge of the set *e* represents a wireless link. These links are not static. The node connectivity changes and hence the topology is also varied. The total number of nodes is $n = |v|$ and they communicate on peer-o-peer basis. Each of the node works as routers. This key features of these networks enable them to be used in places where an infrastructure is not available, such as in desert and on battle grounds, or it can be mentioned that the dynamic nature of these networks and the scarcity of bandwidth in the wireless medium, along with the limited power in mobile devices (such as PDA's and laptops) make routing in these networks a challenging task. The topology of an ad hoc network changes due to the movement of mobile hosts, which may lead to sudden packet losses and delays. In the mobile ad hoc network due to the high mobility, low signal power and limited bandwidth the wireless links are frequently broken and new links are frequently established. Establishment of new link is called Handoff which actually is a result of variation of signal power. Such dynamic network topology presents a significant challenge for the network routing algorithms.

2. RELATED WORK
Several routing algorithms have been proposed for Ad hoc wireless networks. A. Boukerche and his team [4] have made an exhaustive survey on the routing protocols meant for ad hoc networks. They have compared a little bit among the protocols. After analyzing reactive and proactive protocols, Yang *et al.*[1] have proposed that the proactive protocols that implement the hop-by-hop routing technique, e.g Destination-Sequenced Distance Vector (DSDV) and Optimized Link State Routing (OLSR) protocols, are the best choice for mesh networks ([2],[3]). They have also inspected the design requirements for routing link metrics for the mesh networks and related them to the routing techniques and routing protocols. Acampora. A, and Naghshineh. M [5] have discussed about the quality of services of a high speed microcellular network. QoS (Quality of Service) is the basic need of any standard network. Sohan Garg, Payal ,Kansal, Viksit Kumar Sharma [6] have developed a fuzzy logic based routing algorithm where they have put an emphasis on signal power, delay and mobility in wireless ad hoc network.
All these algorithms have their own limitations. These works lay the foundation for the development of our framework in which effort has been made to find out the optimum path for data delivery. In this paper five major parameters of ad hoc network have been considered. These are: signal strength, power consumption by nodes, delay, mobility, and throughput. Fuzzy logic has been applied to get the fuzzified values of the respective parameters. These are used as input parameters of the fuzzy inference system which gives an optimal path as output. The algorithm has been simulated in MATLAB 7.9.0.529 (R2009b) in which we have uses the Fuzzy Logic Tool Box and K-Means clustering technique.

3. METHODOLOGY
3.1 Fuzzy Logic
In fuzzy logic([8] – [11]), unlike standard conditional logic, the truth of any statement is a matter of degree. The notion central to fuzzy systems is that the truth values (in fuzzy logic) or membership values (in fuzzy sets) are indicated by a value on the range [0.0, 1.0], with 0.0 representing absolute False and 1.0

representing absolute Truth. The minimum value is 0 and the maximum value is 1. The value in between the range can be decided by the membership function.

3.1 Π-shaped membership function

This spline-based curve is so named because of its Π shape [7]. The membership function is evaluated at the points determined by the vector x. The parameters *a* and *d* locate the "feet" of the curve, while *b* and *c* locate its "shoulders." The function is given by

$$
\begin{aligned}
y &= 0, && \text{when } x \le a \\
&= 2\{(x-a)/(b-a)\}^2 && \text{when } a \le x \le (a+b)/2 \\
&= 1-2\{(x-a)/(b-a)\}^2, && \text{when } (a+b)/2 \le x \le b \\
&= 1 && \text{when } b \le x \le c \\
&= 1-2\{(x-c)/(d-c)\}^2 && \text{when } c \le x \le (c+d)/2 \\
&= 2\{(x-d)/(d-c)\}^2 && \text{when } (c+d)/2 \le x \le d \\
&= 0 && \text{when } x \ge d
\end{aligned}
$$

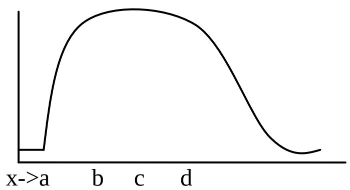

Fig 1: ∏ shaped membership function

3.2 Ad Hoc network parameters

There are several parameters which decides about the performance of the wireless ad hoc network. We have considered the following five parameters.

