Hybrid Optimize Strategy based QoS Route Algorithm for Mobile Ad hoc Networks

Fu Peng and Zhang Deyun

Department of Computer Science and Technology, Xi'an Jiaotong University, Xi'an 710049, China

Abstract: It is very difficult to find feasible QoS (Quality of service) routes in the mobile ad hoc networks (MANETs), because of the nature constrains of it, such as dynamic network topology, wireless communication link and limited process capability of nodes. In order to reduce average cost in flooding path discovery scheme of the traditional MANETs routing protocols and increase the probability of success in finding QoS feasible paths and we proposed a heuristic and distributed route discovery method named RLGAMAN that supports QoS requirement for MANETs in this study. This method integrates a distributed route discovery scheme with a reinforcement learning (RL) method that only utilizes the local information for the dynamic network environment; and the route expand scheme based on genetic algorithms (GA) method to find more new feasible paths and avoid the problem of local optimize. We investigate the performance of the RLGAMAN by simulation experiment bed in NS2. Compared with traditional method, the experiment results showed the network performance is improved obviously and RLGAMAN is efficient and effective.

Key word: QoS route discovery, genetic algorithms, reinforcement learning, mobile ad hoc networks

INTRODUCTION

Mobile ad hoc networks (MANETs) is a kind of new technique of wireless communication for mobile hosts. In MANETs, there is no fixed infrastructure such as base stations. MANETs are self-configuring; there is no central management system with configuration responsibilities[1,2]. All the mobile nodes can communicate each other directly, if they are in other's wireless links radio range. Such kind communications rely on other nodes to relay messages as routers, so every node in MANETs can be regarded as both the host and the router. Mobility and dynamic nodes in an ad hoc network causes frequent changes of the network topology. Ad hoc networks are characterized by a high transmission error probability, which is caused by mobility, the use of wireless links and the limited resources of nodes. Many works has been done on the fields of routing in mobile ad hoc networks[3], but the connections that support Quality of Service (QoS) requirements are always ignored. A lot of algorithms about QoS routing in the wireline network fields cannot be applied directly to MANETs for the performance constraints and dynamic network topology of MANETs. It is important that MANETs provide Qos support routing, such as available bandwidth and acceptable delay[4].

The essential task for Qos routing is to find a feasible path through the network between the source and destination that will have the necessary resources available to meet the QoS constraints[5]. It has been proved that if QoS contains at least two additive metrics, then the QoS routing is a NP-hardness

problem. So it is acceptable and necessary to develop heuristic algorithms to deal with the problem, which is to search optimize solutions with accepted cost. Currently, many studies have been done about QoS routing for MANETs by some researchers such as MANET workgroup of IETF. We had proposed a QoS route discovery method in paper[6] to solve same problem, which is based on ant algorithm and simulated annealing algorithm. In this study, we propose a novel adaptive QoS route discovery method for MANETs, based on Reinforcement learning (RL) and Genetic Algorithm (GA), named RLGAMAN, which integrate two key parts. We introduce a Qos route explore and discovery scheme based on RL to reduce the flooding in MANET. Our approach replace the complete broadcast route explore with RL based route explore, which only utilizes the local network information to decide the feasible QoS paths fit to nature of MANETs. But that may lead to the problem of stagnation route (optimal path is chosen by all the date packets) and incomplete of QoS route discovered by the route explore process. To remedy it, we propose a GA-based path extend algorithmic based, which can help remedy the RL based route discovery process and overcome the problem of local optimization solution to avoid stagnation route. Then a simulation experiments is reported to manifest the performance of our proposal. Finally, we include our main result and discuss extensions that can be made in future work.

QoS routing model for MANETs: The routing protocols for MANETs may be broadly classified as table driven protocols and on demand driven

---

Corresponding Author: Fu Peng, Department of Computer Science and Technology, Xi'an Jiaotong University, Xi'an 710049, China

protocols[3]. Table driven protocols need to maintain the global routing information about the network in every mobile node for all the possible source-destination connection and acquire to exchange routing information periodically. This kind of protocol has the property of lower latency and higher overhead. On-demand routing protocol creates routes only when the source nodes request. When a node requires a route to a destination, it initiates a route discovery process within the network. On-demand routing protocols are characterized as higher latency and lower overhead. A majority of existing research about the QoS route in MANETs is based on the two kinds of route protocols[7]. However, existing studies show that table-driven QoS protocols request globe network state information; and on-demand QoS protocols need initiates a route discovery based on flooding, which are not fit the dynamic and capability constrain in MANETs.

