Route Stability estimation in Mobile Ad Hoc Networks using Learning Automata

Mehdi Zarei, Karim Faez, Javad Moosavi Nya and Morteza Abbaszadeh Meinagh

Abstract- **This paper proposes a novel reverse on-demand routing protocol for mobile ad hoc networks based on best route selection with learning automata. Our proposed protocole capable of operating efficiently under bursty traffic conditions, is introduced. We applied our method in an optimized version of ad hoc on-demand distance vector (AODV) routing algorithm, namely Reverse AODV (R-AODV) routing algorithm. In path discovery phase, the source node select best route that has high stability between available routes. The main idea that used in proposed protocol is changing fitness of routes passing time Also with the awareness of fitness of routes, source node can select best route in sets of available route, when active route fails. For do this we changed R-AODV algorithm in some ways. In our algorithm namely Reverse AODV with learning Automata (RAODVA) routing algorithm, the route request packet not changed and it is like in AODV, but route reply packet must be change for estimation route stability and building multiple routes. Computer simulation using ns-2 simulator is performed with competition to other methods and effectiveness of the proposed method is quantitatively validated.**

Keywords* — *mobile ad hoc networks, AODV routing algorithm, Reverse AODV, Learning Automata.

I. INTRODUCTION

Mobile ad hoc networks (MANET) consist of mobile platform which communicate with each other through wireless links, without infrastructure base stations. Each node not only is a host but also as a router that maintains routes to and forwards data packets for other nodes in the network that may not be within direct wireless transmission range. Topology of a mobile ad-hoc network will often change rapidly; this behavior needs some management and solving problem of this type of networks. If source and destination nodes are not within the transmission range of each other, intermediate nodes are needed to serve as intermediate routers for the communication between the two nodes [1]. Moreover, mobile platform moves autonomously and communicate via dynamically changing network. Thus, frequent change of network topology is a main challenge for many important topics, such as routing protocol robustness, and performance degradation [2, 3].

Mehdi Zarei is a student at the Electrical, Computer and IT Eng. Dep., Qazvin Azad University, Qazvin, Iran. (Email: mehdi.nn@gmail.com).
Karim Faez is a professor at the Electrical Engineering Department, Amirkabir University of Technology, Tehran, Iran, 15914 (Email:kfaez@aut.ac.ir).
Javad Moosavi Nya is a professor at the Electrical Engineering Department, Tabriz University, Tabriz, Iran (Email: nya@yahoo.com).
Morteza Abbaszadeh Meinagh received Master degree at the Electrical, Computer and IT Eng. Dep., Qazvin Azad University, Qazvin, Iran. (Email: m_a138@yahoo.com).

In this paper we first consider the Ad-hoc On-Demand Distance Vector Routing Protocol (AODV) that is uses a demand-driven route establishment procedure and the one optimized version of this algorithms namely Reverse AODV (R-AODV). For designing our routing algorithm, we used learning automata for selection best route between available routes. For path discovery, the route with highest stability selected. Recently, some researchers have applied some types of models for solving routing protocol and present number of protocol for routing in mobile ad hoc networks. Between the types of models, for exapmle the neural networks are used to approximate and control a nonlinear system through an on-line learning process [9, 12, 13] or learning automata, due to the feedback connections in their topologies, are ideal for our problem in this paper for computing fitness of discovered routes dynamically. In these networks, learning algorithm the weights of the networks must to be updated with a dynamical learning algorithm during the control process [14, 11, 10]. With passing time, the fitness of routes changes. The rlearning automata compute fitness of routes periodically. Then we present our routing protocol namely Reverse AODV with learning automata (RAODVA) routing algorithm that is an extension of R-AODV. The RAODVA routing algorithm has good performance for high mobility environment.

