International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 1 Issue 1 –AUGUST 2013

A Reinforcement Learning Framework based Packet Routing in Ad-hoc Networks

Sri Devi Guggilam[#1], Umamaheswararao Batta[*2]

[#]*PG Scholar, Dept. of CSE, VRS & YRN College of Engg. & Technology, Chirala, AP*

[1] sridevihello@gmail.com

[*]*Assistant Professor, VRS & YRN College of Engg. & Technology, Chirala, AP*

***Abstract*—In the wireless mobile ad-hoc networks there is no fixed topology due to the mobility of nodes, interference, multipath propagation and path loss. Hence a dynamic routing protocol is needed for these networks to function properly. The routing in a wireless mobile ad-hoc networks create issues in the learning frameworks and routing in an opportunistic context. To overcome the issues, we propose the distributed adaptive opportunistic routing scheme for multihop wireless ad hoc networks. In this scheme, a reinforcement learning framework to opportunistically route the packets is going to utilizes even in the absence of the reliable knowledge about the channel statistics and network model. This scheme is shown to be optimal with respect to an expected average per-packet reward criterion. The proposed routing scheme jointly addresses the issues of learning and routing in an opportunistic context, where the network structure is characterized by the transmission success probabilities. Particularly this work leads to the stochastic routing scheme and that optimally explores and exploits the opportunities in the wireless networks.**

***Keywords*—Opportunistic routing, reward maximization, wireless ad hoc networks and routing protocol**

I. Introduction

In day today life popularity of mobile networks is goes on increasing and it is an imperative ad hoc networks support quality of services for the real time traffic. In the wireless networks, routing protocols does not consider network parameters other than the shortest path for routing decisions. Recently [30], a number of routing protocols have been proposed for mobile ad hoc networks, as part of the Defense Advanced Research Project Agency's (DARPA) Global Mobile (GloMo) program and the Internet Engineering Task Force (IETF) Mobile Ad hoc Network (MANET) working group. These protocols generally fall into one of two categories: proactive or reactive. Proactive routing attempts to maintain optimal routes to all destinations at all times, regardless of whether they are needed. To support this, the routing protocol propagates information updates about a network's topology throughout the network. In contrast, reactive or on-demand routing protocols determine routes to given destinations only when there is data to send to those destinations. Opportunistic routing is one of the important routing for the multihop wireless ad hoc networks. The researchers found that it is used to overcome the deficiencies of conventional routing as applied in the wireless settings. In the wireless networks, the conventional routings finds only a fixed path along which the packets are forwarded. So that the conventional routings fails to take advantages of the broadcast nature and opportunities provided by the wireless medium and result in unnecessary packet retransmissions. The opportunistic routing decisions, in contrast, are made in an online manner by choosing the next relay based on the actual transmission outcomes as well as a rank ordering of neighbouring nodes.

Opportunistic routing mitigates the impact of poor wireless links by exploiting the broadcast nature of wireless transmissions and the path diversity. The authors in [1] and [6] provided a Markov decision theoretic formulation for opportunistic routing. In particular, it is shown that the optimal routing decision at any epoch is to select the next relay node based on a distance-vector summarizing the expected-cost-to-forward from the neighbors to the destination. This "distance" is shown to be computable in a distributed manner and with low complexity using the probabilistic description of wireless links. The study in [1] and [6] provided a unifying framework for almost all versions of opportunistic routing such as SDF [2], Geographic Random Forwarding (GeRaF) [3], and ExOR [4], where the variations in [2]–[4] are due to the authors' choices of cost measures to optimize. For instance, an optimal route in the context of ExOR [4] is computed so as to minimize the expected number of transmissions (ETX), while GeRaF [3] uses the smallest geographical distance from the destination as a criterion for selecting the next-hop.

The opportunistic algorithms proposed by the researcher for the wireless networks. It is depends on the precise probabilistic model of wireless connections and local topology of the network. However this algorithm based on the probabilistic models, it should be "learned" and "maintained" by practically. In other words, a comprehensive study and evaluation of any opportunistic routing scheme requires an integrated approach to the issue of probability estimation. Authors in [8] provide a sensitivity analysis for the opportunistic routing algorithm given in [6]. However, by and large, the question of learning/estimating channel statistics in conjunction with opportunistic routing remains unexplored.

