

QOSRGA Protocol Using Non-Disjoint Multiple Routes in Mobile Ad Hoc Networks

JIWA ABDULLAH

Department of Communication Engineering
Faculty of Electrical and Electronic Engineering
Batu Pahat, Johor
MALAYSIA
jiwa@uthm.edu.my

Abstract: - The problem of QoS Routing for MANET poses several challenges that must be addressed. To select the most optimal route from source to destination, one has to choose from a set of routes with the corresponding quality of connectivity and resources. A rationale for multiple routes is presented. We argued that QoS routing protocol should exploit the rich non-disjoint node connectivity over disjoint paths to improve the source to target(S-T) connectivity lifetime in a MANET. Non-Disjoint Multiple Routes Discovery (NDMRD) protocol is proposed as a method of accumulating these routes, capturing QoS parameters and disseminates them appropriately. QOSRGA (QoS Routing Using GA) is designed to select the best QoS routes based on QoS metrics such as bandwidth, delay and node connectivity index (*nci*). We outlined the GA process and how its related parameters are chosen. The performances of NDMRD and QOSRGA protocol are presented.

Key-Words: - MANET, QoS Routing, non-disjoint routes, disjoint routes, multiple routes, Genetic Algorithm, Route Accumulation Latency, RREP duplicates, protocol performance, performance metrics

1 Introduction

The goal of QoS routing is to provide an application with a wireless connectivity which could sustain bandwidth, delay, jitter and loss rate requirements. The main objective of QOSRGA is to utilize several feasible multiple paths discovered in a mobile ad hoc networks, then search for the best path using GA. The search is done with four QoS metrics without imposing excessive overheads. Multiple paths have been recognized as an important network feature to improve network reliability[1]. Many papers on mobile ad hoc network routing protocol suggested that their proposed protocols worked correctly, although little were mentioned about the performances[2][3]. Others measured route coupling [4][5][6], the mutual interference of routes in a common-channel multihop ad hoc network. Most of the works on mobile ad hoc multipath restricted their works to small route count, usually two. AOMDV [7] allows up to k link-disjoint RREPs, where one is the quickest path and the others are chosen from the next link-disjoint RREQs. SMR [8] builds two paths from the quickest RREQ and then collects RREQs for a period and chooses a second maximally disjoint path from the first. In a zone-disjoint scheme [9], only two paths are built, but they are not necessarily the minimum. In this paper, it is argued that a QoS routing protocol for MANET

should fully exploit the rich connectivity of the network to improve the overall reliability of packet delivery. Paths with poor quality, shorter node pair connectivity should not be used. In this paper we introduced Non-Disjoint Multiple Routes Discovery (NDMRD) protocol to discover multiple routes and disseminate QoS metrics. Interestingly, a number of routing protocols for ad hoc networks that attempt to take advantage of multiple paths to destinations advocate the use of node-disjoint paths. Section 2 makes the case that disjoint paths are not necessary to improve the reliability of wireless ad hoc networks. Furthermore, Section 3 shows that multiple well-connected loop-free paths offer substantially longer path lifetimes than sets of disjoint paths. Section 4 elaborate on the NDMRD protocol while Section 5 describe the QOSRGA protocol. The performances evaluation are presented in Section 6. Section 7 summarizes the paper.

2 The Rationale of Multiple Routes

2.1 A Collection of Multiple Routes Schemes

Multiple routes between a source and destination are of two types, namely *node-disjoint* and *link-disjoint* multiple routes. Node-disjoint routes do not have any nodes in common, except the source and destination. Fig.1(a) shows a typical non-disjoint network in which it can generate into

Non-Disjoint Multiple Routes Discovery protocol. Fig. 1(b) shows a node disjoint multiple routes with initial two branch. Node-disjoint routes do not have any nodes in common except source and destination as shown in Fig.1(c). Nodes label S and T are source and destination respectively. The advantages of node-disjoint multiple routes are that they may fail independently of each other. Breakage on one route can be corrected by resuming data sessions through the other routes. Fig. 1(d) shows *link-disjoint* multiple routes between S and T, formed with two segments. In Fig.1(e), if node-disjoint multiple routes to destination are available on every node at every routes on the primary route. In Fig.1(f) shows a fail-saved multiple routes.

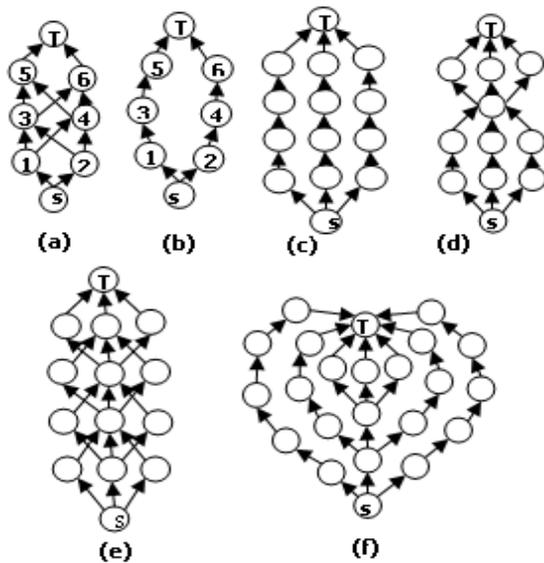


Fig. 1 Different Types of Multiple Routes

2.2 Non-Disjoint Routes vs Disjoint Routes

Disjoint routes are those routes where there are no common nodes in their entire routes except the source and destination nodes. Most of the previous works on MANET multiple paths has restricted the number of potential routes to a small number, typically two. AOMDV [7] allows up to k -link-disjoint RREPs, where one is the quickest path is chosen from the next link-disjoint RREQs. SMR[8] builds two paths from the quickest RREQ and then collects the RREQs for a period and chooses a second maximally disjoint path from the first. We argued that a QoS routing protocols for MANET should fully exploit the rich connectivity of the multiple paths network by considering non-disjoint routes hence improving the reliability of packet delivery.