QoS route discovery can be implemented with distributed routing or source routing. In distributed routing, all the nodes including the source node in the QoS path will run the route algorithm to select the next hop node. In source routing, a QoS route is predetermined by the route algorithm only at the source node. RLGAMAN is a kind of heuristic distributed algorithm mixed some features of source routing. Our QoS route discovery algorithm is implemented with route explore and discovery based on reinforcement learning from source to destination and all the nodes at the reverse path as well as the QoS measurement data are stored and returned by acknowledgment packets to source nodes and the data packets are source routed.

The QoS route discovery scheme is illustrated in Fig. 1. The point A, B, C, ...,H represent the mobile nodes in MANETs. The weight of each edge is expressed with a two-tuples, which denote the available bandwidths and the delay of the relevant link. Suppose we want to find a route from Source Node A to a Destination Node F. For the pure routing algorithm such as Dijkstra, the route <A-G-F> will be the result. QoS route problem is entirely different. The shortest path route <A-G-F> may not be adequate for satisfying the requirement of QoS routing. The problem of QoS routing in MANETs can be described as following formally.
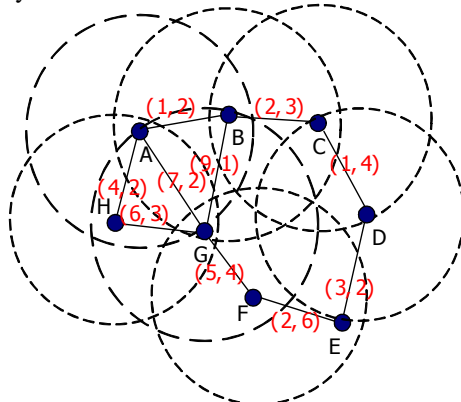

Fig.1:   QoS route discovery in MANETs

A MANETs can be modeled as a undirected graph with weight, G, $G=[V,E]$ V, is the set of the mobile nodes and the E the set of the bidirectional wireless link, for a link $e \in E^{\bullet\bullet}$ have $e = (v_i, v_j)$ and node $v_i \in V, v_j \in V, i \neq j^{\bullet} v_i, v_j$ are the neighbor nodes. G and V are both dynamic set.

For $s \in V$ is the source, $p \in \{V - \{s\}\}$ is the destination, if $e \in E$, define the metric function:

Bandwidth function: $B(e): E \rightarrow R^+$,

Delay function: $D(e): E \rightarrow R^+$

The QoS parameters of path l can be represented as:

$$\begin{cases} D(l) = \sum_{i=1}^{k-1} D(v_i, v_{i+1}) \\ B(l) = \min(B(v_i, v_{i+1}), B(v_{i+1}, v_{i+2}), \cdots, B(v_{k-1}, v_k)) \end{cases},$$ (1)

The QoS route can be described as a optimize problem:

$$(D(l) \leq d) \wedge (B(l) \geq b).$$ (2)

Where, *d* and *b* is the constant parameters.

Frame of RLGAMAN protocol: RLGAMAN use a learning system whose ability to create and maintain routes depends on the availability of inaccurate and local network information.

RLGAMAN needs to maintain some tables: route tables,, network QoS state tables, path tables.
Route tables are located at the source nodes and all the intermediate nodes in a MANET; they store complete route information for every destination as needed. As discussed before, route request packets create paths as a result of their exploration of the network. The route information arrives at the source via revised packets.

Network QoS state tables exists within each packet to store addresses and network metrics. Route request dynamical network QoS state tables with the nodes that they visit. Data packets and acknowledgments use their network QoS state table to list the complete path they follow. Also, packets record their arrival time at each node in their tables. Acknowledgement packets visit the nodes on the reverse path listed in their network QoS state table and update nodes' route table. A network QoS state is a set of QoS statistics about the performance of connections via this mobile node.