In this paper, we mainly focus on the issues of opportunistically routing packets in a wireless multihop network. The problems in the networks are, when zero or erroneous knowledge of transmission success probabilities and network topology is available. By using the reinforcement learning framework, we approach our works to propose a distributed adaptive opportunistic routing algorithm (d-AdaptOR). This routing algorithm that minimizes the expected average per-packet cost for routing a packet from a source node to a destination. This is achieved by both sufficiently exploring the network using data packets and exploiting the best routing opportunities.

The advantages of our proposed algorithm when it is compared to the ordinary conventional routing algorithm, reinforcement learning framework are a low-complexity, low-overhead, distributed asynchronous implementation. The significant characteristics of d-AdaptOR are that it is oblivious to the initial knowledge about the network, it is distributed, and it is asynchronous

The main contribution of this paper is to provide an opportunistic routing algorithm that: 1) assumes no knowledge about the channel statistics and network, but 2) uses a reinforcement learning framework in order to enable the nodes to adapt their routing strategies, and 3) optimally exploits the statistical opportunities and receiver diversity.

The rest of the paper is organized as follows. In Section II, we discuss about the related work of the routing algorithm. In Section III formally introduces our proposed adaptive routing algorithm, d-AdaptOR. In Section IV we summarize about the algorithm used for simulation. In Section V, we present the full simulation study of the d-AdaptOR. Finally, we conclude the paper and discuss future work in Section VI.

II. Related Works

In this section, we briefly discuss the works which is similar techniques as our approach but serve for different purposes.

Jae-Hwan Chang and Leandros Tassiulas [30], in an ad-hoc network of wireless static nodes is considered as it arises in a rapidly deployed, sensor based, monitoring system. Information is generated in certain nodes and needs to reach a set of designated gateway nodes. Each node may adjust its power within a certain range that determines the set of possible one hop away neighbours. Traffic forwarding through multiple hops is employed when the intended destination is not within immediate reach. The nodes have limited initial amounts of energy that is consumed in different rates depending on the power level and the intended receiver. We propose algorithms to select the routes and the corresponding power levels such that the time until the batteries of the nodes drain-out is maximized. The algorithms are local and amenable to distributed implementation. When there is a single power level, the problem is reduced to a maximum flow problem with node capacities and the algorithms converge to the optimal solution. When there are multiple power levels then the achievable lifetime is close to the optimal (that is computed by linear programming) most of the time. It turns out that in order to maximize the lifetime, the traffic should be routed such that the energy consumption is balanced among the nodes in proportion to their energy reserves, instead of routing to minimize the absolute consumed power.

David B. Johnson and David A. Maltz [31], this paper presents a protocol for routing in ad hoc networks that uses dynamic source routing. The protocol adapts quickly to routing changes when host movement is frequent, yet requires little or no overhead during periods in which hosts move less frequently. Based on results from a packet-level simulation of mobile hosts operating in an ad hoc network, the protocol performs well over a variety of environmental conditions such as host density and movement rates. For all but the highest rates of host movement simulated, the overhead of the protocol is quite low, falling to just 1% of total data packets transmitted for moderate movement rates in a network of 24 mobile hosts. In all cases, the difference in length between the routes used and the optimal route lengths is negligible, and in most cases, route lengths are on average within a factor of 1.01 of optimal.

Krishna Gorantala[32], the purpose of the routing protocols is to study, understand, analyze and discuss two mobile ad-hoc routing protocols DSDV and AODV where the first one is a proactive protocol depending on routing tables which are maintained at each node. The other one is a reactive protocol, which finds a route to a destination on demand, whenever communication is needed. Considering the bandwidth, throughput and packet loss, in both DSDV and AODV routing protocols, DSDV is best suited for only smaller networks and AODV is suited for general Ad-hoc networks.

Yu-Liang Chang and Ching-Chi Hsu [33], a network environment an adaptive approach for routing management will be proposed in this paper. In this approach, at first the network infrastructure is constructed by several communication groups, which are called routing groups. A routing group communicates with other routing groups via the boundary mobile hosts as forwarding nodes. In a routing group the mobile hosts are divided, by means of the dominating values, into two groups – one positive cluster and several non-positive clusters. The nodes in the positive cluster maintain the topology information of the routing group. Under such a construction environment, intra-group routing performs uni-casting and gets multiple paths, while inter-group routing performs on group level by propagating the route requests to the boundary clusters, which are called bridge clusters. This routing scheme massively reduces the message complexity that is especially important for system performance under such a resource constraint environment.

