Minimizing the Routing Delay in Ad Hoc Networks through Route-Cache TTL Optimization

Ben Liang and Zygmunt J. Haas

School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853, USA
{liang, haas}@ece.cornell.edu

Abstract. This paper addresses the issue of minimizing the routing delay in ad hoc on-demand routing protocols through optimizing the Time-to-Live (TTL) interval for route caching. An analytical framework is introduced to compute the expected routing delay when the source node has a cached route with a given TTL value. Furthermore, a numerical method is proposed to determine the optimal TTL of a newly discovered route cached by the source node. We present simulation results that support the validity of our analysis.

1 Introduction

Node mobility and the lack of topological stability make the routing protocols previously developed for wireline networks unsuitable for ad hoc networks[9][8][11]. A popular family of ad hoc routing protocols are the reactive routing protocols, also called *on-demand* routing protocols. In these protocols a node is not required to maintain a routing table (although route caches may be kept), but instead a route query process is initiated whenever it is needed. Routing protocols such as ABR, AODV, DSR, the IERP of ZRP, and TORA are examples of reactive protocols[11][1].

In an on-demand routing protocol, a newly discovered route should be cached, so that it may be reused the next time that the same route is requested. However, prolonged storage of a route cache may render it obsolete. When an invalid route cache is used, extra traffic overhead and routing delay is necessary to discover the broken links. Depending on the implementation details, data and/or control packets are delivered over part of the cached route that is still valid, before the broken link can be discovered.[2]

One approach to minimize the effect of invalid route cache is to purge the cache entry after some Time-to-Live (TTL) interval. If the TTL is set too small, valid routes are likely to be discarded, but if the TTL is set too large, invalid route-caches are likely to be used. Thus, an algorithm that optimizes the TTL setting is necessary for the optimal performance of an on-demand routing protocol.

As far as we are aware, there is very little reported work in literature that addresses the issue of ad hoc route-cache TTL optimization. Most existing on-demand protocols,

[1] Due to the page limit, the individual references to these protocols are omitted.

[2] It is possible to employ *proactive* route-cache invalidation initiated by the up-stream node of a broken link, whether or not the link is part of an active route presently delivering data. However, this can lead to large control overhead when the network topology changes frequently. Proactive techniques are outside the scope of this paper.

such as AODV, DSR, and TORA, employ route caching in various forms. In AODV, a discovered route is associated with an "active route time-out" value that dictates the duration within which the route can be used. This time-out value is static and identical throughout the network. In DSR and TORA, a cached route is kept indefinitely (i.e. TTL=∞), until a broken link in the route is detected during data transmission. In this work, we study the TTL optimization adaptive to each cached route.

In [10], case study based on DSR has suggested that route caching can reduce the average latency of route discovery by more than 10-fold. Further simulation studies reported in [1], [7], [4], [2], and [3] have confirmed the effectiveness of route caching in on-demand routing protocols. However, [7], [4], [2], and [3] have also drawn the conclusion that the indefinite route-cache, as is employed in DSR, can lead to many stale routes and hence degrade the routing performance. In addition, [2] and [3] have demonstrated the need for determining a suitable time for the route-cache expiration. The simulation results in [3] have further shown a case study of the optimal route-cache expiry time obtained by exhaustive search. In this work, we approach the problem of adaptive route-cache TTL optimization through analytical studies.

We consider the problem of optimizing the TTL of a cached route in order to minimize the expected routing delay of the next request of the same route (i.e., the same source and destination pair). In Section 2, we explain the network model under consideration. In Section 3, we introduce analytical and numerical frameworks to compute the optimal TTL and the corresponding expected routing delay. In Section 4, we present simulation results that support the validity of our analysis, study the system parameters that affect the optimal TTL, and show the performance gain achieved by the optimal TTL. Finally, concluding remarks are provided in Section 5.

2 Network Model

We consider a mobile ad hoc network consisting of a set V of nodes. At any time instant, an edge (u, v), where $u, v \in V$, exists if and only if node u can successfully transmit to node v. In this case, we say that the link from node u to node v is *up*. Otherwise, the link is *down* or has *failed*.

In the modeling of general communication networks, it is usually assumed that all edge failures are statistically independent [6]. The modeling of dependent link failures generally requires an exponentially large number of conditional probability distributions. Therefore, though unrealistic, the independence assumption greatly simplifies the analysis of network performance. In this paper, we assume that all links have independent and identical up-time distribution $F_u(t)$ and down-time distribution $F_d(t)$.