Route table keep the average round-trip delay, available bandwidth and other QoS metrics to every known destination through every neighbor nodes. This information is required to operate the learning algorithm to decide the output of the packets.

Path table are located at the source nodes and provide the foundation of population for the route discovery algorithm based on GA.

The approach of RLGAMAN is composed of three main schemes: Route exploration from source to

destination, Route registration in the reverse path, Route extend based on GA.

RL-based route explore and discovery: In the RL algorithm used by the RLGAMAN, according to the QoS metric, the observed outcome of a route selection is decided with the corresponding decision of the routing algorithm, which means the better decisions are selected via rewards and bad decisions are eliminated via the punishment. Then, the acknowledgement packets that store the reverse route and the QoS measurement data return to the source node.

The optimal policy can only be determined by adequate exploration of the system. The quality of our policy is directly limited by the quality of the model that it is calculated from. We must sample the state transitions sufficiently often to establish a good model. In our approach, *route exploration* establishment is done on-demand using limited flooding. When we need to establish a route and the destination is not the source node's neighborhood, the node will search for its route table and QoS state tables first. If there is enough information to execute RL-based decision algorithmic, intermediate nodes will forward a route request packet with QoS parameters. If no reply arrives at the explored node in time, the route entry will be deleted at the node and late coming reply packets will be ignored.

GA-based route extend: The route extend based on GA will only run at each source node, to generate and select paths for data packets based on the QoS requirements. The GA population will consist of individuals, which represent paths between the source node and potential destination nodes. We will use a variable length representation, which is expected to allow the GA more flexibility to evolve in response to changes in the network. The fitness of a path is determined from the QoS measurement data returned by acknowledgement packets, which is received in response to sending a data packet along that path.

Route registration: Upon receiving each request packet, the destination will send back an acknowledgement packet to the source along the reverse route. When receiving the acknowledgement packet, each explored intermediate node will check the net state map of ACK packet to update its network state cache. After registration, the nodes are ready to accept the real data packets of the flow.

The RL based route discovery method and GA based route expand method will be discussed in more details as following.

Description of RLGAMAN Algorithm
Reinforcement learning: Reinforcement Learning[8-10] is a general method in machine learning, which deals with the problem of how a system in a dynamic environment can learn to choose optimal actions to achieve its goals and through the learning of

trial-and-error interactions, the system can then attempt to determine the output with the input data. The idea is to adjust parameters in the direction of the empirically estimated gradient of the incensement reward.

The reinforcement learning procedure includes: environment state set, S; actions set, A; a reinforcement function $\delta : S \times A \rightarrow S$; the reward function $r : S \times A \rightarrow R$; state transition policy: $T : S \times A \rightarrow \pi(S)$. (S, A) is the expected instantaneous reinforcement from action $a (a \in A)$ in state $s (s \in S)$. $\pi(S)$ is the set of function over the set S and the learning results, how to make a transition from state s to state s' using action a.
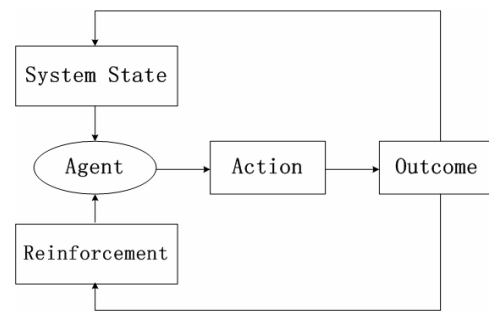

Fig. 2: Reinforcement learning model

Figure 2 illustrates the form of a standard reinforcement model. The action changes the state of the system and the value of this state transition is represented by value of reinforcement. Reinforcement learning problems are usually modeled as Markov decision processes.

RL-based route explore and discovery: The model is Markov if the state transitions are independent of the history of the system. In MANETs, each mobile node acts as a router and a host at the same time and routing information is exchanged periodically or on-demand. Furthermore, the route information depends on the dynamic network state is not very accurate. And it is difficult to collect the accurate QoS information; each mobile node only can obtain the local environment relatively that is not complete and accurate for QoS route compute.