II. LEARNING AUTOMATA

The aim of many intelligent systems is to be able to efficiently work in environments with unknown characteristics [17]. As a result, such systems need to possess the ability to acquire knowledge regarding the behavior of the environment. Learning automata are mechanisms that can be applied to such learning problems. Learning automaton is an automaton that improves its performance by interacting with the random environment in which it operates. The goal of a learning automaton is to find among a set of actions the optimal one, such that the average penalty received by the environment is minimized (this is equivalent to maximization of the average received reward). This obviously means that there exists a feedback mechanism that notifies the automaton about the environments response to a specific action. The operation of a learning automaton constitutes a sequence of repetitive cycles which eventually lead to the target of average penalty minimization. During each cycle, the automaton chooses an action and receives the environmental response (either a rewarding on penalizing one) triggered by the selected action. Based on this response and the knowledge gained from earlier actions the learning automaton determines the selection of the next action. A learning automaton opiates via selection at each discrete time point of an actions that is included in a set of actions $a_1, a_2, ..., a_M$. $p(n) = \{p_1(n), p_2(n), ..., p_M(n)\}$ is a vector representing the probability distribution for the M actions at time n. Obviously, $\sum_{i=1}^{M} p_i(n) = 1$. During each cycle, some action a_i, $1 \le i \le M$ is chosen with probability p_i. Upon selection of and execution of an action a_i at cycle n, the random environment responds with a rewarding or penalizing feedback ci which is used in order to update the probability distribution vector p. After the updating is finished, the automaton selects the next action according to the updated probability distribution vector $p(n+1)$. In cases when the environmental response takes only the values 0 and 1, indicating only reward or penalty respectively, the automaton is known as a P-model one. Since in many cases such a binary response can only lead to a gross estimation of the environment by the automaton, more sensitive schemes have been developed. In such schemes actions can lead to environmental feedback that is neither completely rewarding or penalizing. Automata that operate in such environments are known as Q and S-models. These kind of learning automata work with environmental responses which, after normalization, takes values in the unit interval [0,…,1]. In a Q-model, after an action ai by the automaton, the environmental response can have more than two, still finite however, possible values in the interval [0,…,1]. In an S-model environment, the environmental responses can take continuous values in [0,…,1]. The core of the operation of the learning automaton is the probability updating algorithm, also known as the reinforcement scheme. A general reinforcement scheme has the form of Eq.(1). It can be seen that after receiving feedback for the selected action $a(n)$ at cycle n, $p_i(n+1)$ changes according to weighting coefficients that reflect the distance of the environmental response $\beta(n)$ from a totally rewarding ($\beta(n)=0$) or penalizing ($\beta(n)=1$) one:

$$p_i(n+1) = \begin{cases} p_i(n) - (1-\beta(n))g_i(p(n)) \\ \quad + \beta(n)h_i(p(n)), if a(n) \ne a_i, \\ p_i(n) - (1-\beta(n))\sum_{i \ne j} g_i(p(n)) \\ \quad - \beta(n)\sum_{j \ne i} h_i(p(n)), if a(n) = a_i. \end{cases} \quad (1)$$

The functions g_i and h_i are associated with reward and penalty for action i respectively and $\beta(n)$ is a normalized metric of the environmental response. The lower the value of $\beta(n)$ the more favorable the response. According to the selection made for those functions, a number of different reinforcement schemes arise [17].

III. Reverse AODV Routing Protocol Extension using Learning Automata

R-AODV routing algorithm increased performance and when active route fails, the source node must be select best route between available routes. Stability estimation method concerned in this paper for route selection and increasing performance. Breaking radio links among nodes may easily happen due to the changing network topologies. A good design of the ad hoc routing protocol is needed to overcome the problem. Several ad-hoc routing protocols for MANETs have been proposed in recent years [5]. R-AODV algorithm solved this problem with selection the route with minimum length in available set of route that found. Here we change this stage with our approach. Link stability used in AOSV [6]. In AOSV algorithm for computing link/route stability initially, every node begins to estimate the stabilities of radio links to its neighbors and for keeping track of the link stabilities between a node and its neighbors, each node periodically broadcasts Hello message (HELLO) including the location of the broadcasting node toward its neighbors [6]. In this protocol, when a node receives Hello messages, this node first calculates the distance between neighboring node and itself from the received HELLOs and because it aware distance, evaluates the stability of radio link to the broadcasting neighbor. This information recorded for estimating stabilities of multi-hop routes in follow-up processes. In path discovery process, source node broadcast RREQ that has new link stability field. Intermediate node sends receive RREQs and rebroadcast them. The intermediate nodes rebroadcast only the RREQ with the maximum value in Route Stability among received RREQs. In proposed routing algorithm (RAODVA), when source node want communicate with a destination node, first it broadcast a RREQ packet. This stage is like to AODV algorithm. When destination receives a RREQ message, it broadcast R-RREQ massage to find source node. Each intermediate node that receive R-RREQ message, calculate route stability by following equation:

$$RS_r = \prod_{i \in L_r} ns_i \quad (2)$$

where LS_i denotes the link stability of radio links i in the route r and the link stability LS_i for a radio link i is equal to the probability of received signal power which exceeding a threshold value [6]. When source node receives R-RREQ, it will have multiple route to destination. The source node select stable route to destination. This process illustrated in figure 3. When one intermediate node move and cause link breaks, active route fails and a new route must be selected. In AODV this process done with initializing roué discovery procedure and in R-AODV with selection one available route with minimum hop count. In RAODVA new route selected between available routes with maximum stability. We added link stability field to R-RREQ. When destination node receives first RREQ, it broadcast R-RREQ. Every intermediate node that receives R-RREQ packet, it computes link stability and record it. When source node receives R-RREQ packets, it has information about stability of available routes to destination. So it may select route with highest stability. This information applied for route maintenance when data transmission started. For optimization of RAODV routing algorithm, we applied learning automata. In RAODVA routing algorithm, each available route is equipped with a learning automaton. The protocol operates as follows: After the network feedback is received for the estimation of stability at slot time t, at each route the basic choice probabilities for slot $t, p(t)$. at the beginning of RAODVA operation, at each routes automaton the choice probabilities $p(0)$ are the same for all available routes. At each time t, the basic choice probability $p(t)$ of the selected route is updated according to the network feedback information. If a route's stablity increased during slot t, then its basic choice probability is increased. Otherwise, if routes stability was fix or decreases, its basic choice probability is decreased. RAODVA updates the choice probabilities of routes stability according to the network feedback information. It is proved [?] that the choice probability of each route converges to the probability that this route has high stability.

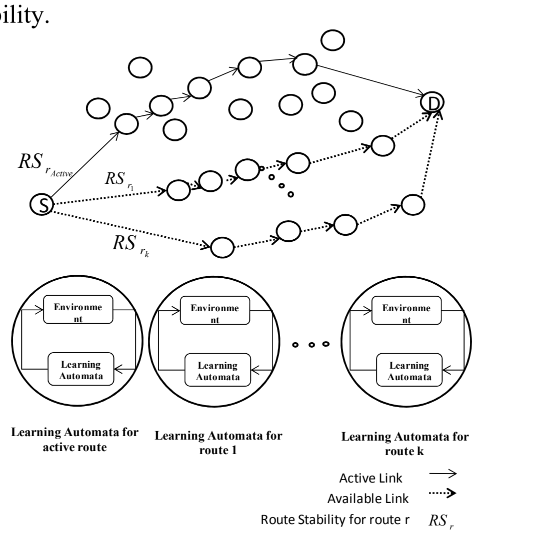

Figure 3. When active route fails or one of route between other available route increases, the source node (S) select the route with highest stability

IV. PATH MAINTENANCE IN RAODVA ROUTING ALGORITHM

When a route established between source and destination, data transmission stage can be start. In high mobility environment, link failure is a common task that will be occurring. The RAODVA routing algorithm suitable for these environments. When data transmission started, learning automata begin learning and periodically generate stability of all available routes. Source node is awareness to fitness of the routes that found with learning automata. If an intermediate node in active route move and link break or fitness of one route in other available set of routes be higher than of active route, source node cane select a stable route instead of failed route. In reverse R-AODV and AODV routing algorithms, source node select new path based on shortest path and when mobile node moves faster, it has not good performance. Here we add link stability parameter to R-AODV algorithm for selection best route between available routes set, when active route fails.

V. SIMULATION RESULTS

We used the ns-2 simulator [16] to implement our routing algorithm for comparison with RAODV and AODV routing algorithms and to check the effectiveness of the proposed method which is quantitatively validated. The area which that we used for implementation of networks is 1000m*1000m and number of nodes for this network is 10, 20, 30, 40, and 50, respectively. For speed of mobility of nodes we used 2, 5, 10, 25, 50, 75m/s, respectively for maximum speed of mobile nodes with 250 m for radio range and the nodes uniformly distributed in simulation area. Random way point used for mobility model and time for each run set to 100 seconds. For evaluation of RAODVA routing algorithm performance, we used two metrics: Delivery Rate which is the ratio of packets reaching the destination node to the total Packets generated at the source node and control packet overhead, when the number of nodes varies. Figure 6 shows packet delivery ratio of AODV, RAODV and RAODVA, by increasing number of nodes brings apparent difference between the three protocols.

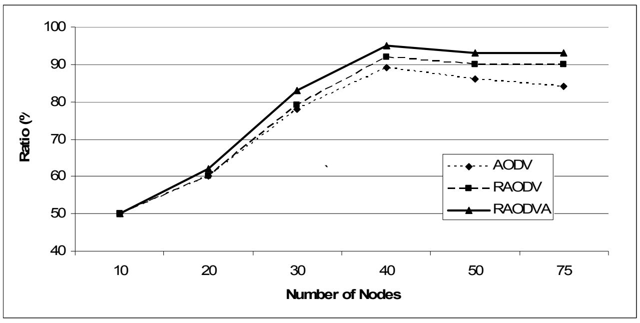

Figure 6. Packet Delivery Ratio, when the number of nodes varies

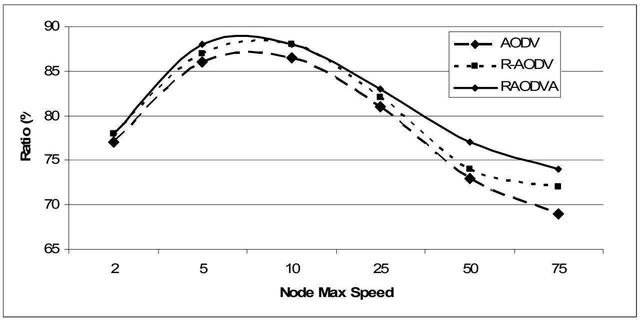

Figure 7. Packet Delivery Ratio, when node speed varies

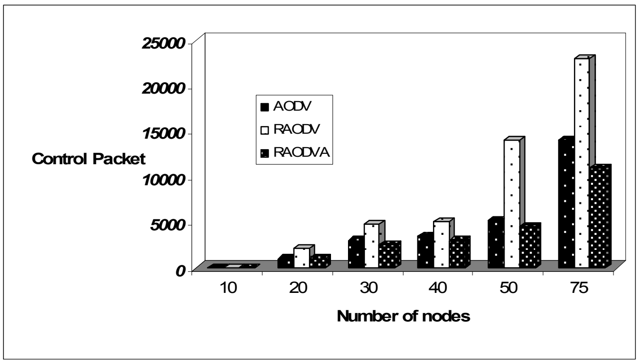

Figure 8. control packet overhead of each protocol, when number of nodes varies

Figure 7 compare the packet delivery ratio of each of the three protocols in varying mobility conditions. In the simulation, all nodes moved at the same specified speed. The graph demonstrates that RAODVA performs the best among the three protocols. Both the AODV and RAODV perform well at low speed but at high speeds, both variations of our proposed protocol perform better than other protocol. Figure 8 shows control packet overhead of each protocol, when number of nodes varies. We can see that RAODVA has lower overhead than AODV and RAODV.

VI. Conclusions

In this paper we present a new protocol for mobile ad hoc networks using learning automata for computing fitness of routes and reverse packet transmission that is an extension of AODV routing algorithm. We changed route replay packet configuration and named it R-RREQ. These packets must be broadcast with destination node for building multiple routes, also recording computing link stability performing when R-RREQ transmute through intermediate nodes. The stability of routes applies in route discovery phase. In data transmission phase, fitness of routes that computes with learning automata, applies for selection the route instead of active route for data transmission. Simulation results shows that this algorithm superior to other version of AODV algorithm. We compared our method with AODV and R-AODV routing algorithm in average to end to end delay and packet delivery ratio. This protocol suitable for environment with high mobility rate and when number of nodes increases it shows good performance.

References

[1] C. E. Perkins and E. M. Royer, "Ad hoc on-demand distance vector routing," in Proc. WMCSA, New Orleans, LA, Feb. 1999, pp. 90–100.

[2] Zhi Li and Yu-Kwong Kwok, "A New Multipath Routing Approach to Enhancing TCP Security in Ad Hoc Wireless Networks" in Proc. ICPPW 2005.

[3] Y.Kim, J.Jung, S.Lee and C.Kim, "A Belt-Zone Method for Decreasing Control Messages in Ad Hoc Netowkrs" ICCSA 2006, LNCS 3982, pp 64-72, 2006.

[4] Chonggun Kim, Elmurod Talipov, and Byoungchul Ahn, "A Reverse AODV Routing Protocol in Ad Hoc Mobile Networks" , LNCS 4097, pp. 522 – 531, 2006.

[5] S.-J. Lee and M. Gerla, "Split Multipath Routing with Maximally Disjoint Paths in Ad Hoc Networks," Proc. ICC 2001, vol. 10, pp. 3201–3205, June 2001.

[6] J. Tarng, B.Chuang and F. Wu , "Link Stability-based Routing Protocol for Mobile Ad Hoc Networks", 2006 IEEE Conference on Systems, Man, and Cybernetics October 8-11, 2006, Taipei, Taiwan

[7] Molter C., Salihoglu U. and Bersini H., "Introduction of an hebbian unsupervised learning algorithm to boost the encoding capacity of Hopfield networks", IEEE, Vol. 9, No.1, Jan 1998.

[8] Fuzzy Systems, and Knowledge Engineering, A Bradford Book The MIT Press Cambridge, Massachusetts London, England, 1996, Massachusetts Institute of Technology, 1998.

[9] Herath K. U. and Hashimoto Sh., "Automated trend diagnosis using neural networks", 0-7803-6583- IEEE, 1186-1191, 2000.

[10] Gers F. A. and Schmidhuber J., "LSTM recurrent networks learn simple context free and context sensitive languages", Transactions on Neural Networks, IEEE, 12(6):1333-1340, 2001.

[11] Sheng Ma. and Chuanyi Ji., "Fast Training of Recurrent Networks Based on the EM Algorithm. Transactions on Neural Networks", IEEE, Vol. 9, No.1, Jan 1998.

[12] J.L.Elman, "Finding Structure in Time," Cognitive Science,vol. 14, pp.179-211, 1990.

[13] M. Meral and N.S Sengor, "System Identification with Hybrid Elman Network," Proc. of the IEEE 12th on Signal Processing and Communications Applications Conference, pp. 80-83, April 2004.

[14] E.W. Saad, D.V. Prokhorov, D.C. Wunsch II, Comparative study of stock trend prediction using time delay, recurrent and probabilistic neural networks, IEEE Trans. Neural Networks 9 (6) (1998) 1456–1470.

[15] Izhak Rubin and Y. C. Liu, "Link Stability Models for QoS Ad Hoc Routing Algorithms," Vehicular Technology Conference, 2003. VTC 2003-Fall. 2003 IEEE 58th.

[16] NS, The UCB/LBNL/VINT Network Simulator (NS), http://www.isi.edu/nsnam/ns/, 2004.

[17] K. S. Narendra and M.A.L. Thathachar, *Learning Automata: An Introduction*, Prentice Hall, Englewood Cliffs, NJ, 1989