This paper [34], describes a recent simulation-based evaluation of several uni-cast routing protocols tailored specifically for mobile ad hoc networks1. Four protocols were evaluated: the Wireless Internet Routing Protocol (WIRP), a Link State (LS) algorithm with constrained LS updates, a Distance Vector variant of WIRP, and Temporally Ordered Routing Algorithm (TORA). The goal was to determine how well these routing protocols worked in specific tactical conditions, supporting a given mix of traffic. Factors varied included tactical scenario, network size, and loading. Tactical scenario included different connectivity's (e.g., "dense" versus "sparse") and link fluctuation rates (e.g., "high" versus "low"). Metrics collected included average end-to-end delay per application, path length, total efficiency, and fraction of user messages received. We found that certain protocols performed better in densely connected networks than in sparser networks (e.g., TORA), while some performed better in sparser networks (e.g., LS). One protocol, WIRP, performed well in both types of networks over the scenarios evaluated. Several remaining issues and areas for continued research are identified.

Amin Vahdat and David Becker [35], in this work, we develop techniques to deliver messages in the case where there is never a connected path from source to destination or when a network partition exists at the time a message is originated. To this end, we introduce Epidemic Routing, where random pair-wise exchanges of messages among mobile hosts ensure eventual message delivery. The goals of Epidemic Routing are to: i) maximize message delivery rate, ii) minimize message latency, and iii) minimize the total resources consumed in message delivery. Through an implementation in the Monarch simulator, we show that Epidemic Routing achieves eventual delivery of 100% of messages with reasonable aggregate resource consumption in a number of interesting scenarios.

In doing so, we build on the Markov decision formulation in [6] and an important theorem in Q-learning proved in [9]. There are many learning-based routing solutions (both heuristic and analytically driven) for conventional routing in wireless or wired networks [10]–[15]. None of these solutions exploits the receiver diversity gain in the context of opportunistic routing. However, for the sake of completeness, we provide a brief overview of the existing approaches. The authors in [10]–[14] focus on heuristic routing algorithms that adaptively identify the least congested path in a wired network. If the network congestion, hence delay, were to be replaced by time-invariant quantities, the heuristics in [10]–[14] would become a special case of d-AdaptOR in a network with deterministic channels and with no receiver diversity. In [15], analytic results for ant routing are obtained in wired networks without opportunism. Ant routing uses ant-like probes to find paths of optimal costs such as expected hop count, expected delay, and packet loss probability. This dependence on ant-like probing represents a stark difference with our approach where d-AdaptOR relies solely on data packet for exploration.

III. PROPOSED WORK

In this paper, we mainly focus on the issues of opportunistically routing packets in a wireless multihop network. The problems in the networks are, when zero or erroneous knowledge of transmission success probabilities and network topology is available. By using the reinforcement learning framework, we approach our works to propose a distributed adaptive opportunistic routing algorithm (d-AdaptOR). This routing algorithm that minimizes the expected average per-packet cost for routing a packet from a source node to a destination. This is achieved by both sufficiently exploring the network using data packets and exploiting the best routing opportunities.

The advantages of our proposed algorithm when it is compared to the ordinary conventional routing algorithm, reinforcement learning framework are a low-complexity, low-overhead, distributed asynchronous implementation. The significant characteristics of d-AdaptOR are that it is oblivious to the initial knowledge about the network, it is distributed, and it is asynchronous

The main contribution of this paper is to provide an opportunistic routing algorithm that: 1) assumes no knowledge about the channel statistics and network, but 2) uses a reinforcement learning framework in order to enable the nodes to adapt their routing strategies, and 3) optimally exploits the statistical opportunities and receiver diversity.

IV. SIMULATION WORKS/RESULTS

We have simulated our system in Java. We implemented and tested with a system configuration on Intel Dual Core processor, Windows XP and using Netbeans 7.0. We have used the following modules in our implementation part. The details of each module for this system are as follows. We have implemented and tested with the 4 modules:

Initialization stage

We consider the problem of routing packets from a source node o to a destination node d in a wireless ad-hoc network of d + 1 nodes denoted by the set = fo; 1; 2; : : : ; dg. The time is slotted and indexed by n _ 0 (this assumption is not technically critical and is only assumed for ease of exposition). A packet indexed by m _ 1 is generated at the source node o at time _m s according to an arbitrary distribution with rate _ > 0.

Transmission Stage

We assume a fixed transmission cost ci > 0 is incurred upon a transmission from node i. Transmission cost ci can be considered to model the amount of energy used for transmission, the expected time to transmit a given packet, or the hop count when the cost is equal to unity.

Acknowledgement Message Passing

We discriminate amongst the termination events as follows: We assume that upon the termination of a packet at the destination (successful delivery of a packet to the destination) a fixed and given positive reward R is obtained, while no reward is obtained if the packet is terminated (dropped) before it reaches the destination.

Relay Stage

Given a successful transmission from node i to the set of neighbor nodes S, the next (possibly randomized) routing decision includes 1) retransmission by node i, 2) relaying the packet by a node j 2 S, or 3) dropping the packet all together. If node j is selected as a relay, then it transmits the packet at the next slot, while other nodes k 6= j; k 2 S, expunge that packet. We define the termination event for packet m to be the event that packet m is either received by the destination or is dropped by a relay before reaching the destination.

V. Conclusion and Future Works

In this paper, we proposed the d-AdaptOR, a distributed, adaptive, and opportunistic routing algorithm whose performances are shown through the practically and optimal with zero knowledge regarding network topology and channel statistics. Under certain idealized assumptions, d-AdaptOR is achieved the performance of an optimal routing with perfect and centralized knowledge about network topology. The d-AdaptOR performance is measured in terms of the expected per-packet reward. We also shown that d-AdaptOR allows for a practical distributed and asynchronous compatible implementation, whose performance was investigated via a detailed set of QualNet simulations under practical and realistic networks. Simulations show that d-AdaptOR consistently outperforms existing adaptive routing algorithms in practical settings.

The long-term average reward criterion investigated in this paper so that the short-term performance is ignored in the paper. The performance of the d-adaptOR overall is fine, but fails to provide a conclusive understanding of the short-term behaviour of d-AdaptOR. An important area of future work comprises developing adaptive algorithms that ensure optimal growth rate of regret. The design of routing protocols requires a consideration of congestion control along with the throughput performance closely related issue. Incorporating congestion control in opportunistic routing algorithms to minimize expected delay without the topology and the channel statistics knowledge is an area of future research.

VI. References

[1] C. Lott and D. Teneketzis, "Stochastic routing in ad hoc wireless networks," in Proc. 39th IEEE Conf. Decision Control, 2000, vol. 3, pp. 2302–2307, vol. 3.

[2] P. Larsson, "Selection diversity forwarding in a multihop packet radio network with fading channel and capture," Mobile Comput. Commun. Rev., vol. 2, no. 4, pp. 47–54, Oct. 2001.

[3] M. Zorzi and R. R. Rao, "Geographic random forwarding (GeRaF) for ad hoc and sensor networks:Multihop performance," IEEE Trans. Mobile Comput., vol. 2, no. 4, pp. 337–348, Oct.–Dec. 2003

[4] S. Biswas and R. Morris, "ExOR: Opportunistic multi-hop routing for wireless networks," Comput. Commun. Rev., vol. 35, pp. 33–44, Oct. 2005.

[5] S. Jain and S. R. Das, "Exploiting path diversity in the link layer in wireless ad hoc networks," in Proc. 6th IEEE WoWMoM, Jun. 2005, pp. 22–30.

[6] C. Lott and D. Teneketzis, "Stochastic routing in ad hoc networks," IEEE Trans. Autom. Control, vol. 51, no. 1, pp. 52–72, Jan. 2006.

[7] E.M. Royer and C. K. Toh, "A review of current routing protocols for ad hoc mobile wireless networks," IEEE Pers. Commun., vol. 6, no. 2, pp. 46–55, Apr. 1999.

[8] T. Javidi and D. Teneketzis, "Sensitivity analysis for optimal routing in wireless ad hoc networks in presence of error in channel quality estimation," IEEE Trans. Autom. Control, vol. 49, no. 8, pp. 1303–1316, Aug. 2004.

[9] J. N. Tsitsiklis, "Asynchronous stochastic approximation and Q-learning," in Proc. 32nd IEEE Conf. Decision Control, Dec. 1993, vol. 1, pp. 395–400.

[10] J. Boyan and M. Littman, "Packet routing in dynamically changing networks: A reinforcement learning approach," in Proc. NIPS, 1994, pp. 671–678.

[11] J. W. Bates, "Packet routing and reinforcement learning: Estimating shortest paths in dynamic graphs," 1995, unpublished.

[12] S.Choi and D. Yeung, "Predictive Q-routing: A memory-based reinforcement learning approach to adaptive traffic control," in Proc. NIPS, 1996, pp. 945–951.

[13] S. Kumar and R. Miikkulainen, "Dual reinforcement Q-routing: An on-line adaptive routing algorithm," in Proc. Smart Eng. Syst., Neural Netw., Fuzzy Logic, Data Mining, Evol. Program., 2000, pp. 231–238.

[14] S. S. Dhillon and P. Van Mieghem, "Performance analysis of the AntNet algorithm," Comput. Netw., vol. 51, no. 8, pp. 2104–2125, 2007.

[15] P. Purkayastha and J. S. Baras, "Convergence of Ant routing algorithm via stochastic approximation and optimization," in Proc. IEEE Conf. Decision Control, 2007, pp. 340–354.

[16] D. P. Bertsekas and J. N. Tsitsiklis, Neuro-Dynamic Programming. Belmont, MA: Athena Scientific, 1996.

[17] S. Chachulski,M. Jennings, S. Katti, and D. Katabi, "Trading structure for randomness in wireless opportunistic routing," in Proc. ACM SIGCOMM, 2007, pp. 169–180.

[18] M. L. Puterman, Markov Decision Processes: Discrete Stochastic Dynamic Programming. New York: Wiley, 1994.

[19] D. P. Bertsekas and J. N. Tsitsiklis, Parallel and Distributed Computation: Numerical Methods. Belmont, MA: Athena Scientific, 1997.

[20] W. Stallings,Wireless Communications and Networks, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 2004.

[21] J.Bicket, D. Aguayo, S. Biswas, andR.Morris, "Architecture and evaluation of an unplanned 802.11b mesh network," in Proc. ACM MobiCom, Cologne, Germany, 2005, pp. 31–42.

[22] M. Kurth, A. Zubow, and J. P. Redlich, "Cooperative opportunistic routing using transmit diversity in wireless mesh networks," in Proc. IEEE INFOCOM, Apr. 2008, pp. 1310–1318.

[23] J. Doble, Introduction to Radio Propagation for Fixed and Mobile Communications. Boston, MA: Artech House, 1996.

[24] S. Russel and P. Norvig, Artificial Intelligence: A Modern Approach, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 2003.

[25] R. Parr and S. Russell, "Reinforcement learning with hierarchies of machines," in Proc. NIPS, 1998, pp. 1043–1049.

[26] P. Gupta and T. Javidi, "Towards throughput and delay optimal routing for wireless ad-hoc networks," in Proc. Asilomar Conf., Nov. 2007, pp. 249–254.

[27] M. J. Neely, "Optimal backpressure routing for wireless networks with multi-receiver diversity," in Proc. CISS, Mar. 2006, pp. 18–25.

[28] L. Breiman, Probability. Philadelphia, PA: SIAM, 1992.

[29] S. Resnick, A Probability Path. Boston, MA: Birkhuser, 1998.

[30] Jae-Hwan Chang and Leandros Tassiulas, Department of Electrical and Computer Engineering & Institute for Systems Research-" Energy Conserving Routing in Wireless Ad-hoc Networks"

[31] David B. Johnson and David A. Maltz, Computer Science Department Carnegie Mellon University5000 Forbes Avenue Pittsburgh, PA

15213-3891-" Dynamic Source Routing in Ad Hoc Wireless Networks"

[32] Krishna Gorantala, Umea University, Department of Computing Science, SE-901 87 UME°A, SWEDEN," Routing Protocols in Mobile Ad-hoc Networks"

[33] Yu-Liang Chang and Ching-Chi Hsu, Department of Computer Science and Information Engineering, National Taiwan University, Taipei 106, Taiwan," Routing in wireless/mobile ad-hoc networks via dynamic group construction"

[34] N. Schult, M. Mirhakkak, D. LaRocca and J. Strater-" Routing in Mobile Ad Hoc Networks"

[35] Gautam Barua and Indraneel Chakraborty, CSE Dept., Indian Institute of Technology North Guwahati, Guwahati 781031, India.-" Adaptive Routing For Ad Hoc Wireless Networks Providing QoS Guarantees"

[36] Amin Vahdat and David Becker-" Epidemic Routing for Partially-Connected Ad Hoc Networks"