Definition 1: Consider a set of m valid routes from source to target as $P_{ST} = \{P_0, P_1, P_2, \dots, P_{m-1}\}$. Each route consists of a collection of nodes, $P_i = [n_0, n_1, \dots, n_i, \dots, n_{k-1}]$. A node non-disjoint set of routes is said to exist if n_i which is not the source or target, is a member of at least two different routes simultaneously. Each route then must have at least one node that is common to any other routes in P_{ST} . One of the characteristics of a non-disjoint network is that, there exist a number of common nodes excluding the S and T. Here, we specify that there must be at least one common node excluding source and target. The non-disjoint multiple paths are still valid if common links exist. If each node has different node qualities in terms of node state, then the combination of all nodes in the route could produce a measure of route quality. In the non-disjoint network the probability of selecting the most reliable routes, is increased. Most of the routing protocols dealing with multiple paths [18] only utilized the disjoint networks. By combining the node state[22] information for each node in the routes and the routes discovered within the non-disjoint network we have a rich choice of possible S-T routes. It is in this respect that the Genetic Algorithm would be used to find the most optimum route with several simultaneous constraints as opposed to other heuristics such as admission control, sequential filtering, metrics ordering or rescheduling principles. This would be consistent with our work on the design of QoS routing. Hence we proposed a Non-Disjoint Multiple Routes Discovery (NDMRD) protocol as part of our QoS routing (QOSRGA) protocol.

2.3 Multiple Routes Reliability Analysis

Consider the networks in Fig. 1(a) and (b) only. The network (b) shows disjoint S-T connectivity and the network (a) shows rich non-disjoint connectivity. If we consider each node to have mobile probability p , then using reliability calculation we can determine the S-T reliability of the two networks. Using the method of inclusion/exclusion on minimum paths shown by [2], the reliability polynomials are:

$$R(\text{disjoint}) = 2p^4 - p^8 \tag{1}$$

$$R(\text{non-disjoint}) = 2p^4 - p^8 + (6p^4 - 12p^6 - 8p^7 + 15p^8 + 12p^9 - 20p^{10} + 8p^{11} - p^{12}) \tag{2}$$

The disjoint routes in Fig. 1(b) have two minimum paths $\{S-1-3-5-T\}$, $\{S-2-4-6-T\}$. The non-disjoint routes in Fig. 1(a) has eight minimum paths $\{S-1-3-5-T\}$, $\{S-1-4-6-T\}$, $\{S-1-3-6-T\}$, $\{S-1-4-5-T\}$, $\{S-2-4-6-T\}$, $\{S-2-3-5-T\}$, $\{S-2-3-6-T\}$, $\{S-2-4-5-T\}$.

Fig. 2 plots the network reliability for the disjoint and non-disjoint routes configurations. As expected, the non-disjoint configuration has a significantly higher reliability.

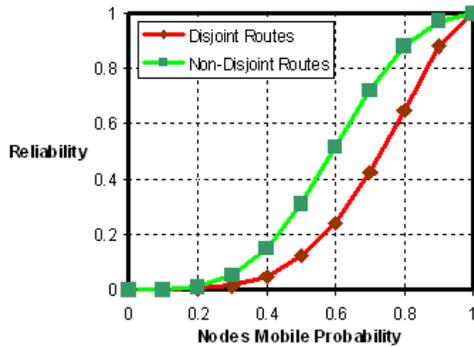


Fig. 2 Route Reliability vs Mobile Probability

Nasipuri *et al.* [18] uses a non-disjoint multipath approach in their routing protocol. In the proposed protocol, it is assumed that the primary path be k hops. Each node along the primary path has an alternate disjoint route to T, so there are $k+1$ minimum path. In their first protocol, it has only two minimum paths. This explains the phenomenon they observe that the rate of path discovery decreases as the path length increases. It is because with each extra hop along the primary path, they add another minimum path. The method of by [2] in the computation of time for node connectivity is being adapted. Following [18], let ensure that the cumulative distribution function(CDF) for link operation is $F(t) = 1 - e^{-\lambda t}$. For a series of k links, the CDF is $F_s(t) = 1 - e^{-k\lambda t}$. For a set of m parallel paths, each with a CDF of $F_s(t)$, $F_p(t) = (F_s(t))^m$. Using these results, the CDF for the disjoint network in Fig. 1(b) is $F_{disjoint}(t) = (1 - e^{-4\lambda t})^2$. Using the relation that the expected value is

$E[X] = \int_0^{\infty} 1 - F(x) dx$ then the mean lifetime of the disjoint routes is,

$$E_{disjoint}[X] = \int_0^{\infty} 2e^{-4\lambda t} - e^{-8\lambda t} dt = 3/(8\lambda). \quad (3)$$

To analyze the non-disjoint routes, we use the equation from [2], reliability is,

$$R = \sum_{j=1}^h (-1)^{j+1} \cdot \sum_{I \subseteq \{1, \dots, h\} | |I|=j} \text{Prob}[E_I] \quad (4)$$

where E_I is the event that all paths P_i with $i \in I$ operate no longer than time t . Let n be the number of distinct links in E_I , then $\text{Prob}[E_I] = 1 - e^{-(n\lambda t)}$. It is required that all paths with n distinct links operate no longer than time t is exactly the same as a series path of n links. This will yield an equation almost

identical to equation (3), except that each term ap^b will be replaced by $(-ae^{-(b\lambda t)})$. Then,

$$F_{Non-Disjoint}(t) = 1 - 8e^{-4\lambda t} + 12e^{-6\lambda t} + 8e^{-7\lambda t} + 14e^{-8\lambda t} - 12e^{-9\lambda t} + 20e^{-10\lambda t} - 8e^{-11\lambda t} + e^{-12\lambda t} \quad (5)$$

$$E_{Non-Disjoint}[X] = \int_0^{\infty} 1 - F_{Non-Disjoint}(t) dt = 44/(77\lambda) \quad (6)$$

Comparing the two equations, we find that the non-disjoint routes last, on average, 1.52 times longer than the disjoint routes. Repeating the same calculations for 3-hop, 5-hop and 6-hop routes we get the comparisons of ratios. The ratios of node connectivity time of non-disjoint to disjoint networks for 3, 4, 5 and 6-hops are given as 1.28, 1.52, 1.79 and 2.01 respectively. While it is difficult to generalize the mean S-T connectivity lifetime to an arbitrary network, the trend is obviously favoring the non-disjoint construction.

3 The Rationale of Multiple Routes

3.1 Multiple Routes Discovery and Routes Accumulation

When a source node wants to communicate with a destination node, it checks its route table to confirm whether it has a valid route to the destination. If so, it sends the queued packets to the appropriate next hop towards the destination. However, if the node does not have a valid route to the destination, it must initiate a route discovery process. To begin the process, the source node creates RREQ packet. This packet contains message type, source address, current sequence number of source, destination address, the identification, flow id and route list. The identification is incremented every time the source node initiates a RREQ. In this way, the broadcast identification and the address of the source node form a unique identifier for the RREQ. The first aim of NDMRD protocol is to generate a set of ST routes which are node non-disjoint. This set of routes is to be accumulated at the source node. To achieve this, each node must receive duplicates of RREQ packet. If we set each node to only allow one duplicate of RREP to be forwarded, then we would have a set of routes that are disjoint. To achieve non-disjoint routes, we could allow as many duplicates as permitted by the memory. For example, consider a scenario in Fig. 3 consisting of five nodes. S is the source and T is the target node where {1, 2, 3} are neighbors to S, {S, 2, T} are neighbors to node 1 and also to node 3. Node S is out of range to T and 3 is out of range to 1.