The route algorithms based on RL have received some attention[11,12] in wireline network. In the RL algorithm used by the RLGAMAN, according to the QoS metric, each mobile node acts as an agent, which must make certain decisions, how to find a feasible path for some new connection arrival. The agent outcome (route selection) of a decision is used to reward or punish the corresponding decision of the routing algorithm so that good decisions are selected via rewards, while bad decisions are eliminated via the punishment. Then, the acknowledgement packets that

store the reverse route and the QoS measurement data return to the source node. The value of following a policy π with parameters ξ is the expected cumulative discounted (by a factor of $\lambda \in [0,1)$) reward value that can be written as:

$$V^{\pi} \blacklozenge s_t \blacklozenge = \sum_{i=0}^{\infty} \lambda^i r_{t+i} \qquad (3)$$

$$G(\xi) = \sum_{i=0}^{\infty} \lambda^i \xi_{t+i} \qquad (4)$$

the evaluate function is:

$$Q(s_t, a_t) = \xi_t + \lambda \max Q(s_{t+1}, a_{t+1}) \qquad (5)$$

and the n-step truncated return is

$$r_t^{(n)} = \sum_{i=0}^{n-1} r_{t+i} + r^n V_{t+n}(x_{t+n}) \qquad (6)$$

The optimal policy can only be determined by iterative compute with the reinforcement Q value. The quality of the policy is directly limited by the quality of the model that it is calculated from. We must sample the state transitions sufficiently often to establish a good model. In our approach, route exploration establishment is done using limited flooding. When we need to establish a route and the destination is not in the source's neighborhood list, the source node will search for its route table and network state table first. If there is enough information to execute RL-based decision algorithmic, intermediate nodes will forward a route request packet. If no reply arrives at the explored node in time, the route entry will be deleted at the node and late coming reply packets will be ignored.

GA-based route extend: Reinforcement learning belongs to swarm intelligent and is easy to lead to the problem of stagnation route. In the approach we discuss in this study, GA will generate and select paths for the packets that carry payload based on the QoS goal. That extends the RL based route discovery method discussed above to help find new feasible path and to avoid the problem of local optimize. There are some existing route methods based on GA. The study[13] use GA for routing, but aim of it is to find the more paths in wireline network. Genetic algorithm is a kind of approach based on simulated evolution. The key components of GA include genetic coding, genetic operator, fitness function and selection[9,14]. GA can be descript formally as following:

$$GA := SEA(e, f, S, E, \Sigma) \qquad (7)$$

Where, *e* denotes 0/1 encoding, *f* denotes fitness function that is used to evaluate the individuals, *S* denotes selection operators, *E* denotes the evolution operators include the crossover operator and mutate operator, Σ is the parameters space that serve to adjust the evolution direction of GA. The genetic cycle can be illustrated as:

$$P^*(t+1) = \arg\{f(P^*(t)), f(A(t+1)))\} \qquad (8)$$

The $P*(t)$ is the best individual in the $P(t)$ generation and $A(t+1)$ is the best individual in the $P(t+1)$ generation

The GA population will consist of individuals, which represent paths between the source node and potential destination nodes. Each individual of a population of individuals represents a potential solution to the problem to be solved. We will use a variable length representation, which is expected to allow the GA more flexibility to evolve in response to changes in the network[12]. In the approach we discuss in this study, a GA will run at each source node, to generate and select paths for the packets that carry payload based on the QoS goal. The fitness of a path is determined from the QoS measurement data returned by an Acknowledgement packet that is received in response to sending a packet carry payload along that path, which is received in response to sending a data packet along that path.

For the MANETs as shown in Fig.1. Consider node A as a source node and node H as a destination node. The chromosome can be coded as the sequence of φ_{AH}:

$$\varphi_{AH} = (AHGBCDEF) \qquad (9)$$

the goal value is:

$$(B(\varphi))^{\sigma_B} + (d - D(\varphi))^{\sigma_D} \qquad (10)$$

Where σ_B and σ_D are constants used to show relative importance about Delay and Bandwidth in QoS routing. In the selected population, it is used to select some individuals in order to carry out the genetic operations. All the individuals are ranked with their fitness. The selection probability is decided based on the rank. When goal value is larger, the individual fitness is higher. Through the RL-based route discovery method, many feasible paths are discovered and Acknowledgement packets bring back valid routes to the source; thus, our approach provides a way of generating new sequence φ using broadcast route request packets and reinforcement learning based search for feasible path. Crossover operation and mutation operation are also introduced in our method, but won't be discussed in details.