We assume that route requests to a destination node n_d arrive at the source node n_s as a stream that has identically distributed inter-arrival intervals with a general distribution $F_a(t)$. We consider only non-trivial networks where the average time between topology changes is smaller than the average route-search delay.[3] Therefore, we assume that the route-request inter-arrival time is much larger than the route-search delay, since, otherwise, a valid route is already found at the last route request. Namely, a burst

[3] Otherwise, the only suitable routing approach is to flood data packets throughout the network.

of data packet train sent to a common destination within a very small time frame would constitute a single route request.

When a route request is made due to a data packet arrival, if n_s has a cached route to n_d, it immediately sends out the data packet using the cached route. If the cached route is valid, we assume that this operation does not incur any routing delay. However, if the cached route is invalid, the intermediate node on the up-stream end of a failed link notifies n_s via a route-error packet. In this case, and in the case that n_s does not have a cached route to n_d, the pre-defined routing protocol [4] is employed to search for a new route to n_d. We further assume that n_s renews or re-computes the TTL of a cached route to n_d each time a packet is successfully sent through the cached route. A cached route is purged when its TTL expires.[5]

We assume that all data and control packet transmissions across a link incur an average delay of L seconds.[6]

3 Optimizing the Route-Cache TTL to Minimize Routing Delay[7]

3.1 Computing the Expected Routing Delay

Suppose the source node n_s has a cached route to the destination node n_d, which is validated by the last route request and has a TTL value of T seconds. Let D be the number of hops in this route.

Let the next route request to n_d arrive at time t_a after the n_s-to-n_d route is cached. Then, from Section 2, t_a has distribution $F_a(t)$. Let $f_a(t)$ be the density function of t_a, and let $f_a{}^*(s)$ be the Laplace transform of $f_a(t)$. Furthermore, let $f_c(t)$ be the density function of t_a given $t_a < T$. Then, the Laplace transform of $f_c(t)$ is $f_c{}^*(s) = -\frac{1}{F_a(T)} \sum_{\xi \in \text{poles of } f_a{}^*(s-z)} Res_{z=\xi} \frac{1-e^{-zT}}{z} f_a{}^*(s-z)$, where $Res_{z=\xi}$ denotes the residue at the pole $z = \xi$.

Let $f_u(t) = dF_u(t)/dt$ be the density function of the link up-time and $f_u{}^*(s)$ be its Laplace transform. The residual lifetime of a link in the cached route has the density function $f_r(t) = \frac{1}{\mu_u}[1 - F_u(t)]$, where μ_u is the mean up-time of a link. The Laplace transform of $R_r(t) = 1 - \int_0^t f_r(\tau)d\tau$ is $R_r{}^*(s) = \frac{1}{s} - \frac{1}{\mu_u s^2}[1 - f_u{}^*(s)]$.

Let X_i be the minimum of the residual lifetimes of the first i links in the cached route. Let $f_{X_i}(t)$ and $F_{X_i}(t)$ be its density and distribution functions, and let $f_{X_i}{}^*(s)$

[4] The exact mechanism of the on-demand routing protocol is not important here.

[5] Some on-demand protocols allow an intermediate node that has a cached route to the destination reply to the source initiated route request. Such schemes have been shown to significantly improve the routing performance. However, the quantitative effect of the stale routes provided by the intermediate nodes is not well understood. Therefore, in this work, we only consider the TTL of route caches kept by a source node.

[6] For example, in the case study of [10], the average delay is shown to be 14.5 ms/hop. Although packets may be different in length, in a wireless ad hoc network operating at medium to high load, the predominant factor in the aggregate delay of packet transmission across a link is the queuing delay in the MAC layer due to the contention of the shared wireless medium.

[7] Due to the page limit, in this section, we only give a brief outline of the important results and leave out the details to the long version of this paper.

and ${F_{X_i}}^*(s)$ be the corresponding Laplace transforms, respectively. Let $R_{X_i}(t) = 1 - F_{X_i}(t)$ and $R_r(t) = 1 - F_r(t)$, and ${R_{X_i}}^*(s)$ and ${R_r}^*(s)$ be their Laplace transforms, respectively. Then ${f_{X_i}}^*(s)$ can be determined through the following recursion: ${R_{X_i}}^*(s) = -\sum_{\xi \in \text{poles of } {R_r}^*(s-z)} Res_{z=\xi}\, {R_{X_{i-1}}}^*(z){R_r}^*(s-z)$, along with ${f_{X_i}}^*(s) = s{F_{X_i}}^*(s) - {F_{X_i}}^*(0+) = 1 - s{R_{X_i}}^*(s)$.

Let $Q_i(T)$ be the probability that, when a route request arrives before the TTL expires, the first i links of the cached route have not failed. We can obtain $Q_i(T) = -\sum_{\xi \in \text{poles of } {f_{X_i}}^*(-s)} Res_{s=\xi}\, \frac{{f_c}^*(s)}{s}{f_{X_i}}^*(-s)$. Define $Q_0(T) = 1$. The expected routing delay of the next route request, when the TTL of a D-hop cached route is set to T, is $C(T) = 2L\left[D + F_a(T)\sum_{i=1}^{D} i\,(Q_{i-1}(T) - Q_i(T)) - F_a(T)Q_D(T)D\right]$.

The above analytical framework provides a means for evaluating the expected routing delay given the TTL value. However, it is very likely that the optimal TTL value is more important to a system designer. In the next section, we provide a numerical method to compute the optimal TTL.

3.2 Determining the Optimal Route-Cache TTL

Let $q(\tau)$ be the probability that a given link in the cached route is still up at time τ after the last route request. The expected routing delay as defined in the previous section has the following alternate form: $C(T) = 2LD - 2L\int_0^T \left[2Dq^D(\tau) - \frac{q^D(\tau)-1}{q(\tau)-1}\right] f_a(\tau)d\tau$.

Since $q(\tau)$ is a decreasing function of τ and $0 \le q(\tau) < 1$, it is easy to verify that $C(T)$ is a convex function of T. Therefore, if we let $g[q(T)] = -\frac{1}{2Lf_a(T)}\frac{dC(T)}{dT} = 2Dq^D(T) - \frac{q^D(T)-1}{q(T)-1}$, the minimum of $C(T)$ is achieved when $g[q(T)] = 0$. Therefore, the optimal value of $q(T)$ is the root in $[0,1)$ of a function of the form $g(x) = 2Dx^D - \frac{x^D-1}{x-1}$. Given any value of D, a numerical method such as bisection or the Newton's method can be used to find this root. Since $q(T) = 1 - F_r(T)$, once the optimal value of $q(T)$ is determined numerically, the optimal TTL value can be found by reversing the density function of the residual lifetime of a link.

The above illustrates an important property of the optimal TTL: it does not depend on $f_a(t)$. This property significantly reduces the computational requirement of the adaptive, real-time route-cache TTL optimization performed by individual nodes in an ad hoc network.

4 Simulation and Numerical Evaluation

4.1 Simulation Model and Output Analysis

A simulation model is developed to validate the analytical model. It represents the link establishments and breakages in an ad hoc network based on the network model described in Section 2. In particular, we present the simulation results for a 300-node network where the link up and down-times between any pair of nodes are exponentially distributed with mean values $\mu_u = 1$ and $\mu_d = 48.8$ (i.e., the average node degree is 6). Given a source node, the destination node is chosen randomly with uniform distribution among all other nodes in the network. For a chosen source and destination node

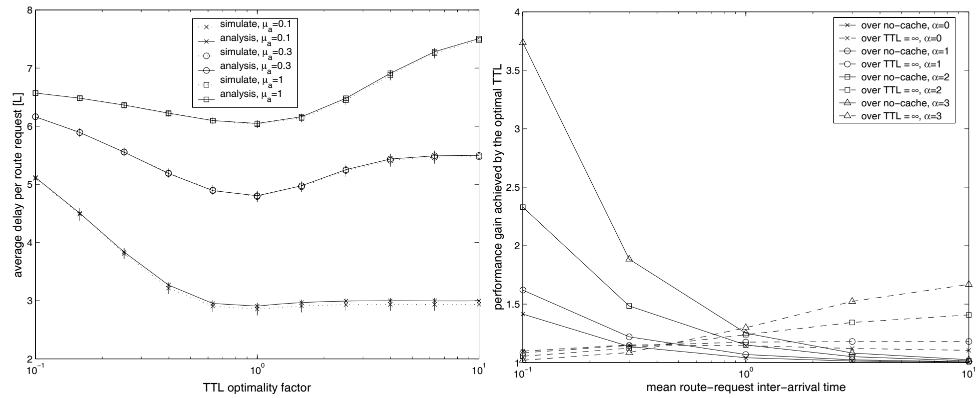

Fig. 1. (a) Expected routing delay (normalized to L) vs. the TTL optimality factor γ. The vertical lines represent the 99.95% confidence intervals. **(b)** Performance gain achieved by using the optimal route-cache TTL over TTL=0 and TTL=∞.

pair, the route-request inter-arrival time has distribution $f_a(t) = \frac{1}{\mu_a} e^{-\frac{t}{\mu_a}}$. We further define a TTL optimality factor γ, such that, when a new route is cached, its TTL is set to γT_{opt}, where T_{opt} is the optimal TTL value found as described in Section 3.2. The comparison between our analytical and simulation results is illustrated in Fig. 1(a).

Figure 1(a) validates both the analytical and the simulation models. In particular, the simulation results demonstrate that the minimal routing delay is indeed achieved at $\gamma = 1$, as expected from the analysis. The computed average delay per route request in some cases is 2% higher than the corresponding simulation outcome. This is due to the pessimistic assumption in the analytical model that once a link in a cached route fails, it does not become up again at the time of the next route request.

Figure 1(a) also suggests that the optimal TTL determination is the most important when the route-request inter-arrival time is moderate compared with the mean link-failure time. For systems with different parameter values, the results are similar to Fig. 1(a) and are omitted.

4.2 Performance Gain of the Optimal TTL

Using the proposed analytical framework, we can quantitatively study the advantage of optimizing the route-cache TTL. Due to the page limit, we are unable to show all results. In Fig. 1(b), we illustrate the performance gain of using the optimal TTL over the no route-cache system (TTL=0) and the never-expiring route-cache system (TTL=∞)[8] , for different values of μ_a[9] and various traffic locality. In describing the traffic locality, we have used a power law distribution as follows. Let π_D be the probability that a given

[8] We define the performance gain as the ratio between the expected delay of using a non-optimal TTL and the expected delay of using the optimal TTL.

[9] We have scaled time such that $\mu_u = 1$. Therefore, $1/\mu_a$ represents the relative frequency of the route requests to the frequency of topology variation. Also note that the analytical results are valid for any μ_d as long as $\mu_d >> \mu_u$.

route request is made to a destination of D hops away. If D is upper-bounded by D_{max}, the probability distribution function of D is defined as $\pi_D = \frac{D^{-\alpha}}{\sum_{i=1}^{D_{max}} i^{-\alpha}}$, where a larger value of α indicates a higher level of locality. In this example, $D_{max} = 20$.

Figure 1(b) demonstrates that the performance gain is a fast increasing function of α. As a point of reference, when $\alpha = 3$ and $\mu_a = 1$, using the optimal TTL can reduce the routing delay of either a non-caching system or a never-expiring caching system by approximately 25%. Therefore, route-cache optimization is especially important in the design of *scalable* on-demand routing protocols for large mobile ad hoc networks, where it has been proven that the traffic pattern must be localized[5].

5 Conclusions

We have presented analytical and numerical methods to determine the expected routing delay and the optimal route-cache TTL for on-demand routing. The analysis is based on a random-graph model of mobile ad hoc networks. Our analytical results agree very well with the simulation results.

Through the proposed analytical framework, one can study the routing delay of a network given various system parameters. The results of our analysis have shown that the optimal route-cache TTL does not depend on the route-request frequency or inter-arrival distribution. Furthermore, our numerical results have demonstrated that optimizing the route-cache TTL is the most effective when the traffic pattern is localized.

References

1. J. Broch, D. A. Maltz, D. B. Johnson, Y.-C. Hu, and J. Jetcheva, "A performance comparison of multi-hop wireless ad hoc network routing protocols," *ACM/IEEE MOBICOM*, 1998.
2. S. R. Das, C. E. Perkins, and E. M. Royer, "Performance comparison of two on-demand routing protocols for ad hoc networks," *IEEE INFOCOM*, 2000.
3. Y.-C. Hu and D. B. Johnson, "Caching strategies in on-demand routing protocols for wireless ad hoc networks," *ACM/IEEE MOBICOM*, 2000.
4. G. Holland and N. Vaidya, "Analysis of TCP performance over mobile ad hoc networks," *ACM/IEEE MOBICOM*, August, 1999.
5. P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Information Theory*, vol. 46, no. 2, March 2000.
6. J. J. Kelleher, "Tactical communications network modeling and reliability analysis: overview," JSLAI Report JC-2091-GT-F3, November 1991.
7. P. Johansson, T. Larsson, N. Hedman, B. Mielczarek, and M. Degermark, "Scenario-based performance analysis of routing protocols for mobile ad-hoc networks," *ACM/IEEE Mobicom*, August 1999.
8. J. Jubin and J. D. Tornow, " The DARPA packet radio network protocols," *Proceedings of IEEE (Special Issue on Packet Radio Networks)*, vol. 75, pp. 21-32, January 1987.
9. B. M. Leiner, D. L. Nielson, and F. A. Tobagi, "Issues in packet radio network design," *Proceedings of the IEEE*, vol. 75, pp. 6-20, January 1987.
10. D. A. Maltz, J. Broch, J. Jetcheva, and D. B. Johnson, "The effect of on-demand behavior in routing protocols for multihop wireless ad hoc networks," *IEEE JSAC - Special Issue on Wireless Ad Hoc Networks*, vol. 17, no. 8, pp. 1439-1453, August 1999.
11. C. E. Perkins, ed., *Ad Hoc Networking*, Addison-Wesley Longman, 2001.