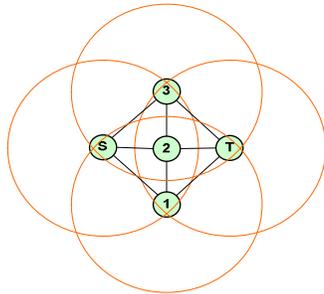


Fig. 3 A Five Nodes Scenario

Definition 2: The **RREQ duplicate** is defined as the number of times RREQ packet from a particular source with the same identification arrived at intermediate nodes and was forwarded.

To initiate the route discovery, node S starts transmission of packet RREQ. The maximum duplicates for RREQ is set to 10. When RREQ is generated at S, the packet is then broadcasted. Node 1, node 2 and node 3 will receive the packet and forward it to node T. When the RREQ packet has arrived at node T, RREP packet is then generated and sends in reversed direction to the route list of the RREQ packet. After a time, t the source node, S produces a set of non-disjoint routes as { [S-3-T], [S-2-T], [S-1-T], [S-3-2-T], [S-3-2-1-T], [S-2-3-T], [S-2-1-T], [S-1-2-T], [S-1-2-3-T] }. It is observed that the RREQ packet goes through node 2 seven times. If we limit the number of duplicates to 5 for example, the number of non-disjoint routes returned will be 6. Figure 4(a) illustrates the process of route accumulation and Figure 4(b) shows the outcome as a set of routes in the node's routing table. The number of routes will also be reduced if the time t is set less than before.

Definition 3: The **route accumulation latency** is defined as the length of time allowed for a source node to accept a number of RREP packets destined to this node with the same *id* as that RREQ which originated from it.

The symbol, t represents the route accumulation latency. The value of t chosen must ensure a good number of routes are accumulated. The number of RREP packets received represents the number of node non-disjoint routes discovered. Hence there are two parameters that governed the number of non-disjoint routes to be accumulated: (1) the maximum number of RREQ duplicates and (2) the route accumulation latency. The question is how do we determine the value of the number of RREQ packet duplicates and the route accumulation latency? The second aim of the NDMRD protocol is to facilitate the functions of the node state monitoring protocol in the updating, disseminating and accumulating the

QoS route parameters. The node state monitoring protocol is described in [22]. In this route discovery procedure, the source node transmits a Route Request(RREQ) packet, identifying the target for which the route is needed. An intermediate node receiving the RREQ packet retransmits the packet if it has not yet forwarded a copy of it.

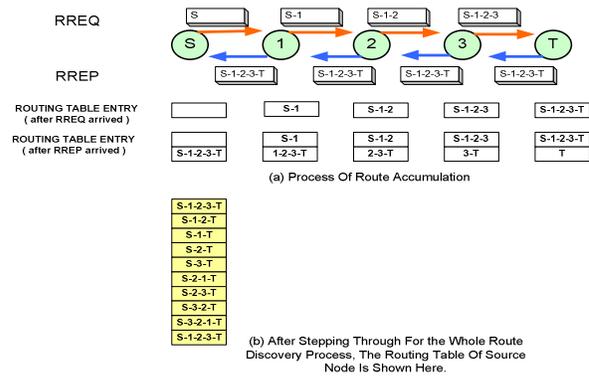


Fig. 4 Route Accumulation at Source Node

When the target node has received the RREQ packet, it returns a Route Reply (RREP) to the source. The RREP packet then traverses the route taken by the RREQ, in the opposite directions of the RREQ and is propagated towards the source. To reduce the frequency of performing route discovery and to limit the flooding of the network by forwarding the RREQ mechanism the protocol sends the periodic connectivity (CONN) packet at the interval rate of one per second. CONN packet is transmitted with Time-to-Live (TTL) set to 1, to avoid flooding. This will ensure nodes within the transmission range are connected as neighbor nodes and to activate the node state cache, while nodes that have moved out of range, will have their node state cache deactivated, avoiding stale information. When the source node starts transmitting RREQ packets, the sending of periodic CONN packets is stopped until the transmission session is completed or the route is broken. For each individual route discovery attempt, each node forwards the RREQ base on the following conditions, (1) should this be the first occasion, that RREQ has arrived at the node, then it forwards the packet to the next one-hop nodes accordingly; (2) if it's not the first time, then this is the RREQ duplicate. The RREQ duplicate counter is then incremented and if the counter is more than the maximum number of RREQ duplicates allowed, the RREQ packet is destroyed; (3) if no more hop, the RREQ packet is destroyed. The route lifetime is indicated by the node connectivity index(nci) in the form suitable for QoS routes determination [22].

4 NDMRD Protocol

The Opnet[17] implementation of NDMRD protocol consists of Send Route Request module, Send Route Reply module, Received Route Request module and Received Route Reply module.

4.1 Send Route Request Module

The RREQ packet is activated when event packet arrival has occurred. When a packet arrived from upper layer, it is an application packet. The protocol then checks the Routing Table and determines whether any route exists to the destination. If no route exists, then the RREQ packet is initiated. This creates the QOSRGA packet, issues an id, encapsulates the packet into IP, sets TTL and lastly sets the Originate Request Table. The RREQ packet is then broadcasted.

4.2 Received Route Request Module

The Received Route Request module is activated when type RREQ packet arrived at the node. If the node is an intermediate node, the packet is checked for RREQ duplicates, using the list in the Forward Request Table. Then, the packet is registered and allowed for up to a maximum number of the RREQ duplicates. This will result in the creation of non-disjoint routes. If TTL is 1, destroy the packet, otherwise set the entry to Forward Request Table. Then rebroadcast the RREQ, with this node address add to the address list of RREQ packet. If the node is a source node, then destroy the packet, the node is receiving its own RREQ packet. If the node is the destination of the packet, then initiate a Send Route Reply operation.

