Analyzing Path Accumulation for Route Discovery in Ad hoc Networks

Karim Seada, Cédric Westphal, Charles Perkins

***Abstract*—Despite the large amount of research in ad hoc networks, there are still several limitations on the scalability and performance of routing protocols, which require a more fundamental understanding of some of the basic mechanisms. In this paper, we focus on one of these mechanisms, which is *route discovery*, and analyze its extent and overhead in discovering routes to nodes in the network. [5] has augmented route discovery in AODV with a mechanism called *path accumulation*, that allows additional routes to be propagated through the network. We provide detailed analysis and simulations for the potential benefits of path accumulation and apply it on one of the prominent ad hoc routing protocols, AODV. The results show that path accumulation can improve the performance of AODV and reduce its overhead significantly.**

***Index Terms*—Wireless ad hoc networks, route discovery, path accumulation.**

I. INTRODUCTION

Ad hoc networks allow wireless mobile devices to communicate directly without the need of a preexisting infrastructure. Packets in ad hoc networks may need to go through multiple hops from a source to a destination, which requires nodes to act as relays, forwarding the packets of other nodes. Ad hoc networks routing protocols fit broadly in two categories: proactive routing protocols, and reactive protocols. Proactive protocols, such as OLSR [6] track the topology changes of the network so as to keep route information between any source and destination available at all times. Reactive protocols such as AODV [1] or DSR [7], initiate a new route as needed, at the start of a connection. Proactive protocols are better suited for low mobility environment, while reactive protocols focus fit dynamic networks better.

A main operation in a reactive ad hoc network protocol is for a source node to discover a route to its destination. In this paper, we consider on-demand routing protocols, in which this route discovery is performed only when a packet needs to be sent to a destination and no route already exists. Previous studies have shown the scalability of AODV and its robustness to mobility [2], [3]. In addition, it is one of the earliest protocols to be standardized for ad hoc networks [4]. In AODV, the route request is normally flooded to the whole network which makes it the most expensive operation in terms of overhead and makes mechanisms that reduce the amount of route discovery valuable for improving performance.

In this paper we study path accumulation, as proposed in [5], and analyze its potential in reducing the route discovery overhead. The idea of path accumulation is to include route information about the intermediate nodes in the control packets, allowing nodes receiving the control packets to learn routes to additional nodes. These routes are then stored for a period of time, called the Active Route Time-out, and are available without initiating a route discovery. This potentially reduces the need for flooding route requests when communication to these destinations is needed.

Reducing the route discovery overhead is critical in wireless ad hoc networks: a broadcast of RREQs through the network creates the so-called *broadcast storm* [8], in which a group of nodes attempts to communicate all at the same time over the same wireless channel. This creates contention for the medium, congestion, packet losses for the control packets, delays for the packet which make it through, and an overall degradation of the performance of the network. Reducing the number of route requests not only saves bandwidth, but also reduces the impact of the broadcast storms.

Reducing the number of RREQs means increasing the *efficiency* of each RREQ: fewer RREQs accomplish the same discovery work. This means that, in a dynamic network, the routes collected by the RREQs will be fresher, and the network will be more reactive to dynamic changes.

We model the extent of network discovered when

Karim Seada and Charles Perkins are with the Nokia Research Center, Palo Alto, CA. Cédric Westphal is with Nokia Network, Mountain View, CA. {karim.seada, cedric.westphal, charles.perkins@nokia.com}

using path accumulation and compare to route discovery without path accumulation, which gives us an indication of the potential reduction in control overhead. We perform detailed ns2 simulations with wireless MAC and physical layer models and show that the results have trends similar to the analysis. We also show that path accumulation can improve the average path length and reduce the data load by optimizing the routes between nodes.

In Section II we explain how path accumulation functions in more detail. Section III contains the analysis of route discovery overhead and Section IV contains the simulation results. The evaluation of the data path length is presented in Section V and some conclusions are presented in Section VI.

II. PATH ACCUMULATION

In AODV, a source that has a packet to a destination needs first to discover a route to this destination. If the source has no route already in its routing table to that destination, it sends a route request (RREQ) which is flooded to the whole network. When the destination or a node that has a route to the destination receives this RREQ, it replies back with a RREP. When the source receives the RREP it can start sending data packets along the route from which the RREP arrived. Intermediate nodes receiving the RREQ create entries in their routing table to the source, similarly nodes receiving the RREP create entries to the destination. The routing table entry contains the next-hop to the corresponding node. Accordingly, a route between a source and a destination is constructed hop-by-hop along the path taken by the RREP and data packets do not need to contain the whole route. For more details about AODV, see for instance [1].

The idea of path accumulation [5] is to allow nodes receiving and forwarding control packets to record their identity in the packet and eventually learn about other nodes in the path between the source and destination. Each RREQ and RREP contains a source route for the nodes along the path, so that each node can have a routing table entry to the rest of the nodes. The main benefit of obtaining the additional routing entries is to reduce the route discovery overhead by eliminating some of the RREQs that would be required to discover these nodes. Since RREQs are the major source of control overhead due to flooding the whole network, any reduction in RREQs is expected to improve the performance significantly. The tradeoff is that the RREQ and RREP packet header will become larger to accommodate the source route.

Figure 1 (inspired by [5]) clarifies the difference between route discovery without path accumulation and route discovery with path accumulation. Without path accumulation, intermediate nodes learn only about the source and destination, while with path accumulation each node learns about every other node in the path.

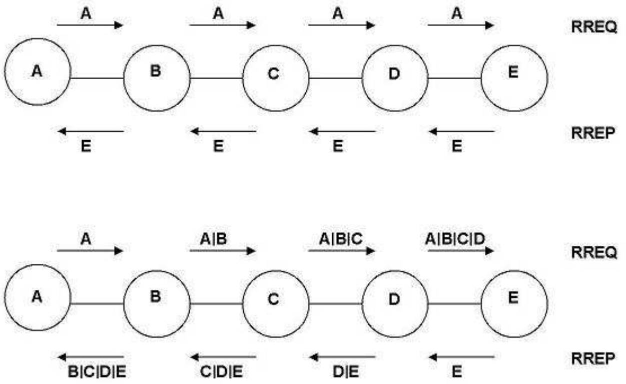

Fig. 1. Route Discovery Without and With Path Accumulation

III. ROUTE DISCOVERY OVERHEAD

We evaluate the route discovery overhead by counting the number of RREQs required to discover new paths in the network. First, we show an approximate analysis for the asymptotic number of RREQs required to discover all paths in the network. We define a path as a source-destination route entry. The total number of paths is $n(n-1)$ which is $O(n^2)$, where n is the number of nodes. Without path accumulation, at each RREQ, each node learns a path to the source and nodes along the source-destination route learn a path to the destination. Thus the number of paths discovered by a RREQ is $O(n)$. Therefore, to discover all paths, $O(n)$ RREQs are required.

With path accumulation, at each RREQ, each node learns about all nodes along its path from the source and nodes along the source-destination route learn about all nodes in that route. Thus the number of paths discovered by a RREQ is $O(n\sqrt{n})$. Therefore to discover all paths $O(\sqrt{n})$ RREQs are required.

In the following sections we provide more detailed analysis and simulation results for the number of paths discovered at each RREQ.

A. Analysis

We perform analysis for the percentage of paths discovered by each RREQ without path accumulation and with path accumulation.

1) Without Path Accumulation: The total number of paths in the network are $b = n(n-1)$, where n is the number of nodes. Consider the network is initially empty of any route information.

Theorem III.1: Given any $1 > \epsilon > 0$, the asymptotic number of RREQs required to discover a fraction $(1-\epsilon)$ of all the paths in an n-node static ad-hoc network is $O(n)$.

Proof: Let y_1 be the number of new paths discovered by the first route request, y_2 the number of new paths discovered by the second one, and generally, y_k the number of new paths discovered by the k^{th} RREQ.

In addition, assume that at time 0, all nodes discover paths to their direct neighbors either through Hello messages, or when neighbors broadcast the first RREQ: $y_0 = dn$, where d is the node degree (average number of neighbors)

If no paths are known in the network, a RREQ should discover $a = (n-d-1) + (l-1)$. The first term are the paths learned by all nodes receiving the RREQ to the source (subtracting the neighbors since they are already counted in y_0), and the second term are the paths to the destination learned by nodes forwarding the RREP. l is the average number of hops between two nodes.

When there are paths already known, the equation should be scaled down by the paths already known:

$$\begin{aligned} y_1 &= a\frac{b-y_0}{b} \\ y_2 &= a\frac{b-(y_0+y_1)}{b} \\ y_k &= a\frac{b-\Sigma_{i=0}^{k-1}y_i}{b} \end{aligned} \tag{1}$$

This implies

$$\begin{aligned} y_{k+1} - y_k &= -a\frac{y_k}{b} \quad \text{that is} \\ y_{k+1} &= (1-\frac{a}{b})y_k \text{ for } k > 0, \text{ or} \\ y_k &= a(1-\frac{a}{b})^{k-1}\frac{b-dn}{b} \end{aligned} \tag{2}$$

Since at least 1 path has to be discovered by a new route request, that of the destination, y_k is modified to prevent equation 2 from decreasing below 1 as $k \to \infty$:

$$y_k = \max\left[a(1-\frac{a}{b})^{k-1}\frac{b-dn}{b}, 1\right] \tag{3}$$

Using 3, we can formally check the asymptotic behavior of the number of route requests required to discover all the paths: set $\alpha > 0$ to be a constant to be determined later. Summing over the first αn terms, we get:

$$\begin{aligned} \Sigma_{i=1}^{\alpha n} y_i &= a\Sigma_{i=1}^{\alpha n}(1-\frac{a}{b})^i \\ &= a\frac{1-(1-a/b)^{\alpha n}}{1-(1-a/b)} \\ &= a\frac{1-e^{-\alpha}}{\frac{1}{n}} \\ &\sim (1-e^{-\alpha})n^2 \\ &= O(n^2) \end{aligned} \tag{4}$$

where we used $a = O(n), b = O(n^2), y_0 = O(1)$ and $\lim(1-1/n)^n = e^{-1}$.

Picking α such that $\alpha > -\log(\epsilon)$ completes the proof. ▮

2) With Path Accumulation: A similar result can be demonstrated for path accumulation, with the asymptotic behavior $O(\sqrt{n})$ instead of $O(n)$.

Theorem III.2: Given any $1 > \epsilon > 0$, the asymptotic number of RREQs required to discover a fraction $(1-\epsilon)$ of all the paths in an n-node static ad-hoc network is $O(\sqrt{n})$.

Proof: We only present a sketch of the proof, as it is similar to that of Theorem III.1, with the main difference that a is now replaced by a_{pa} defined as:

$$a_{pa} = (n-d-1)*(l-1) + l(l-3). \tag{5}$$

All nodes forwarding the RREQ (except neighbors of the source which already know the source) will learn about paths to all other nodes in their route to the source, and nodes forwarding the RREP will learn about paths to all other nodes (except their neighbors and the source which is already counted) in their route to the destination.

The equation

$$y_k = \max\left[a_{pa}(1-\frac{a_{pa}}{b})^{k-1}\frac{b-dn}{b}, 1\right] \tag{6}$$

is obtained in a similar manner.

As in the previous section, we can confirm that about $\sqrt{n}$ RREQs are required to find out all n^2 routes: summing of the first $\sqrt{n}$ terms yield a result which is $O(n^2)$ ∎

Theorem III.1 and theorem III.2 basically state that path accumulation in AODV-PA reduces the number of RREQs by a factor $\sqrt{n}$. This is a dramatic improvement, as it means that each RREQ will gather $\sqrt{n}$ more routes. This means that the cost of incurring a broadcast storm is off-set by gathering $\sqrt{n}$ more routes, making it more worthwhile. This also means that, in a dynamic environment, the information of the last $\sqrt{n}$ route requests is sufficient for AODV-PA, not that of the last n as for AODV. The last $\sqrt{n}$ RREQs are of course more fresh than the last n, providing more recent routes to the network that are less likely to be broken, reducing even further the need for RREQs.

IV. SIMULATIONS

In order to validate our analytical model, we conducted simulations of the model defined by equation 3 for AODV and equation 6 for AODV-PA.

We run NS-2 [9] simulations of static networks of 100 and 1000 nodes at a constant density of around 8 neighbors/node in order to count the number of paths discovered at each RREQ and compare the simulation results to the analysis. The average number of hops is around 5 in the 100-node network and 12 in the 1000-node network. The simulation uses an 802.11b with CTS/RTS MAC and a two-ray ground propagation model at the physical layer. Each node sends exactly one packet to a random destination. If a route entry already exists to the destination, no RREQ is generated.

Figure 2 shows the percentage of paths discovered as a function of RREQs in the 100-node network. The simulation results have similar trends to the analysis both without and with path accumulation. At some points, the analysis exceeds the simulation due to the effect of MAC contention and collisions on the simulation results in reducing the amount of paths discovered. Figure 3 shows similar trends in the 1000-node network. Without path accumulation, the simulation and analysis almost match, while with path accumulation the discovery of the last portion of undiscovered paths in the simulation becomes more challenging, as the packets need to propagate more hops dealing with collisions and packet drops.

Our analysis for AODV-PA seems to provide only a first order approximation. It is our goal to refine the analytical model to obtain the same perfect match between analysis and simulation for AODV-PA as we do for AODV.

The reduction in the number of route discovery packets shown in the previous analysis and simulations comes at the cost of increasing the RREQ and RREP packet size as a result of including extra entries for intermediate path nodes. The reduction in channel access should have a much more significant effect than this increase in packet size due to the overhead of channel contention and sending extra packets.

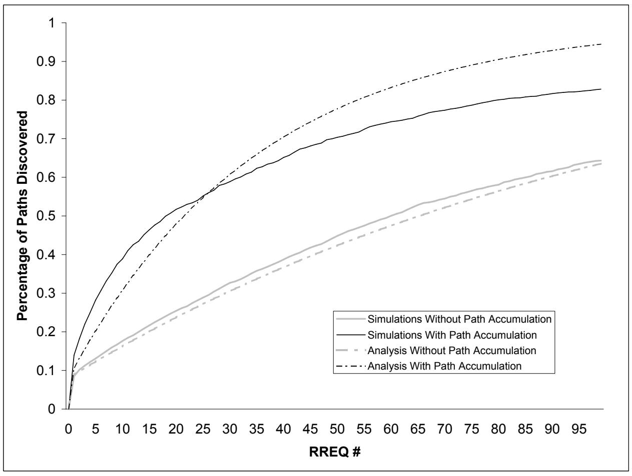

Fig. 2. Number of Paths Discovered as a Function of the Number of Route Requests in a 100-Node Network

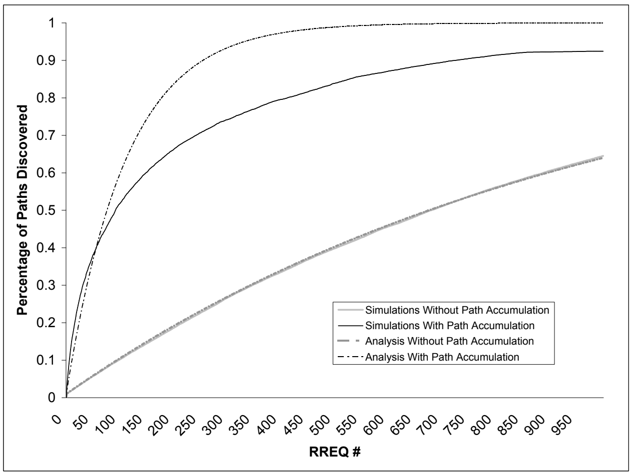

Fig. 3. Number of Paths Discovered as a Function of the Number of Route Requests in a 1000-Node Network

V. DATA PATH LENGTH

To view the influence of AODV-PA on reducing the broadcast storm problem, we look at the path which is eventually found by the routing protocol. As we stated above, flooding the network creates a burst of contention for the channel, which results in packets being lost. The RREQ which arrives first at the destination, and thus triggers the Route Reply, might not be the one following the optimal path in terms of path length.

In addition to the reduction in route discovery overhead, path accumulation improves the data path length in terms of number of hops, reducing the number of transmissions required in delivering a packet between a source and a destination. The reason is that the additional route information contained in the packet header allows nodes to optimize their entries and learn about shorter routes to different destinations.

We run NS-2 simulations of sizes varying from $n = 100$ to $n = 1000$ nodes in a static network at a constant density of about 8 neighbors/node. We compute the average path length without and with path accumulation. We compare the path length to the average shortest path computed using Dijkstra (average number of hops ranges from around 5 hops at 100 nodes to 12 hops at 1000 nodes) and show the ratio. Figure 4 shows path accumulation improving the path length around 5% at 100 node up to 7% at 1000 nodes. As the network size increases, the path length optimality worsens and the difference widens in favor of path accumulation. For flows with heavy traffic, an extra hop can lead to a significant reduction in network load, thus again underlying the benefit of path accumulation in this context.

VI. CONCLUSION

In this paper, we studied the potential benefits of path accumulation using analysis and detailed simulation. The results show the possible reduction in route discovery overhead and data load. We constructed an analytical model to compute the number of RREQs required to build the routing information for AODV and AODV-PA. The model shows that, for a network with n nodes, AODV-PA requires only $\sqrt{n}$ RREQs in a static network, compared with n for AODV.

We validated the model through NS-2 simulations, and looked at an immediate and practical consequence of the reduction in the number of RREQs: AODV-PA finds routes with shorter path length on average, despite not following an exhaustive search, as AODV does.

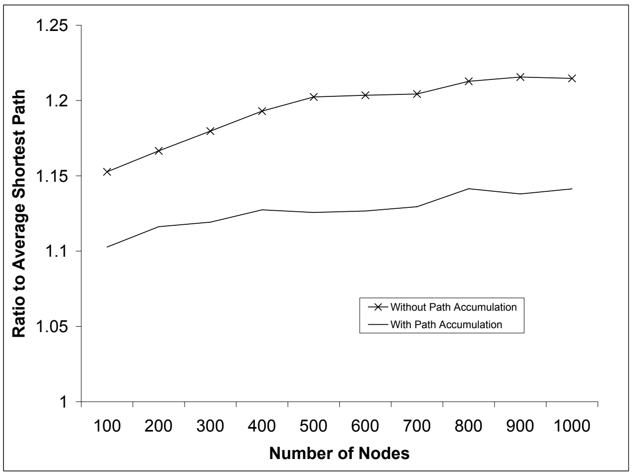

Fig. 4. Path Length With and Without Path Accumulation

The paper opens some new research areas that we intend to cover in the future:

- First, providing as good an analytical model for AODV-PA as we do for AODV is our primary research target.
- Second, the current analysis considers a static network which is important for understanding the base performance. Adding mobility will make the analysis more challenging and we plan to extend the analytical model to include mobility and apply it to a wider variety of networks.
- Third, path accumulation has the drawback of increasing the control packet size. We believe that reducing the access of wireless channel by reducing the number of packets sent has a more significant effect than the size of the packet, nevertheless we plan to evaluate networks with larger loads and higher congestion to verify the viability of path accumulation under these conditions.

REFERENCES

[1] Charles Perkins , Elizabeth M. Royer. *Ad hoc On-Demand Distance Vector Routing.* Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications, New Orleans, LA, February 1999, pp. 90-100.

[2] Josh Broch, David A. Maltz, David B. Johnson, Yih-Chun Hu, and Jorjeta Jetcheva. *A Performance Comparison of Multi-Hop Wireless Ad Hoc Network Routing Protocols.* In Proceedings of the Fourth Annual International Conference on Mobile Computing and Networking (MobiCom'98), ACM, Dallas, TX, October 1998.

[3] Samir R. Das, Charles E. Perkins, Elizabeth M. Royer and Mahesh K. Marina. *Performance Comparison of Two On-demand Routing Protocols for Ad hoc Networks.* IEEE Personal Communications Magazine special issue on Ad hoc Networking, February 2001, p. 16-28.

[4] Charles E. Perkins, Elizabeth M. Belding-Royer, and Samir Das. *Ad Hoc On Demand Distance Vector (AODV) Routing.* IETF RFC 3561.

[5] Sumit Gwalani, Elizabeth M. Belding-Royer, and Charles Perkins. *AODV-PA: AODV with Path Accumulation* IEEE International Conference on Communication, ICC'03.

[6] Thomas Clausen and Philippe Jacquet, *Optimized Link State Routing Protocol (OLSR)*, IETF RFC3626, October 2003.

[7] David Johnson, David Maltz and Yih-Chun Hu, *The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks (DSR)*, IETF MANET working group, draft-ietf-manet-dsr-10.txt, work in progress, July 2004.

[8] Sze-Yao Ni, Yu-Chee Tseng, Yuh-Shyan Chen and Jang-Ping Sheu, *The broadcast storm problem in a mobile ad hoc network* in Proceedings of the 5th annual ACM/IEEE international conference on Mobile computing and networking, pp.151-62, 1999.

[9] Steven McCanne and Sally Floyd, *NS-2 Network Simulator* http://www.isi.edu/nsnam/ns.