Proceedings of the IEEE Workshop on Mobile Computing Systems and Applications, December 1994.

Routing in Ad Hoc Networks of Mobile Hosts

David B. Johnson

Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15213-3891

dbj@cs.cmu.edu

Abstract

An ad hoc *network is a collection of wireless mobile hosts forming a temporary network without the aid of any centralized administration or standard support services. In such an environment, it may be necessary for one mobile host to enlist the aid of others in forwarding a packet to its destination, due to the limited propagation range of each mobile host's wireless transmissions. Some previous attempts have been made to use conventional routing protocols for routing in ad hoc networks, treating each mobile host as a router. This position paper points out a number of problems with this design and suggests a new approach based on separate* route discovery *and* route maintenance *protocols.*

1. Introduction

Mobile hosts such as notebook computers, featuring powerful CPUs, large main memories, hundreds of megabytes of disk space, multimedia sound capabilities, and color displays, are now easily affordable and are becoming quite common in everyday business and personal life. At the same time, network connectivity options for use with mobile hosts have increased dramatically, including support for a growing number of wireless networking products based on radio and infrared.

With this type of mobile computing equipment, there is a natural desire and ability to share information between mobile users. Often, mobile users will meet under circumstances that are not explicitly planned for and in which no connection to a standard wide-area network such as the Internet is available. For example, employees may find themselves together in a meeting room; friends or business associates may run into each other in an airport terminal; or a collection of computer science researchers may gather in a hotel ballroom for a workshop or conference. Requiring each user to connect to a wide-area network in such situations, only to communicate with each other, may not be possible due to lack of facilities, or may be inconvenient or impractical due to the time or expense required for such connection.

These kinds of networks of mobile hosts have become known as *ad hoc networks*. An ad hoc network is a collection of wireless mobile hosts forming a temporary network without the aid of any centralized administration or standard support services regularly available on the wide-area network to which the hosts may normally be connected. For example, the Internet Engineering Task Force (IETF) is nearing completion of a proposed standard for mobile host internetworking in the Internet, involving the support services of a *home agent* on a mobile host's home network, and a *foreign agent* on the foreign network currently being visited by the mobile host [8]. In an ad hoc network, no home agent or foreign agent is available.

Some form of routing protocol is in general necessary in such an environment, since two hosts that may wish to exchange packets might not be able to communicate directly. For example, Figure 1 illustrates a simple ad hoc network of three mobile hosts using wireless network interfaces. Host *C* is not within the range of host *A*'s wireless transmitter (indicated by the circle around *A*) and host *A* is not within the range of host *C*'s wireless transmitter. If *A* and *C* wish to exchange packets, they may in this case enlist the services of host *B* to forward packets for them, since *B* is within the overlap between *A*'s range and *C*'s range. The maximum number of network hops needed to reach another mobile host in any practical ad hoc network is likely to be small, but may often be greater than one as shown here. The routing problem in a real ad hoc network may be even more complicated than this example suggests, due to the inherent nonuniform propagation characteristics of wireless transmissions, and since any or all of the hosts involved may move at any time.

This research was supported in part by the Wireless Initiative of the Information Networking Institute at Carnegie Mellon University.

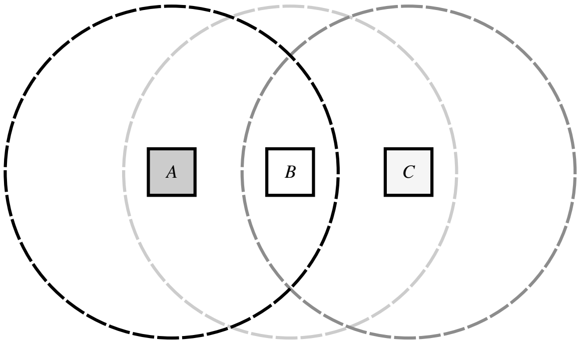


Figure 1 An ad hoc network of three wireless mobile hosts

This position paper considers the problem of routing in ad hoc networks and suggests a new approach to this routing based on separate *route discovery* and *route maintenance* protocols. Section 2 discusses the use of conventional routing protocols in ad hoc networks and describes a number of important problems with such protocols in this environment. Section 3 then outlines a new approach for ad hoc network routing, based on initially discovering a suitable route from a sending mobile host to some destination mobile host in an ad hoc network, and for then maintaining this route as the sender, the destination host, or other mobile hosts within the ad hoc network move or change status. Finally, Section 4 presents conclusions.