4.3 Send Route Reply Module

Send Route Reply module will be activated on the arrival of RREQ packet if the node is the destination node. RREP packet is then created with the source address from the IP datagram of the RREQ packet. The protocol then extracts the Node State information from the Node State Table of this node and insert into the RREP packet. The value of node bandwidth, medium access delay, nci and end-to-end delay were piggybacked onto the RREP packet. The IP datagram is then created and subsequently encapsulates the RREP packet. Following this, RREQ is copied into the Forward Route Request Table. The protocol allowed the

destination node to permit the RREQ duplicates. This allows the non-disjoint routes to be returned. Since the node is the destination node, RREP is then unicasts to the next-hop neighbour following the reversed routes obtained from the RREQ packets.

4.4 Received Route Reply Module

If the node is the intermediate node, then the QoS parameters are obtained from the Node State Table and insert into the RREP packet. The QoS parameters from the previous hop are extracted from the RREP packet into this node's QoS parameter's list. The route traversed by RREP packet is then copied into the node's Routing Table. After setting the RREP packet into the IP datagram, it is then forwarded to the next hop. If the node is the source node, the QoS parameters list are extracted from the RREP packet and inserted into the node's QoS parameter's list. The QoS parameters matrix are generated, given as Eqn. 1, Eqn. 2, Eqn. 3 and Eqn.4, where

$$D = \begin{bmatrix} D_{0,0} & \cdots & D_{o,k-1} \\ \vdots & \ddots & \vdots \\ D_{k-1,0} & \cdots & D_{k-1,k-1} \end{bmatrix} \quad (7)$$

$$C = \begin{bmatrix} nci_{0,0} & \cdots & nci_{o,k-1} \\ \vdots & \ddots & \vdots \\ nci_{k-1,0} & \cdots & nci_{k-1,k-1} \end{bmatrix} \quad (8)$$

$$BW = [B_0 \ B_1 \ B_2 \ \cdots \ B_{k-1}] \quad (9)$$

$$d = [d_0 \ d_1 \ d_2 \ \cdots \ d_{k-1}] \quad (10)$$

where $D_{i,j}$ is the end-to-end delay, nci_{ij} is the node connectivity index, B_i is the node bandwidth and d_i is the medium access delay. The matrices representing the end-to-end delay, node connectivity index, node bandwidth and medium access delay respectively. Finally the RREP packet is destroyed.

4.5 Received Data Module

If there are still more hops in the data packet, the current node is not yet the destination. After running some data processing maintenances, unicast the data packet to the next hop according to the data packet route list. On the other hand if it is the destination node, then the packet will be sent to the upper layer.

4.6 Received Error Module

If the packet is of type RERR, then all routes in the node Routing Table that have link from error source address to the unreachable node address were removed. This might prevented from using stale information regarding the current Node State.

4.7 Received Connectivity Packet Module

The partially periodic transmission of CONN packet is explained in Section 3.1. The format of the packet is of the form similar to RREQ packet but with the identification set to zero. On receiving the packet type RREQ, the protocol checks the identification field. If its '0', then the packet is CONN packet otherwise it's a RREQ packet. If it's a CONN packet, it is then destroyed. The purpose of CONN packet is to ensure the monitoring protocol can always measure the topology changes during the non-transmitting phases and maintain the Node State.

5 Implementation of QOSRGA

5.1 The Design of QOSRGA

The proposed QOSRGA is based on source routing, effectively select the most viable routes in terms of bandwidth availability, end-to-end delay, media access delay and the sum of *nci*. The NDMRD protocol initially determined a number of potential routes. The returning RREP packets extract the QoS parameters from each node along the routes. GA then operates on this set of routes and the corresponding set of QoS parameters. The operation of GA is shown in Figure 5.

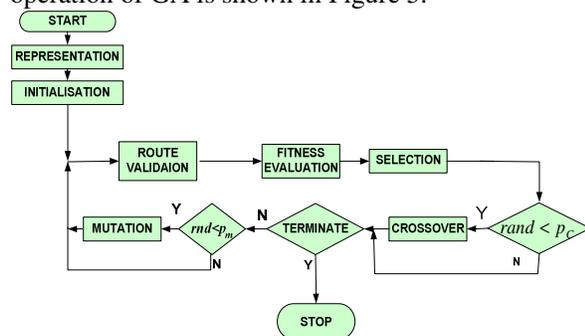


Fig. 5 Flowchart Showing the Operation of GA

5.1.1 Chromosome Representations

The chromosome consists of sequences of positive integers, which represent the identity of nodes through which a route passes. Each locus of the chromosome represents an order or position of a node in a route. The gene of the first and the last locus is always reserved for the source node, *S* and

destination node, *T* respectively. The length of the chromosome is variable, but it should not exceed the maximum length $|V|$, where $|V|$ is the total number of nodes in the network [14]. It is unlikely that more genes are needed than the total number of nodes to form a route.

5.1.2 Limited Population Initialization

GA process typically starts with a large number of initial populations which has better chances of getting good solutions. In QOSRGA, the initial populations are accumulated by NDMRD protocol. In a MANET system, with 5 nodes, the possible number of solution are calculated as 10 according to the formula $n(n-1)/2$ [13]. One approach is to generate the initial solutions randomly and then remove the invalid solutions before being fed to the GA module. Furthermore the infeasible solutions can only be eliminated after the connectivity matrix is obtained by the multiple route discovery algorithms. Clearly a set of useful solutions are extracted, before being processed by the GA module. This set of solutions has the characteristics of non-disjoint multiple routes

5.1.3 Fitness Calculation

Fitness calculation is most crucial in the GA operation, where best route can be identified. In our case the least value of fitness constitute the lowest cost and the one that is to be chosen. The fitness value of routes is based on various QoS parameters; bandwidth, node delay, end to end delay and the *nci*. Clearly it can be classified as multiple-objectives optimisation problem. According to M. Gen [13], each objective function can be assigned a weight and then the weighted objectives combined into a single objective function. For our QOSRGA protocol, the weighted-sum approach can be represented as follows. The fitness function operates to minimise the weighted-sum *F*, which is given as,

$$F = \alpha.F_1 + \beta.F_2 + \gamma.F_3, \tag{11}$$

where,

$$(a) \quad F_1 = \sum_{i=1}^m \sum_{\substack{j=1 \\ j \neq i \\ (i,j) \neq (s,t)}}^{|s \rightarrow t|} C_{ij} \cdot Q_{ij} \tag{12}$$

$$(b) \quad F_2 = \left(\sum_{i=1}^m \sum_{\substack{j=1 \\ j \neq i}}^{|s \rightarrow t|} D_{ij} \cdot Q_{ij} + \sum_{j=1}^{|s \rightarrow t|} d_j \cdot N_j \right) - D_{Qos} \tag{13}$$

$$(c) \quad F_3 = \begin{cases} 1/B_i & \text{if } B_i - B_{Qos} > 0 \\ 1000 & \text{if } B_i - B_{Qos} \leq 0 \end{cases} \tag{14}$$