Upon receiving a route request, the destination will send back an acknowledgement to the source along the reverse route. When receiving this acknowledgement packet, each explored intermediate node will check the network QoS table to update its route table. After registration, the nodes are ready to accept the real data packets of the relevant data flow. In this study, we will only consider a type of MANET whose topology is not changing so fast to make the QoS routing meaningless. There may exist transient time periods when the required QoS is not guaranteed due path breaking or network state changed. The required QoS would be ensured when the paths remain unbroken.

Many multimedia applications accept such kind of QoS techniques.

Simulation and analysis: The ad hoc network is not a kind of practical technique yet, so our research is carried out based on simulations for condition constrain. Our simulation bed is based on NS2(network simulator 2) simulator and the simulation environment configuration is as following:

Table 1:   Simulation environment configuration

Simulation environment	Option and parameter
Flat Size	500 x500 meters.
Number of Nodes	80
MAC Layer	IEEE 802.11, peer to peer mode
Transport layer	User Datagram Protocol (UDP)
Traffic	Traffic enters the simulated network with the packet size of 512 bytes.
Mobility model	Waypoint
Simulation Time	500 sec

The simulations were used to compare RLGAMAN with a traditional QoS route using on-demand route policy based on flooding, in which the route request packets are broadcast completely.

In Fig. 2 and 3 we summarize the simulation results we have obtained. The RLGAMAN method uses the Delay and Link bandwidth QoS parameters. The Delay means the time it takes a packet to go from one node to another one. The link bandwidth shows the available bandwidth in corresponding link. The simulations compared a network with ordinary QoS route method based on on-demand route discovery scheme such as AODV with our RLGAMAN route discovery scheme.
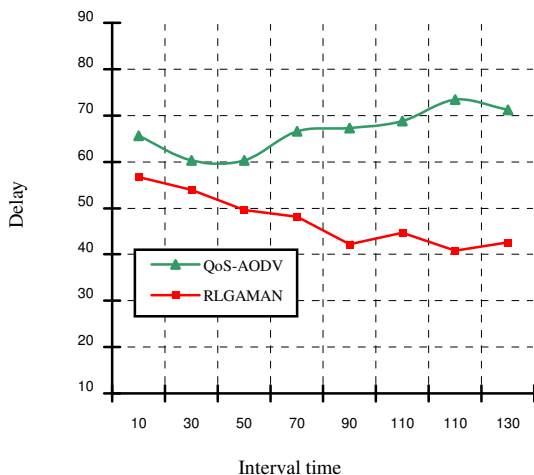

Fig. 2: End to end delay

The Fig. 2 refers to end-to-end packet delay from source to destination, Fig. 3 presents throughput of network under different QoS routing policy. The simulation results show the RLGAMAN method improves the network performance with respect to the traditional method.
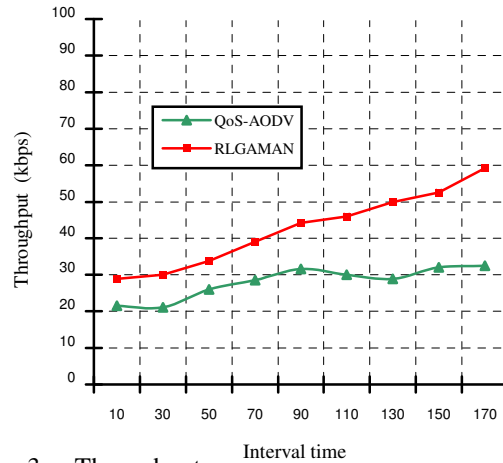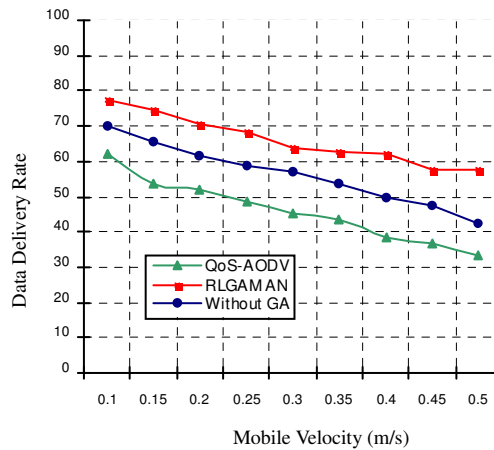