3.2.1 Delay

A network path with minimum delay is preferred over the others. It is worth noting that if intra-flow and inter- flow interferences, queuing delays, and link capacity are not taken into consideration, then delay minimization often ends up being equivalent to path length or hop-count minimization.

3.2.2 Energy Consumption

Energy consumption is a major issue in all types of wireless networks where the battery lifetime constrains the autonomy of network nodes. A protocol, if chooses path with an unreliable link, it would probably produce longer delay due to higher retransmission rates, that ultimately results in raise in energy consumption (along with computational processing overhead of aggressive control packets). For energy saving, most of the work focuses on the communication protocol design. For example, the routing protocol ZigBee [12] uses a modified Ad Hoc On-demand Distance Vector (AODV) routing to be used by low-power devices. By adapting transmission power to the workload, Real-time Power-Aware Routing (RPAR) protocol [13] reduces communication delays. and the link capacity possibly need to be captured by the link metric, as Yang *et al.*[1] have presented in their work.

3.2.3 Throughput

It is desirable to implement a wireless routing protocol with the maximum probability of data delivery, minimum probability of data loss. So, in wireless networks, the attempt has always been to choose an end-to-end high capacity path. A protocol can achieve maximum throughput:

(a) directly by maximizing the data flows,
(b) indirectly by minimizing interference or retransmissions,
(c) by allowing the multiple rates of data delivery to coexist in a network, where a channel with high capacity of data transmission is used over each link. End-to-end delay may also be reduced as a direct result of larger bandwidth.

3.2.4 Mobility

This is a very important feature of MANET where there is always a high probability that a moving node trying to communicate with another moving node. This mobility actually inversely effects the quality of the signal received by the nodes.

3.2.5 Signal Strength

This parameter is actually one of the most dominating parameters of Wireless network. Higher the signal power better the quality of signal received. Therefore the node has to be aligned with the higher signal power.

3.3 K-means clustering

K-means clustering is one of the simplest unsupervised learning algorithms that can solve the well known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed a priori. The main idea is to define k centroids, one for each cluster. These centroids shoud be placed in such a way that this algorithm aims at minimizing an objective function. In this case a squared error function has been used as objective function. A square error function is described as follows

$$j = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i^{(j)} - c_j \right\|^2$$

where $\left\| x_i^{(j)} - c_j \right\|^2$ is a chosen distance measure between a data point $x_i^{(j)}$ and the cluster centre c_j, is an indicator of the distance of the *n* data points from their respective cluster centres. The algorithm is composed of the following steps:

1. Place K points into the space represented by the objects that are being clustered. These points represent initial group centroids.
2. Assign each object to the group that has the closest centroid.
3. When all objects have been assigned, recalculate the positions of the K centroids.
4. Repeat Steps 2 and 3 until the centroids no longer move. This produces a separation of the objects into groups from which the metric to be minimized can be calculated.

4. IMPLEMENTATION

Step1:

Here Fuzzy Logic has been used for routing and management of an ad hoc wireless network. The proposed fuzzy logic based routing algorithm takes into account of five input variables, signal power, mobility, delay, throughput and energy consumption . The absolute value of each of these parameters can take a large range at different points on the network. In this paper the normalized values for each parameter has been considered.

Step 2:

Now, 'crisp' normalized values have been converted into fuzzy variables. For this, three fuzzy sets have been defined for each variable. The fuzzy sets, low (from 0 to 0.4), medium (from 0.2 to 0.8) and high (from 0.6 to 1.0) have been used for the input variables. The normalized value of each parameter is mapped into the fuzzy sets. Each value of the input parameters will have some grade of membership for each set. The value is either 'low', or 'medium' or 'high' that will be decided depending on the maximum grade of membership it will have in the fuzzy sets. The membership values that have been defined for each of the fuzzy set for any particular input variable are 'π' in shape. The variation of input parameters: signal strength and throughput according to the PI-membership function have been shown in Fig. 1.2 and Fig. 1.3. All the input parameters will have the same kind of characteristics.