The fitness function is a weighting function [13] that measures the quality and the performance of a specific node state. A fitness function must include and correctly represent all or at least the most important parameters that affect QoS routing. The next issue is the decision on the importance of each parameter on the QoS routing protocol as a whole. The significance of each parameter is defined by setting appropriate weighting coefficients to a , b and g in the fitness function that will be minimized by the GA operations. The values of these coefficients are determined based on their equal importance towards the overall QoS Routing as follows: $a=10e^{-3}$, $b=10e^{-4}$ and $g=10e^{-3}$. With regard to the function which involved bandwidth, it is required to find the minimum bandwidth among the nodes and compare this with the demand bandwidth, B_{QoS} . If the minimum bandwidth is less than the B_{QoS} , we set the fitness to a high value so that in the selection process it will be eliminated. By doing so we have simultaneously eliminated all the nodes where the bandwidth is limited, the total delay is more than the typical delay and when the nci is high.

5.1.4 Mobile Nodes Crossover

Crossover examines the current solutions in order to find better ones. Physically, the crossover operation in the QoS routing problem plays the role of exchanging each partial route of two chosen chromosomes in such a manner that the offsprings produced by the crossover represent only one route. This dictates selection of one-point crossover as a good candidate scheme for the proposed GA. One partial route connects the source node to an intermediate node, and the other partial route connects the intermediate node to the destination node. The crossover between two dominant parents chosen by the selection gives a higher probability of producing offsprings having dominant traits. But the mechanism of the crossover is not the same as that of the conventional one-point crossover. In the proposed scheme, the two chromosomes chosen for crossover should have at least one common gene except for source and destination nodes. It is not a requirement that they be located at the same locus. That is, the crossover is independent of the node position in routing paths.

5.1.5 Restoration Function

The crossover operation may generate infeasible chromosomes that violate the constraints, causing loops to be generated in the routing paths. The restoration method is employed in the proposed GA

which eliminates the lethal genes. It thus can cure all the infeasible chromosomes.

5.1.6 Route Mutation

Mutation is used to randomly change the value of a number of the genes within the candidate chromosomes. It generates an alternative chromosome from a selected chromosome. It can be seen as an operator charged with maintaining the genetic diversity of the population, thereby keeping away from local optima. Mutation may also induce a subtle bias in which it generates an alternative partial route from the mutation node to the destination node. Indeed by the process of mutation, harmful effects may vanish altogether.

5.1.7 GA Parametric Evaluations

Selecting genetic algorithm parameters such as population size, mutation rate and crossover rate is very difficult task. Besides that the selection algorithm must be rightly chosen. Each combination of parameters may produce a variety of outcomes. Haupt *et al* [16], outlined a general procedure for evaluating these parameters, after which reasonably suitable parameters are adopted for the specific application. In our case, we considered four selection methods namely the roulette wheel selection (RWS); tournament selection (TS); stochastic universal selection (SUS) and elitism technique [ET]. Next, the parameters P_c , P_m and population size are considered. We need to examine the performances of each and select our preferences. We use Matlab to initially designed GA-based routing algorithm without the QoS function. The route selection is based on the shortest path without considering the bandwidth, delay and node connectivity index. The cost for each path is randomly generated. The main objective is to examine all the GA parameters that are useful for our protocol designed and would use them in the designed of our QoS route algorithm. Hence, in this section we consider a mobile network consisting of 20 nodes, randomly distributed within a perimeter of 1000m by 1000m. Each node has a transmission range of 250m.

a) Population Size

The effect of population is investigated by fixing the mutation rate ($P_m=0.01$) and changing the population size. The simulation is run for 2000 generations. The minimum cost in each generation is recorded and the average minimum cost C_{AMC} is evaluated over the range from 0 until the 2000th generation. Fig. 6 shows the plot C_{AMC} for four different selection methods (with $\mu = 0.05$ for

Elitism). It shows that in RWS, a population size in excess of 700 produces a significantly low cost. This is because with a large population, the RWS method finds it easier to choose the low cost individuals. Consequently, the probability of a low cost individual being selected becomes low. Apart from this, with a large population size there are too many sectors within the wheel making the probability of selecting each sector smaller. The most significant result is that of the tournament selection and elitism. With a population size of approximately 10, it produces very low C_{AMC} . Hence the best choice of selection method would be the tournament selection and elitism. In fact we could use a population as low as 20 and still produce good fitness. Finally, the Tournament selection is opted.

b) Crossover Probability and Mutation Probability

Another set of very important parameters for the GA implementation are the crossover probability P_c and the mutation probability P_m . The parameters determine how many times crossovers occurred and how many times mutations occurred within a transmission period. The occurrence of crossover and mutation increases the convergence rate. De Jong[16] tested various combinations of GA parameters and concluded that mutation is necessary to restore lost genes but this should be kept at a low rate for otherwise the GA degenerates into a random search. Further study by Schaffer *et al* [20], suggest that the parameters should have these recommended ranges; population size of 20 ~ 30; mutation rate of 0.005 ~ 0.1 and crossover rate of 0.75 ~ 0.95. Another study by Haupt [21] concluded that the best mutation rate for GA's lies between 5% and 20% while the population size should be less than 16. For our case, where GA operation is done online, the value of P_c and P_m is taken to be between 0.4 and 0.9 and between 0.05 and 0.2 respectively. The choice of these parameters should produce a reasonably high efficiency packet transmission. We limit the population size up to the number of routes discovered. The limit is also imposed on the number of generations. Haupt [16] provide useful guidelines on when to stop the algorithm. We designed the algorithm so that it stops when the value of the route does not change and also we restricted the maximum number of generations up to 20.