2. Conventional Routing Solutions

2.1. Description

A natural method for trying to provide routing in an ad hoc network is to simply treat each mobile host as a router and to run a conventional routing protocol between them [3, 9]. In effect, mobile host *B* in Figure 1 acts as a router between the "network" directly reachable by *A* and the "network" directly reachable by *C*. Host *A* transmits its packets for *C* to *B*, which then forwards them on to *C*. Conventional routing protocols are based on either *distance vector* or *link state* algorithms [10].

In *distance vector* routing, each router maintains a table giving the distance from itself to all possible destinations. Each router periodically broadcasts this information to each of its neighbor routers, and uses the values received from its neighbors to compute updated values for its own table. By comparing the distances received for each destination from each of its neighbors, a router can determine which of its neighbors is the correct "next hop" on the shortest path toward each destination. When presented a packet for forwarding to some destination, each router simply forwards the packet to the correct next hop router. By transmitting routing table updates more frequently such as when any information in the table changes, the algorithm converges more quickly to the correct path (for example, when a link comes up or goes down), but the overhead in CPU time and network bandwidth for transmitting routing updates increases. Examples of distance vector routing protocols include the routing protocol used in the DARPA Packet Radio Network [3]; the original routing protocol for the ARPANET [6]; RIP (used in parts of the Internet [1], in Novell's IPX [15], and in Xerox's XNS [16]); and RTMP (used in AppleTalk) [14].

In *link state* routing, each router maintains a complete picture of the topology of the entire network. Each router monitors the cost of the link to each of its neighbor routers, and periodically broadcasts an update of this information to all other routers in the network. Given this information of the cost of each link in the network, each router computes the shortest path to each possible destination. When presented a packet for forwarding to some destination, each router forwards the packet to the next hop router based on its current best path to that destination. Link state routing protocols converge much more quickly as conditions in the network change, but generally require more CPU time (to compute the complete shortest path to each possible destination) and more network bandwidth (to broadcast the routing update from each router to all other routers in the entire network) than distance vector algorithms. Examples of link state routing protocols include the "new" routing protocol that replaced the original protocol for the ARPANET [5], IS-IS (adopted by ISO as a standard routing protocol) [2], and OSPF (used in parts of the Internet) [7].

2.2. Problems

Although using either type of conventional routing protocol in an ad hoc network, treating each mobile host as a router, may often work, there are a number of problems with this approach:

- Transmission between two hosts over a wireless network does not necessarily work equally well in both directions. Even though host *A* in Figure 1 may receive a routing update from *B* indicating that *B* is closest to *C* and thus would be the first hop on *A*'s shortest path to *C*, host *A* may in fact be unable to transmit a packet back to *B*. Figure 1 represents the transmission range of all hosts as equal and uniform on all sides of the host, but radio and infrared propagation does not always work so nicely in reality. Thus, some routes determined by conventional routing protocols may not work in some environments.

- Many "links" between routers seen by the routing algorithm may be redundant. Rather than a single router (mobile host B) between A and C, there may be many mobile hosts within A's range, all perhaps equally good for forwarding packets to C. Wired networks, on the other hand, are usually explicitly configured to have only one (or a small number) of routers connecting any two networks. The redundant paths in the wireless environment unnecessarily increases the size of routing updates that must be sent over the network, and increases the CPU overhead required to process each update and to compute new routes.

- Periodically sending routing updates wastes network bandwidth. Often, nothing will change from one routing update to the next, but each router (mobile host) must continue to send periodic updates so that other routers will continue to consider routes through that router as valid. Routing updates from mobile hosts outside each other's transmission range will not interfere with each other, but where many mobile hosts are within transmission range of each other, their routing updates will consume each other's network bandwidth.

- Periodically sending routing updates wastes battery power. Most mobile hosts in an ad hoc network will be operating on battery power, and transmitting each packet expends a significant amount of battery power (transmitting a packet, in effect, launches a portion of the host's battery power into the air). Although receiving a packet generally requires less power than sending one, the need to receive these periodic routing updates effectively prevents a host from conserving its own battery power by putting itself into "sleep" or "standby" mode when not otherwise busy.

- Finally, conventional routing protocols are not designed for the type of dynamic topology changes that may be present in ad hoc networks. In conventional networks, links between routers occasionally go down or come up, and sometimes the cost of a link may change due to congestion, but routers do not generally move around dynamically, shifting major portions of the network topology back and forth. Mobile hosts, though, may be characterized by such dynamic movement, because they are, after all, mobile. Convergence to new, stable routes after such dynamic changes in topology may be quite slow, particularly with distance vector algorithms. The speed of convergence may be improved by sending routing updates more frequently, but such a shift only wastes more bandwidth and battery power when topology changes are less dramatic.

Some of these problems can be addressed by improvements to the routing protocol [9], but many of the problems still remain. This position paper, instead, takes the view that a new approach to routing in ad hoc networks is needed.

3. Route Discovery and Maintenance

The problem of routing can be divided into the two areas of *route discovery* and *route maintenance*. In order for one host to communicate with another, it must initially discover a suitable route to use in sending packets to that destination. As long as conditions remain unchanged, this route should then continue to work for as long as it is needed. However, as the status of different links or routers used in this route change, changes in the route may be necessary, or a new route may need to be discovered.

3.1. Route Discovery

A very simple and efficient method of route discovery suitable for use directly in some ad hoc networks is the Internet's Address Resolution Protocol (ARP) [11]. ARP is designed for dynamically translating a host's network protocol address (such as an IP address [12]) to its MAC-level address (such as an Ethernet address). A host attempting to translate another host's address broadcasts a query packet onto its local network, which is answered by the target host giving its MAC address; other hosts on the local network receiving the query do not reply. The returned MAC address is then cached by the host for use in sending future packets to this destination.

In an ad hoc network, if the source and target mobile hosts are both within transmission range of each other, a simple ARP query is all that is needed to find a "route" to the target host; the returned MAC address may be used directly to transmit packets to that host. In this case, no periodic routing updates are needed, providing substantial savings in network bandwidth and battery power requirements for all involved. What is needed to make this approach a general solution to route discovery in ad hoc networks is a technique for extending this to the case in which the source and target may not be within range of each other, while still preserving the simplicity and efficiency of the protocol as much as possible in the case in which they are.

One possible solution is to send a request packet (similar to ARP) but to propagate the request using some form of flooding, in order to reach other mobile hosts beyond the sender's transmission range. As the request propagates, each host adds its own address to a route being recorded in the packet, before broadcasting the request on to its neighbors (any host within range of its wireless broadcast transmission). When receiving a request, if a host finds its own address already recorded in the route, however, it discards that copy of the request and does not propagate that copy further. This scheme is similar to that used for finding source routes in source routing bridges in IEEE 802 LANs. The Tucson Amateur Packet Radio (TAPR) work also used source routing in wireless networks, but the routes were staticly configured by the user [4].

Since many mobile hosts may be within transmission range of each other, though, there may be many duplicate copies of each request propagated. To largely eliminate these duplicates, each request should contain a unique *request id* from the original sender; each host keeps a cache giving the request id and sender address of recently forwarded requests, and discards a request rather than propagating it if it has already propagated an earlier copy of the same request id. Thus, each host will only propagate the first copy of each request that it receives. This will usually be the copy that came to it along the shortest path from the original sender (since it arrived first), and thus is most useful in finding the shortest path to the final target. This scheme could easily be extended, though, to include the length of the path in the request id cache and to propagate a later copy of the same request if it somehow arrived over a shorter path than the earlier copy.

As mentioned previously, although more than one network hop may be needed to reach another mobile host in an ad hoc network, the maximum number of hops needed it is likely to be small. The number of duplicate requests propagated can thus be further reduced by limiting the maximum number of hops over which any route discovery packet can be propagated. When processing a received route discovery request, a mobile host should discard the request rather than forwarding it if it is not the target of the request and if the route recorded in the packet has already reached the maximum length.

When the query packet ultimately reaches the target host, the complete route from the original sender to this host will have been recorded in the packet. In order to be of use to the original sender, though, the route must then be returned to the sender. The target host may attempt to reverse the recorded route to reach the original sender, or may use the same route discovery procedure to find a route back to the original sender; the route from the original sender to this target should be returned to the sender in the new query packet used for its own route discovery.

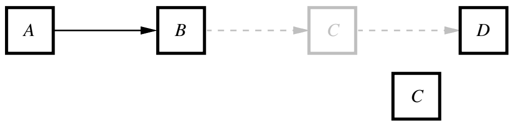
This route discovery exchange between the two end mobile hosts could optionally be piggybacked on the first data packets sent between them. For example, when opening a TCP connection, separate packets are usually used to exchange SYN and ACK control bits between the two end hosts of the connection [13]; the route discovery information could easily be carried in these same packets. If the end mobile hosts use a flow setup protocol to reserve resources or bandwidth for a specified quality of service between them, this route discovery exchange could likewise be integrated with the flow setup exchange. Such a flow setup could also be used to establish state for the route in each of the mobile hosts along the path, so that the entire route need not be included on each packet sent along the path.

Mobile hosts should cache routes discovered in this way for use in sending future packets to that same destination. The overhead of the protocol can be further reduced by making more extensive use of such caching. If a mobile host has cached a route listing some number of hops to a destination mobile host, then the shortest route to each of the hops listed on that route is naturally a prefix of that route. In this case, no new route discovery is needed for this mobile host to communicate with any of the other mobile hosts named as hops on any of the routes currently in its cache. Also, as a mobile host forwards packets, it will be able to observe many other routes to other mobile hosts, since each packet contains a route. By examing the routes on packets that it forwards, a mobile host may be able to cache routes to new destinations or to obtain updated information to destinations already in its cache. Furthermore, since transmissions in a wireless network are necessarily broadcast transmissions, a mobile host may be able to learn new routing information from the route contained in any packet that it can receive, even if the packet is not explicitly addressed to this host.

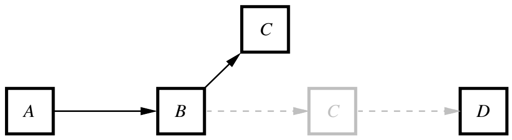
3.2. Route Maintenance

Conventional routing protocols integrate route discovery with route maintenance by continuously sending the normal periodic routing updates. If the status of a link or router changes, the periodic updates will eventually reflect the changes to all other routers, presumably resulting in the computation of new routes. With the separate route discovery approach as outlined in Section 3.1, a link or router going down would instead cause the route to mysteriously stop working with no feedback to the sender. The role of the route maintenance protocol is to provide this feedback, and to allow the route to be modified or a new route to be discovered in this case.

In an ad hoc network, a route may also stop working if one or more of the mobile hosts along the route simply move. For example, Figure 2 illustrates two possible scenarios in which the movement of a mobile host causes an existing route to stop working. Assume that mobile host A has been sending packets to mobile host D using a route through mobile hosts B and C. Figure 2 (a) shows the case in which C has moved out of range of B, breaking the route on to D. Figure 2 (b) shows a different scenario in which C has moved such that it is now out of range of its next hop



(a) Host *C* has moved, breaking the route at *B*



(b) Host *C* has moved, leading the route away from *D*

Figure 2 Route changes due to host movement

on to *D*; in this case, *C* after moving is still within range of *B*, but it has led the route away from *D*.

In many wireless networks, route maintenance can be provided with very little overhead. Since wireless networks are inherently less reliable than wired networks, many wireless networks utilize a hop-by-hop acknowledgement at the data link level in order to provide early detection and retransmission of lost or corrupted packets. In these networks, the problem of route maintenance is quite simple, since at each hop, the sender can determine if that hop of the route is still working. If the data link level reports a transmission problem for which it cannot recover (for example, because the maximum number of retransmissions it is willing to attempt has been exceeded), all that is needed is to report this error back to the original sender to cause that host to reinvoke the route discovery procedure to find a new route. It may also be possible for the intermediate host experiencing the error to instead use the route discovery procedure itself to extend the existing route (up to itself) on to the correct target.

If the wireless network does not support such lower-level acknowledgements, an equivalent acknowledgement signal may be available in many environments. After sending a packet to the next hop mobile host, the sender may be able to hear that host transmitting the packet again, on its way further along the path. For example, in Figure 1, host *A* may be able to hear *B*'s transmission of the packet on to *C*. In addition, existing packets coming in the reverse direction along the same link (such as transport or application level replies or acknowledgements from the original destination) could also be used as an acknowledgement that the route (or that hop of the route) is still working. As a last resort, a bit in the packet header could be included to allow a host transmitting a packet to request an explicit acknowledgement from the next-hop receiver. If no other acknowledgement signal has been received in some time from the next hop on some route, the host could use this bit to inexpensively probe the status of this hop on the route.

4. Conclusion

This position paper has proposed a new method for routing in ad hoc networks based on separate *route discovery* and *route maintenance* protocols. The performance of this approach depends on a number of factors such as how often mobile hosts in such an environment attempt to communicate with other mobile host for which they have no cached route (when route discovery is needed) and how often mobile hosts move enough for existing routes to stop working (when route maintenance is needed). Perhaps the most important factor in the protocol affecting performance is how well the propagation of redundant copies of a route discovery request by any mobile host can be reduced through methods such as discarding a request if this mobile host is alreay listed on the route in the request or if this mobile host has recently processed a request with this same request id, through limiting the maximum length of a route, and through aggressive route caching and full use of information in the cache.

A number of options remain to be resolved in the design of the protocols described here. We are currently building a packet-level simulation with which to evaluate these options and to study the behavior and performance of the system. We are also exploring additional areas related to ad hoc networking, such as the routing between an ad hoc network and a wide-area network such as the Internet: if one or more of the mobile hosts in an ad hoc network are also connected to the Internet [8], it is possible for other mobile hosts in the ad hoc network to communicate with Internet hosts, but additional routing support is needed for them to learn an appropriate route to these Internet hosts and for Internet hosts to be able to route packets into the ad hoc network.

References

[1] C. Hedrick. Routing Information Protocol. Internet Request For Comments RFC 1058, June 1988.

[2] International Standards Organization. Intermediate system to intermediate system intra-domain routing exchange protocol for use in conjunction with the protocol for providing the connectionless-mode network service (ISO 8473). ISO DP 10589, February 1990.

[3] John Jubin and Janet D. Tornow. The DARPA packet radio network protocols. *Proceedings of the IEEE*, 75(1):21–32, January 1987.

[4] Philip R. Karn, Harold E. Price, and Robert J. Diersing. Packet radio in the amateur service. *IEEE Journal on Selected Areas in Communications*, SAC-3(3):431–439, May 1985.

[5] John M. McQuillan, Ira Richer, and Eric C. Rosen. The new routing algorithm for the ARPANET. *IEEE Transactions on Communications*, COM-28(5):711–719, May 1980.

[6] John M. McQuillan and David C. Walden. The ARPA network design decisions. *Computer Networks*, 1(5):243–289, August 1977.

[7] J. Moy. OSPF version 2. Internet Request For Comments RFC 1247, July 1991.

[8] Charles Perkins, editor. IP mobility support. Internet Draft, IETF Mobile IP Working Group, 21 October 1994. Work in progress.

[9] Charles E. Perkins and Pravin Bhagwat. Highly Dynamic Destination-Sequenced Distance-Vector Routing (DSDV) for Mobile Computers. In *Proceedings of the SIGCOMM '94 Conference on Communications Architectures, Protocols and Applications*, pages 234–244, August 1994.

[10] Radia Perlman. *Interconnections: Bridges and Routers*. Addison-Wesley, Reading, Massachusetts, 1992.

[11] David C. Plummer. An Ethernet address resolution protocol: Or converting network protocol addresses to 48.bit Ethernet addresses for transmission on Ethernet hardware. Internet Request For Comments RFC 826, November 1982.

[12] J. B. Postel, editor. Internet Protocol. Internet Request For Comments RFC 791, September 1981.

[13] J. B. Postel, editor. Transmission Control Protocol. Internet Request For Comments RFC 793, September 1981.

[14] Gursharan S. Sidhu, Richard F. Andrews, and Alan B. Oppenheimer. *Inside AppleTalk*. Addison-Wesley, Reading, Massachusetts, 1990.

[15] Paul Turner. NetWare communications processes. *NetWare Application Notes*, Novell Research, pages 25–81, September 1990.

[16] Xerox Corporation. Internet transport protocols. Xerox System Integration Standard 028112, December 1981.