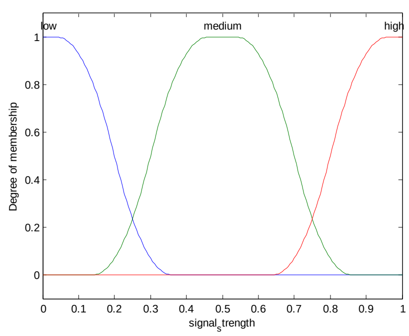

Fig. 2: input variable 'signal strength'

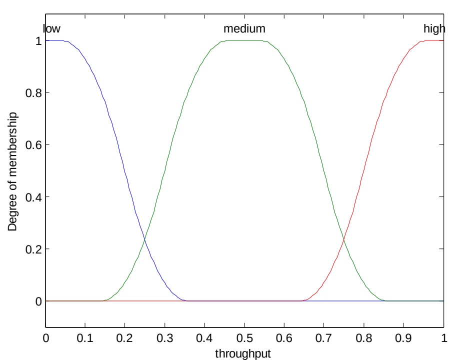

Fig. 3: input variable 'throughput'

Step 3:

The rules of inference have been formed. There are five network parameters. These parameters are used as inputs. Each parameters have been classified in three classes: low, medium and high. Therefore total 243 (3^5) number of rules have been devised. The rules have been furnished in Table 1. Due to space constraint only some of the rules are mentioned there.

Table1
Rules

Rules	Sig. strength	Mobility	Delay	Thrpt.	Energy consumption	Route
Rule 1	Low	Low	Low	Low	Low	Sub-Optimal
Rule 2	Low	Low	Low	Low	Medium	Sub-Optimal
Rule 3	Low	Low	Low	Low	High	Below-optimal
Rule 4	Low	Low	Low	Medium	Low	Sub-optimal
Rule 5	Low	Low	Low	Medium	Medium	Sub-optimal
Rule 6	Low	Low	Low	Medium	High	Below-optimal
Rule 7	Low	Low	Low	High	Low	Optimal
Rule 8	Low	Low	Low	High	Medium	Sub-optimal
Rule 9	Low	Low	Low	High	Medium	Sub-optimal
Rule 10	Low	Low	Low	High	High	Below-optimal

Step 4:

The crisp values of the network parameters have been taken. Thereafter these crisp values have been fuzzified. An output linguistic variable is used to represent a route.

Step 5:

These fuzzy input values are applied on the rule base (furnished in Table 1). Based on the fuzzy value of the network parameters, the particular rule has been selected and accordingly the output linguistic value of the rule has been obtained. This output linguistic value of the rule base decides about the route of the network.

Step 6:

The proposed optimal routes are based upon the fuzzy rules for different ranges of the metric availability. The routes are defined as below optimal (from 0 to 0.4), suboptimal (from 0.2 to 0.8) and optimal (from 0.6 to 1.0) between two mobile hosts. The below optimal indicates not optimal path, the sub optimal indicates good path and the optimal path indicates the best path.

Step 7:

The proposed routing algorithm has been applied using different routing metrics. These routes have to selected based on parameters: mobility, signal power, throughput, power consumption and delay requirements of the network.

Step 8:

The defuzzified crisp value for selected variable has been calculated from the derived algorithm.

It is necessary to select the optimal and suitable route from source to the destination based on mobility, signal Strength, delay, throughput and energy consumption. The system has been developed based on the fuzzy inference system. The major components of the system consist of the knowledge base, decision making, fuzzification and defuzzification. In this fashion 243 rules have been formed depending on the three values 'low', 'medium' and 'high' of the five input variables viz. 'signal strength', 'mobility', 'delay', 'throughput' and 'energy consumption'. The output of the system is the path or route which is either 'below optimal' or 'sub-optimal' or 'optimal' depending on the effect of the input parameters. 'optimal' means the best path for communication in present condition.

Step 9:

It is known that higher values of 'signal strength' and 'throughput' makes the route better and that of 'delay', 'mobility' and 'energy consumption' affects the route adversely. These parameters have been converted into corresponding linguistic values and based on the fuzzy rule base system the cumulative value of the 'route' has been calculated.

Step 10:

The input network parameters have been converted into certain numeric form whose values have been chosen randomly based on the nature of the effect of the parameters (adverse effect or positive effect) on the route. From these values the cumulative value of the path has been formed. These cumulative values actually represent the cumulative effect of the parameters on the quality of the route. Table 2 shows the cumulative values for the rules shown in Table 1.

Step 11:

K-Means clustering technique has been applied on the cumulative values which represent the quality of the route. 3 clusters have been formed where cluster1 represents the 'Sub-optimal' route, cluster2 represents the 'Below-optimal route' and cluster3 represents the 'Optimal' route. It is assumed that the cells in wireless ad hoc network have been clustered with respect to the input parameters and hence the route passing through the cluster3 cells will be the 'Optimal' in the present condition.

Table 2
Rule and corresponding cumulative values

Rule	Cumulative values
1	5.6
2	5
3	4.4
4	6.2
5	5.6
6	5
7	6.8
8	6.2
9	5.6
10	5

4 RESULTS

This routing framework has been tested with many test sets. Some of the test values of the input parameters and the corresponding route have been furnished in Table 3.

Table 3
Test values

Tests	Sig. strength	Mobility	Delay	Thrpt.	Energy consumption	Route
Test 1	(high)	(low)	(medium)	(high)	(low)	6.8 (optimal)
Test2	(medium)	(medium)	(low)	(medium)	(high)	3.2 (below optimal)
Test 3	(low)	(high)	(medium)	(high)	(low)	5.4 (sub-optimal)

Fig. 4, Fig. 5 and Fig. 6 show the 3D plot of the input parameters 'signal strength' and 'throughput' on the horizontal axis and output 'route' on the vertical axis The result for cluster3 (Optimal route) has been shown in Fig. 4. The result for cluster2 (Below optimal route) is shown in Fig. 5. The results for cluster1 (Sub optimal route) is shown in Fig. 6. This graph can be generated with respect to any two of the five input parameters on horizontal plane and output 'route' on the vertical plane.

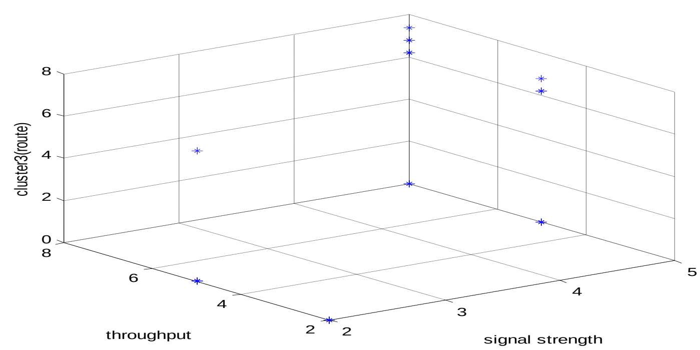

Fig. 4 Output 'Route' w.r.t 'throughput' and 'signal strength' in cluster3 for Test1

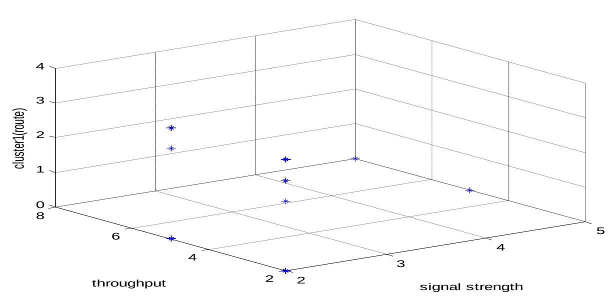

Fig 5 Output 'Route' w.r.t 'throughput' and 'signal strength' in cluster2(below optimal) for Test2

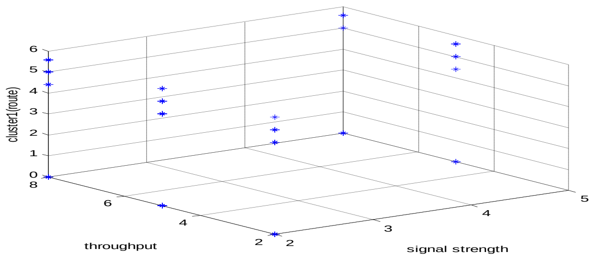

Fig. 6 Output 'Route' w.r.t 'throughput' and 'signal strength' in cluster1 (sub optimal) for Test3

5 CONCLUSION

Due to the unprecedented growth in the scale and diversity of mobile computing devices, new horizons for wireless connectivity has come into view. In this paper an effort has been made to apply the concept of K-Means clustering for selection of optimal route. Limited trials have been made to show that the protocol is functional and effective. It is obvious that further experimentation is needed to accurately assess the practical effectiveness of the protocol in a medium to large size network.

7. REFERENCES

[1] Y. Yang, J. Wang, and R. Kravets, "Designing Routing Metrics for Mesh Networks", WiMesh, 2005.

[2] C. E. Perkins and P. Bhagwat, "Highly dynamic Destination-Sequenced Distance-Vector routing (DSDV) for mobile computers," SIGCOMMComput. Commun. Rev., vol. 24, pp. 234-244, 1994.

[3] T. Clausen and P. Jacquet, "Optimized Link State Routing Protocol (OLSR)," ed, 2003.

[4] Azzedine Boukerche, Begumhan Turgut, Nevin Aydin, Mohammad Z. Ahmad, Ladislau Boloni, Damla Turgut, "Routing Protocols in Ad Hoc Networks: A survey", Computer Networks (55), ELSEVIER, 2011, pp 3032-3072

[5] Acampora. A, and Naghshineh. M: "Control and Quality-of-Service provisioning in high speed microcellular networks", IEEE Pers Communication, 1994, 1, (2), pp. 36-42

[6] Sohan Garg, Payal ,Kansal, Viksit Kumar Sharma: "A Fuzzy Logic based network Dependant Routing Algorithm for Ad Hoc Wireles Network", Int. Journal Engg. Research and Application (IJERA), vol. 1 (1), pp. 1-10, 2010

[7] Fuzzy Logic ToolBox, MATLAB

[8] Gian Nagib & Wahied G. Ali, "Network Routing Protocol using GA", International Journal of Electrical Engineering & Computer Sciences, IJECS-IJENS-Vol :10 No:2, pp40-44.

[9] Anant Oonsivilai, Wichai Srisuruk, Boonruang Marungari, Thanatchai Kulworawanichong, "Tabu Search Approach to solve Routing issues in Communication networks", Proceedings of International Conference of World Academy of Science, Engineering & Technology, 53,2009, pp1174-1177.

[10] Q. Song and B. S. Chissom, "Forecasting enrollments with fuzzy time series part I", Fuzzy Sets and Systems 54(1993) 1 - 9.

[11] J. Sullivan and William H. Woodall, "A Comparison of Fuzzy Forecasting and Markov Modelling", Fuzzy Sets and Systems 64(1994) 279 - 293.

[12] Boughanmi, N. and Song, Y.Q., "A New routing metric for satisfying both energy and delay constraints in wireless sensor networks". Journal of Signal Processing Systems for Signal, Image, and Video Technology, p137–143, 2008.

[13] O. Chipara, Z. He, G. Xing, Q. Chen, X. Wang, C. Lu, J. Stankovic, and T. Abdelzaher, "Real-time Power-Aware Routing in Sensor Networks", Forteenth IEEE International Workshop on Quality of Service, 2006.