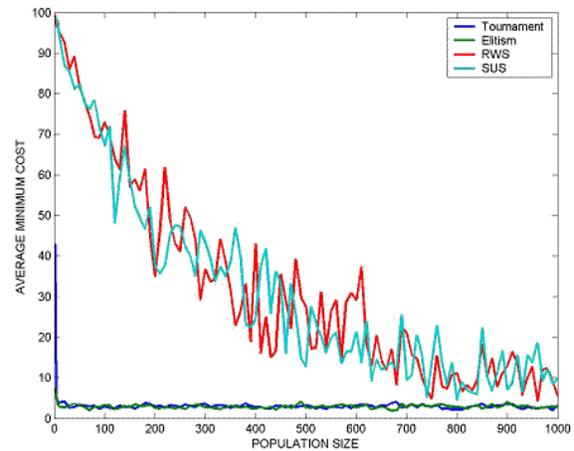


Fig. 6 Average Min Cost Against Population Size

We run simulation experiments online by considering a MANET scenario running QOSRGA protocol, with 20 nodes placed within an area of 1000 meter x 1000 meter. Each node has a radio propagation range of 250 meters and channel capacity of 2 Mbps. We initiated 10 sources transmitting CBR with data payload of 512 bytes. The nodes move randomly with random waypoint mobility model. For each point, 10 simulations run were done and for each run it executed for 200 seconds. We conduct two set of simulation experiments, one for calculation of crossover probability and the other for mutation probability. We conduct the experiment using the source traffic rate of 40 kbps, 100 kbps and 900 kbps. The aim of the first set is to identify exactly the possible values of P_c which would give the best results. Our metric is the transmission efficiency. It is defined as the ratio of average throughput of all nodes to the average load of all the nodes in the network. We varied P_c but set P_m constant as 0.1. The results are shown in Figure 7. For P_c with values from 0.4 to 0.8 the transmission efficiency is more than 80%.

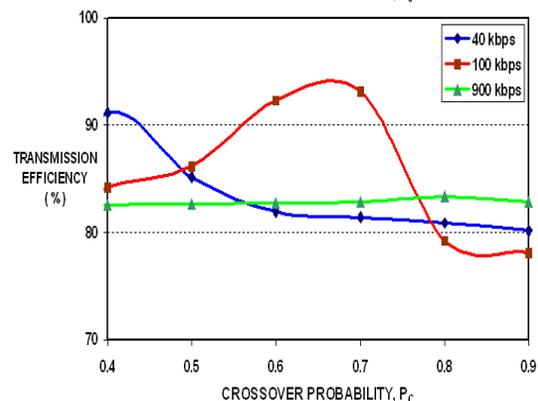


Fig. 7 Effect of P_c on Transmission Efficiency

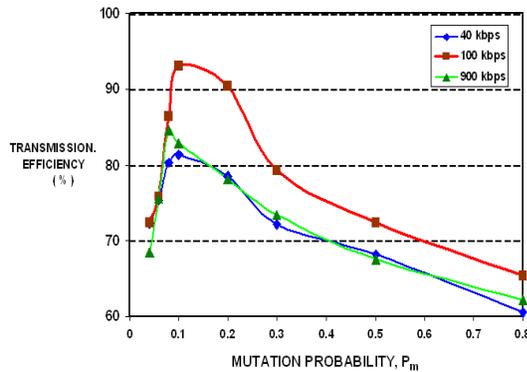


Fig. 8 Effect of P_m on Transmission Efficiency

For 100 kbps CBR source, the maximum efficiency occurred when P_c is approximately 0.65 and for 40 kbps CBR source it is 0.4. The 900 kbps CBR source, the efficiency does not deviate very much. Hence as a general guideline we choose the value of P_c as 0.7 for all our future simulation experiments. For mutation probability, we run a similar simulation, with crossover probability fixed at 0.7 and varies the mutation probability from 0.04 to 0.8. Figure 8 shows the result of mutation probability. Mutation probability produces highest transmission efficiency when it is 0.1, which is more than 80 % for all three different traffic rates. Hence it is concluded that the crossover probability and mutation probability can be taken as 0.7 and 0.1 respectively.

6 Performance Evaluations

6.1 Performance Metrics

The following metrics [19] are used in varying scenarios to evaluate the three different protocols.

6.1.1 Average Number of Routes

This metric is useful since the protocol undertake to generate as many routes as possible. We varies the number of RREQ duplicates and RAL to understand its effect on the number of routes. In the implementation of the protocol, we could tweak the value of RREQ duplicates and RAL.

6.1.2 Average packet delivery ratio

Since our study is essentially based on bandwidth measurement, we propose a metric which expresses the efficiency of bandwidth, as an average packet delivery ratio. We defined the average packet delivery ratio (APDR) as the ratio between the total packets generated by every node to the total received packets at the upper layer within the nodes in the system. We expressed it in terms of a percentage.

6.1.3 Average total end to end delay of packets

This includes all possible delays from the moment the packet is generated to the moment it is received by the destination node. The statistic of average delay of all the packets received during the simulation time is taken and then divided by the average total number of packets arrived at every receiving node. This gives the average delay of a packet.

6.1.4 Total Average Throughput

In this context the throughput is defined as the total number of bits (in bits/sec) forwarded from the WLAN layers to higher layers in all WLAN nodes of the network. To find the average throughput of a single node one has to divide by the number of nodes in the system.

6.2 Performance of NDMRD protocol

6.2.1 Number of Routes Discovered

The strength of the NDMRD protocol is the number of routes that it discovered as a result of RREQ packets send and RREP packets received. The total number of routes that are discovered depends largely on two related variables, the Route Accumulation Latency and the number of RREQ duplicates. The number of routes available and the Node State information that are piggy-backed to RREP packet will enhance the ability to discover the routes based on the given metrics. The multiple routes discovery is done within the time specified as Route Accumulation Latency. The total number of routes is obtained by counting the total number of RREPs from the destination. Simulation experiment is done to ascertain the value of RREQ duplicates and Routes Accumulation Latency that need to be specified for our QoS routing protocol. The setup consists of 20 nodes with Random Waypoint Mobility model, representing the nodes that move randomly inside a field configuration of 1000m x 1000m. The objectives of the experiment are three folds: (1) to study the effect of Route Accumulation Latency on the number of routes discovered; (2) to study the effect of RREQ duplicates on the number of routes discovered; (3) to study the relationship between the numbers of routes discovered and the overall performance of the NDMRD protocol for one and five CBR sources when considering for various number of routes. The overall performance of the NDMRD protocol will indicate to us the optimum value of Route Accumulation Latency and the number of RREQ duplicates that need to be set.

6.2.2 The Effect of RREQ Duplicates

Figure 9 shows the number of routes accumulated as a function of the duplicates. When number of duplicates is set to zero, an average of 2 routes are obtained. Maximum number of routes is accumulated when duplicates is set to 10 until 15. After 15 duplicates the number of routes starts to drop to 15 routes and when 20 duplicates are used, the routes discovered reduced to 16. Initially, no congestion occurred, so many routes are accumulated. As the RREQ duplicates are increased, the broadcast and the unicast packet that are generated greatly increased the traffic congestion. Hence the number of initial routes discovered is less.

6.2.3 Route Accumulation Latency

Figure 10 shows the effect of changing the Route Accumulation Latency on the number of initial routes. The rate of change of the number of routes from RAL=0.05 until RAL=0.1 is very steep. The number of initial routes accumulated is increasing proportionately. Beyond that, the rate of change of the number of routes decreased tremendously. When RAL is set more than 0.1 seconds, the duplication of RREQ packet becomes more nuisance than necessary. It generated more congestion and caused RREP packet to drop more often. For RAL beyond 0.1, the rate of change of number of routes is very small. We concluded that the value of RAL to be set in all future experiments would be between 0.08 to 0.1 seconds.

6.2.4 Performances of NDMRD Protocol

In this section we present the performance of NDMRD protocol with APDR as metrics as shown in Figure 11. The source data rate used is 40 kbps. Two sets of experiments were done, one with a single CBR source and the second is five CBR sources. Generally, APDR for single CBR source is better than APDR for five CBR sources. The difference in performance is due to congestion of five-sources scenario is more prominent than a single source scenario. The performance of NDMRD protocol also depends on the number of initial routes accumulated. It is observed that the smaller number of routes produce slightly better APDR compared to the higher number of routes. It

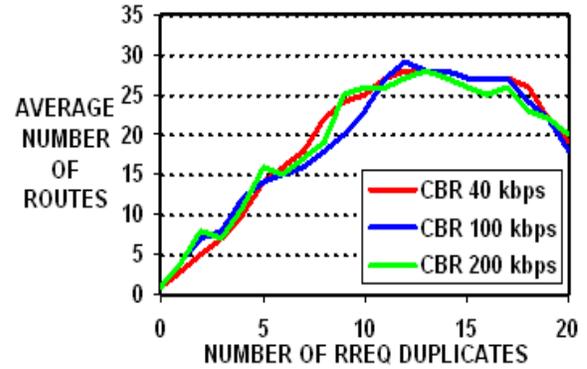


Fig. 9 Effect of RREQ Duplicates on Routes

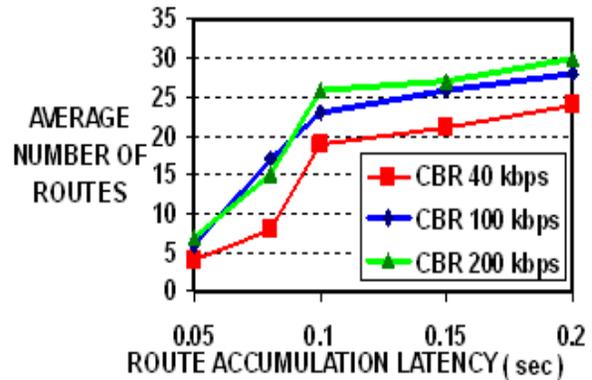


Fig 10 Effect of Route Accumulation Latency on Routes

can be attributed to the fact that the larger number of routes takes more time to generate QoS route compared to the lesser one. However the percentage difference is approximately 10%. This observation is very significant in that our setting of RAL and RREQ duplicates should reflect the possible 10% gain in term of APDR. But since our simulation experiments are based on only 40 kbps, 100 kbps and 200 kbps then as a guideline, the value of RREQ duplicates should be set between 5 and 10 duplicates and the value of RAL between 0.08 and 0.1 second.

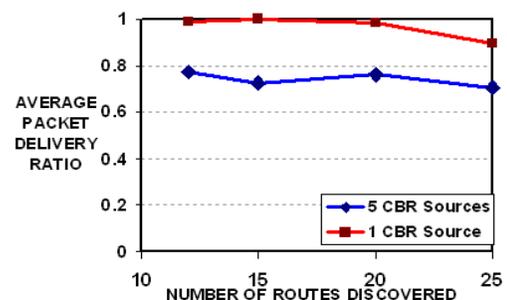


Fig. 11 Effect of Number of Routes on APDR

6.3 Performance of QOSRGA protocol

6.3.1 Impact of Source Traffic Rate Variation

The simulation experiments are carried out by keeping the maximum node velocity constant at 2 m/s, with 40 nodes. This is to study the effect of varying the source traffic rate from 20kbps to 1200 kbps. In the simulation environment 10 nodes are set as the sources to random destination nodes. The traffic rate of the sources is varied by configuring the source node with exponentially distributed inter arrival rate.

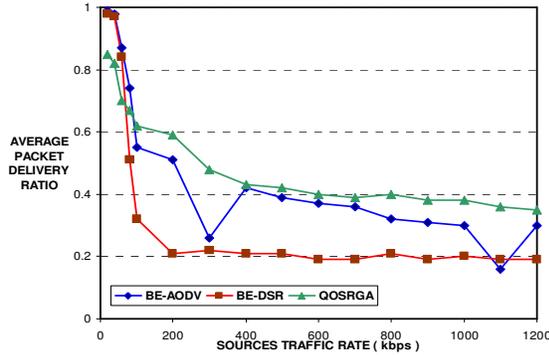


Fig. 12 Effect of Sources Traffic Rate on APDR

6.3.2 Average Packet Delivery Ratio

The traffic that we considered originated from the upper layer process model. It takes into account the data transmission rate and the control packet transmission. Control packet transmission must be considered since these packets also load the network. In order to compare the APDR, the average of total traffic send and average of the total traffic received were recorded for each traffic rate of 20 kbps to 1200 kbps. For each traffic load, the simulation repeat for 10 runs and the average readings are recorded. Each run we set the seed with different values, so as to diversify the simulation output. Hence each point in the graph is a result of 10 runs. The ratios of average total packets received to average total packets sent are taken for each traffic rate. A plot of APDR against source traffic rates is shown in Figure 12. At low source traffic rate all protocol shows similar results. The plots climbed down rapidly until the 40% where each produce different rate. BE-DSR goes further until 20% where it stays constant. QOSRGA performed a little bit better than BEAODV. When searching for the routes, QOSRGA readily acquired network information that includes the bandwidth availability and the node connectivity index (*nci*). It had chosen the route which has less probability of breaking in the near future. It chooses route which is more reliable than the other two BE protocol. When the source traffic is more than the 400 kbps, the

congestion has caused the ratio to stabilize at approximately 40%.

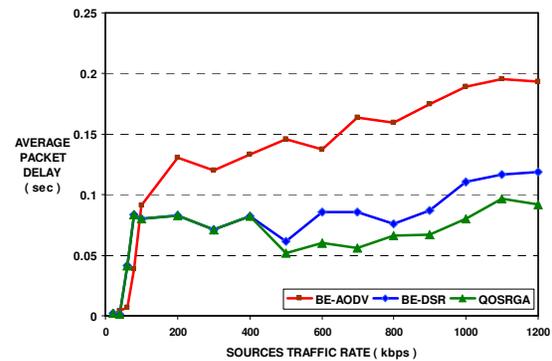


Fig. 13 Effect of Sources Traffic Rate on Average Packet Delay

6.3.3 Average End to End delay of data packets

Figure 13 depicts the variation of the average end-to-end delay as a function of the traffic rate. It can be seen that the QOSRGA protocol has a lower average delay than BE-DSR and BE-AODV under all source traffic rates. The primary reason is that the number of route discoveries is reduced in QOSRGA. Although QOSRGA has a low number of route discoveries, its delay also decreases gradually with increase in the traffic rate. An increase in the traffic rate leads to higher network load traffic in MANET. But because QOSRGA was set to choose enough bandwidth first, then due to the GA process the route selection always tries to locate the route with enough bandwidth without dropping packets as happens for BE-DSR and BE-AODV. Beyond 800 kbps, the congestion was so high that the average delay for the QOSRGA protocol increases. From 20 kbps until 1200 kbps the average packet delay for QOSRGA is less than 100 ms, which is within the standard QoS requirements.

7 Conclusion

The paper addressed the problem of QoS Routing for MANET. The goal of QoS routing is to select the most optimal route to send data packet against the constraint of bandwidth, delay and mobility. We have shown that the most viable way of selecting routes is to select from multiple non-disjoint routes. These routes couple the corresponding quality of connectivity and resources could be applied to GA algorithm. The NDMRD protocol was presented as part of complete QOSRGA protocol. The working of QOSRGA was outlined and the corresponding results were given. The proposed protocol using GA could contribute to the better understanding of how QoS routing in MANET can be properly designed.

References:

- [1] Tarique, Mohammed, et al. "Survey of multipath routing protocols for mobile ad hoc networks." *Journal of Network and Computer Applications* 32.6 (2009): 1125-1143.
- [2] Mosko, Marc, and J. J. Garcia-Luna-Aceves. "Multipath routing in wireless mesh networks." *Proc. IEEE Workshop on Wireless Mesh Networks (WiMesh), Santa Clara, USA*. 2005.
- [3] H. Rangarajan and J. Garcia-Luna-Aceves, "Using labeled paths for loop-free on-demand routing in ad hoc networks," *MobiHoc*, 2004.
- [4] M. Pearlman, Z. Haas, P. Sholander, and S. Tabrizi, "The impact of alternate path routing for load balancing in mobile ad hoc networks," *Proc. ACM MobiHoc*, 2000, pp. 3–10.
- [5] K. Wu and J. Harms, "On-demand multipath routing for mobile ad hoc networks," *Proc. EPMCC*, 2001.
- [6] S. Roy, D. Saha, S. Bandyopadhyay, T. Ueda, and S. Tanaka, "Improving end-to-end delay through load balancing with multipath routing in ad hoc wireless networks using directional antenna," *Proceeding. Of 5th International Workshop, LNCS v2918*, Jan. 2003, pp. 225 – 234.
- [7] M. Marina and S. Das, "On-demand multipath distance vector routing in ad hoc networks," in *Proc. ICNP*, 2001, pp. 14–23.
- [8] S.-J. Lee and M. Gerla, "Split multipath routing with maximally disjoint paths in ad hoc networks," *IEEE ICC*, 2001, pp. 3201–3205.
- [9] W.H. Liao, Y.C. Tseng, J.P. Sheu, and S.L. Wang. "A Multi-Path QoS Routing Protocol in a Wireless Mobile Ad Hoc Network." In *Proc. Of IEEE ICN'01: International Conference on Networking, Part II*, pages 158–167, July, 2001.
- [10] Imrich Chlamtac, Marco Conti, Jennifer J.-N. Liu, "Mobile ad hoc networking: imperatives and challenges," *Adhoc*, vol. 1, pp. 13-64, 2003. 2003.
- [11] Manet Working Group Charter, "Internet Engineering Task Force," vol. 2005, 2005.
- [12] S. B. Lee, A. Gahng-Seop, X. Zhang and A. T. Campbell, "INSIGNIA: An IP-based quality of service framework for mobile ad hoc networks," *Journ PADC*, vol. 60, pp. 374-406, 2000. 2000.
- [13] M. Gen and R. Cheng, *Genetic Algorithms and Engineering Optimization*. Canada: Wiley-Intersciences Publication, 2000,
- [14] C. Y. Lee and E. K. Antonsson, "Variable length genomes for evolution strategies," in *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2000)* 2000,
- [15] J. Inagaki, M. Haseyama and H. and Kigajima, "A genetic algorithm for determining multiple routes and its applications," in *Proceedings of the IEEE International Symposium on Circuits and Systems*, 1999, pp. 137-140.
- [16] R. L. Haupt and S. E. Haupt, "Practical Genetic Algorithms". Canada: Wiley-Interscience, 1998,
- [17] Opnet Modeler. Available: www.opnet.com
- [18] A. Nasipuri, R. Castaneda, and S.R. Das, "Performance of multipath routing for on-demand protocols in mobile ad hoc networks," *ACM/Baltzer Mobile Networks and Applications Journal*, Vol 6, 2001, pp. 339-349.
- [19] S. Corson and J. Macker, "Mobile ad hoc networking (MANET): Routing protocol performance issues and evaluation considerations," *IETF RFC 2501*, January 1999. Available: <http://www.ietf.org/rfc/rfc2501.txt> 46.
- [20] J.D. Schaffer; R.A. Caruana; L.J. Eshelman and R.Das, "A study of control parameters affecting online performance of genetic algorithms for function optimization" *Proceedings of the third international conference on Genetic algorithms*, 1989, pp51-60, George Mason University, Morgan Kaufmann Publishers Inc.
- [21] Haupt, R.L. , Optimum population size and mutation rate for a simple real genetic lgorithm that optimizes array factors, *Antennas and Propagation Society International Symposium*, 2000. IEEE, Vol.2, Iss., 2000, pp:1034-1037 vol.2
- [22] J. Abdullah and D.J. Parish, "Impact of QoS Routing Metrics for MANETs in the Pervasive Computing Environment." *Proc. IEEE WOCN2007*, Singapore.