Fig. 3:   Throughput

Fig. 4:   Data delivery rate

Figure 4 shows the comparison of data delivery ratio for AODV, RLGAMAN without GA and RLGAMAN in various mobile velocities. As mobility increases the data delivery ratio of all the methods will drop down. That means that the required QoS is not guaranteed due path breaking or nodes move to other position. However, the delivery ratio of RLGAMAN drops down more slowly, which means that our method has better performance.

CONCLUSION

Most of the existing routing protocols in MANETs have been focused on only best effort data traffic. Routing schemes, which can support connections with QoS requirements, have only recently begun to receive attention. In this study, we propose a novel ad hoc route discovery method for QoS routing in MANETs, based on Reinforcement learning (RL) and Genetic Algorithm (GA), named RLGAMAN, which is composed of two main pivotal parts, the reinforcement learning based route discovery and GA based route expand.

From the results of our simulate work gathered so far, it can be aid that RL techniques can play an important role in controlling flooding in route searching to improve performance of network in environments in which the route selection is only based on local network information. We have reported simulations under various load and packet loss conditions that indicate that the GA approach can provide improvements to network QoS to remedy the deficiency of RL.

The problems of resource reservation, admission control, QoS violation recovery as well as the traffic control proposed above will be taken into account in our further coming research.

REFERENCES

1. David, R. and G.N. Ignas, 2003. Ad hoc networking in future wireless communications. Computer Communications, 6: 36-402.
2. Anonymous, 2004. IETF.: Mobile ad hoc networks charter, <http://www.ietf.org/html.charters/manet-charter.html>.
3. Royer, E.M., 1999. A review of current routing protocols for ad hoc mobile wireless networks. IEEE Personal Communications, pp: 46-554.
4. Chakrabarti, S. and A. Mishra, 2001. QoS issues in ad hoc wireless networks. IEEE Communication Magazine, 39: 142-148.
5. Chen, K., S.H. Shah and K. Nahrstedt, 2001. Cross-layer design for data accessibility in mobile ad hoc networks. J. Wireless Personal Communications, 21: 49-76.
6. Fu, P. and D.Y. Zhang, 2005. A heuristic and distributed QoS route discovery method for mobile ad hoc networks. WAIM2005, LNCS 3739, pp: 428-439.
7. Chen, S. and K. Nahrstedt, 1999. Distributed quality-of-service routing in ad hoc networks. IEEE J. On Selected Areas In Communications, 17: 1-18.
8. Kaelbling, L.P., M.L. Littman and A.W. Moore, 1996. Reinforcement learning: A survey. J. AI Research, 4: 237-277.
9. Tom. M. Mitchell, 2003. Machine Learing. McGraw-Hill, China Machine Press, Beijing.
10. Peshkin, L., 2001. Reinforcement learning by policy search. Ph.D. Thesis, Brown University, Providence, RI 02912.
11. Peshkin, L. and V. Savova, 2002. Reinforcement learning for adaptive routing. Neural Networks, IJCNN '02. Proc. Intl. Joint Conf., 2: 1825-1830.
12. Gelenbe, E., Z.G. Xu and E. Seref, 1999. Cognitive packet networks. Tools with Artificial Intelligence Proc. 11th IEEE Intl. Conf., pp: 47-54.
13. Xu, Z.G., Wu and S. Annie, 2004. Adhoc-like routing in wired networks with genetic algorithms. Ad Hoc Networks, 3: 255-263.
14. Goldberg, D.E., 1989. Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley.