A Review and Comparison of Efficient Flooding Schemes for On-demand Routing Protocols on Mobile Ad hoc Networks (MANETs)

Bassam M.S. Waheed, Ayyoub A. Moghanjoughi, Sabira Khatun, Borhanuddin B. Ali

Wireless Networks Routing Research Group, Department of Computer & Communications Systems Engineering, Faculty of Engineering, University Putra Malaysia

Abstract: Since the basic components of ad hoc wireless networks are mostly battery-operated portable devices, power conservation is one of the central issues of such networks. Power-conservative designs for ad hoc networks pose many challenges due to the lack of central coordination facilities. Existing on-demand routing protocols perform route discovery by flooding the network with a query message requesting a route to the destination. Flooding is used because of its simplicity and greater success in finding the best route between the source and destination available at that time of route discovery. However, as flooding involves querying all reachable network nodes, frequent flooding can rapidly deplete the energy reserved at each node. In addition to consuming significant portions of the available network bandwidth. Further, as the number of communicating nodes increases, more congestion, contention, and collisions can be expected. This paper reviews and compares approaches for optimizing bandwidth efficiency of route discovery, where several efficient flooding schemes have been presented based on different techniques to solve the problems related with the traditional blind flooding.

Key words: Efficient flooding, broadcast storm problem, Mobile Ad-hoc networks

INTRODUCTION

Ad hoc wireless networks have received widespread attention in recent years. Unlike wired networks or cellular networks, ad hoc wireless networks have no fixed networking infrastructure. As indicated in figure 1, an ad hoc wireless network consists of multiple nodes that maintain network connectivity through wireless communications (Kmar et al., 2008). This connectivity is enabled via radio transmissions generated by a set of cooperating nodes. Ad hoc wireless networks have applications in areas where it is not economically practical or physically possible to provide a wired networking infrastructure. For example, in a battlefield situation it is not possible to install a conventional network in hostile territory. In such a situation an ad hoc wireless network offers a promising solution. While in the commercial sector mobile ad hoc networks could be used in rescue operations in remote areas, or when local coverage must be deployed quickly at a remote construction site, ad hoc networking could also serve as wireless public access in urban areas, providing quick deployment and extended coverage (Wu and Tseng, 2007).

Fig. 1: An ad hoc wireless network
Ad hoc networks are characterized by scarce resources (e.g., bandwidth, battery power etc.), lack of any established backbone infrastructure, and a dynamic topology. The scarcity of resources, lack of an infrastructure for performing routing, and the constantly changing topology render conventional routing protocols inappropriate for the target environment (Hyas et al., 2003; Li et al., 2003). A challenging but critical task that researchers have tried to address over the past few years was to develop suitable and efficient routing protocols that suit with the characteristics of ad hoc networks.

The traditional on-demand routing protocols like ad hoc on-demand distance vector (AODV), temporally ordered routing algorithm (TORA), dynamic source routing (DSR) etc. perform the flooding using broadcast relays which use local broadcast and this could lead to performance degradation due to the following reasons:

- High protocol overhead: The number of messages is of the order of the number of nodes in the network (Sivakumar et al., 2003).
- Unreliability of broadcasts: researchers in (Johansson et al., 1999; Bharghavan et al., 1994) have demonstrated the inefficacy of local broadcasts, because of their unreliability, to convey information to all nodes in wireless environments. Johansson et al., (1999). showed that, for moderately connected graphs, the fraction of nodes in a network that receives a flood goes down to around 80% in the worst case. This figure can be expected to go down even more in sparsely connected topologies. As a result, route requests might not reach the destination in a single flooding attempt resulting in multiple floods.
- Interference due to broadcasts: Since local broadcasts are not transmitted with the full (RTS–CTS) handshake they can experience collisions with the other packets (including other broadcasts) in the network resulting in an overall reduction in the network utilization (Sivakumar et al., 2005).

Excessive bandwidth overhead produced by query flooding during route discovery has inspired protocol designers to find new ways to reduce the frequency and spread of route discovery. The rest of the paper is organized as follows. Section 2 presents the available techniques in AODV and DSR to reduce the overhead of route discovery process. Section 3 presents the recent proposed techniques by individual researchers. Section 4 presents the comparisons of the presented protocols, and finally conclusion of paper comes in Section 5.

**Overview of AODV and DSR techniques:**

AODV has several attractive features that enable reliable routing over mobile ad hoc networks (MANETs). These features can summarized as follows

- Supressing duplicate route requests: due to the broadcast nature of transmissions, a node’s transmission may be heard by several near by nodes, it is possible that a node may receive multiple copies of the same route request (RREQ) message rebroadcast by its neighbors. For reasons of bandwidth efficiency, processing and forwarding of duplicate RREQs are normally not desirable. Therefore intermediate nodes process and forward only the first copy of the RREQ and discard other copies of the same RREQ that arrive later. To ascertain whether a RREQ received has been previously processed, nodes keep track of each RREQ’s source IP address and RREQ ID, which is a sequence number generated by the source to uniquely identify a particular RREQ (Perkins and Royer, 1999).
- Intermediate nodes replying to route requests: during the process of route discovery, any intermediate node receiving the RREQ and having a route to the destination may send a route reply (RREP) message to the source, instead of continuing to forward the RREQ to the destination. This helps to quench or stops the spread of the query flooding at the intermediate node also it may reduce the route discovery overhead and latency (Perkins and Royer, 1999).
- Expanding Ring Search: Its technique used by AODV to delay the triggering of flooding, that it searches in increasingly larger neighborhoods, by sending out successive RREQs, each with a larger time-to-live (TTL) values that limits how far a RREQ can traverse from the source. In this technique, several attempts at route discovery may be conducted. The source initially conducts a route discovery that searches only the region within some limited hops from itself. When a route to the destination cannot be found after some timeout, the source attempts another route discovery, but with a greater search scope than the preceding attempt by increasing the TTL of the RREQ. Therefore this technique is effective only if the destination can be found within the initial attempts (Basagni et al., 2004).
- Local repair: AODV includes an additional provision to allow an intermediate forwarding node to initiate a route rediscovery to the destination upon detection of a link failure. This process is typically known as “local repair,” in which the intermediate node upstream the link failure sends out a RREQ, the TTL of
which is set to the remaining hop-distance to the destination, added with an increment value. Data packets
are buffered at this node during the route rediscovery and sent as soon as the route is repaired. If after
some timeout, no route to the destination could be found, the buffered data packets would be dropped,
and a route error (RERR) message would be sent to the source, which then attempts a source-initiated
route discovery. Local repair is useful, in particular in large networks where routes could be long and thus
more susceptible to link failures (Basagni et al., 2004).

DSR protocol has the following set of features that can be demonstrated in the following points.

- Caching Overheard Routing Information: one distinct feature of DSR is the caching of overheard routing
  information available from the headers of passing-by packets. Nodes with network interface operating in
  promiscuous mode are able to eavesdrop on packets that may not be addressed to them. Routing
  information learned from these packets can be stored for future use by the nodes themselves. Hence, nodes
  may perform less route discovery since the desired routes may be already available in their caches.
  However, the promiscuous learning of routes may incur high processing overhead, as well as requiring
  large memory storage (Johnson and Maltz, 1996).
- Preventing Route Reply Storm: another feature of DSR is the deliberate introduction of delay to an
  intermediate node that is sending a RREP message. Due to the nature of broadcast transmission, many
  nodes around the broadcasting node may receive the RREQ and send RREPs simultaneously. This may
  result in what is dubbed a RREP “storm,” in which a great number of node s attempt to send RREPs from
  their caches simultaneously, causing local congestion and excessive packet collisions. Having some nodes
  delay sending their RREPs may mitigate this problem. The delay time is specified to be: $d = H(h – 1 + r)$,
  where $h$ is the number of hops of the returned route, $r$ is a random number between 0 and 1, and
  $H$ is a small constant delay to be introduced per hop. Notice that shorter the route, the earlier the node
giving this route would reply (Basagni et al., 2004).
- Non-propagating route requests: Similar to the expanding ring search in AODV, DSR has a variant of such
  an approach, but it limits the number of route discoveries to two attempts. In the initial attempt, the source
  sends a “nonpropagating” RREQ with a hop limit of 1 (i.e., TTL = 1) to look for either the destination
  or some node with a route to the destination within its immediate neighborhood. If a route cannot be
  found, the source node sends a “propagating” RREQ with no hop limit, which essentially floods the
  network. (Sivakumar et al., 2005).

Efficient Flooding Schemes:
Flooding is one of the most important operations in mobile ad hoc networks. It is used to propagate the
control messages in most ad hoc routing protocols. In flooding, a node transmits a message to all of its
neighbors. The neighbors in turn transmit to their neighbors and so on until the message has been propagated
to the entire network (Johnson and Maltz, 1996). However blind flooding can become very inefficient because
flooding increases link overhead and wireless medium congestion. In a large network, with heavy load,
this extra overhead can have severe impact on performance and should be eliminated, where in spite of
suppressing the duplicated route requests in both AODV and DSR, some of rebroadcasts may be redundant.
As shown in figure 2 that is if node A broadcasts a new RREQ to Nodes B and C, which in turn rebroadcast
to Node D. Hence, node D receives two copies of the same RREQ, one of which is redundant.

Fig. 2: Broadcast storm
Moreover, if Nodes B and C are close to each other and both transmit at the same time, channel contention could occur. Further, RTS/CTS exchange is not used in broadcast transmission. If the underlying MAC does not provide collision detection capability, packet collisions could be damaging. The resulting redundancy, contention and collisions constitute what is called the “broadcast storm” problem (Ni et al., 1999). In this paper we present and compare between different efficient flooding schemes that can be categorized according to their strategy in solving the broadcast storm problem.

**Fig. 3:** Categorization of efficient flooding schemes for routing protocol in wireless ad hoc networks

**One-hop Neighbor Information:**

This scheme focuses on implementing efficient flooding based on 1-hop neighbor information only to make the routing protocol easy to implement and light weight in overhead. Authors in (liu et al., 2007) achieve local optimality in two senses, first: the number of forwarding nodes is minimal, second: the time complexity is the lowest. The time complexity for computing the forwarding nodes in each step is \( O(n \log n) \), which is the lower bound \( (n \) is the number of neighbors of a node), the proposed scheme assumes all nodes in the network have the same transmission range \( R \) thus the network can be represented as a unit disk graph \( G(V,E) \). Each node \( v \) in \( V \) has a unique ID, denoted by \( id(v) \). Let \( N(v) \) denotes the set of neighbor nodes of \( v \) and \( F(v) \) denotes the forwarding set of node \( v \). That is, nodes in \( N(v) \) are within the transmission range of \( v \) and can receive signals transmitted by \( v \). Node \( v \) needs to know the information of its direct neighbors, including their IDs and their geographic locations. The 1-hop neighbor information can be easily obtained from the HELLO messages periodically broadcast by each node. The basic idea of this flooding scheme is that when a node (source node) has a message to be flooded out, it computes a subset of its neighbors as forwarding nodes and attaches the list of the forwarding nodes to the message. Then, it transmits (broadcasts) the message out. After that, every node in the network does the same as follows: Upon receiving a flooding message, if the message has been received before, it is discarded; otherwise, the message is delivered to the application layer and the receiver checks if itself is in the forwarding list. If yes, it computes the next hop forwarding nodes among its neighbors and transmits the message out in the same way as the source. The message will eventually reach all the nodes. When a node \( s \) receives a flooding message for the first time and it appears in the forwarding list attached to the message (\( s \) could be the original source of the message), \( s \) is designated as a forwarding node and it computes the next hop forwarding nodes from its neighbors. Since \( s \) only has 1-hop neighbor information, it does not know who the 2-hop neighbors are. To achieve 100 percent deliverability, the forwarding set \( F(s) \) must cover the entire neighbor’s area of \( s \) according to the following definition:
Minimize $F(u)$ such that $\bigcup_{v \in F(u)} d(v) = \bigcup_{u \in N(s)} d(u)$.

Figure 4. Shows that node $s$ has three neighbors: $u$, $v$, and $w$. Since $d(u) \cup d(v) \cup d(s)$ makes up the neighbor’s area of $s$, it is enough to cover all the of $s$’s 2-hop neighbors if only $u$ and $v$ forward the message.

Fig. 4: Neighbor’s area of node $s$

To minimize $F(s)$, every node in $F(s)$ must contribute to the neighbor’s boundary of $s$; otherwise, this node can be removed from $F(s)$ without affecting the coverage area of $F(s)$. Therefore, computing the minimal $F(s)$ is to find a subset of $N(s)$ such that every node in the subset contributes to the neighbor’s boundary of $s$. The $F(s)$ computed above is only locally optimal based on the 1-hop information of $s$. When a node $u$ receives the flooding message from $s$ (we call $s$ the parent of $u$) and $u$ is a forwarding node nominated by $s$ (i.e., $u \in F(s)$), the computing of $F(u)$ can be further optimized based on the information of $F(s)$, which is attached to the flooding message from $s$. This is because some nodes in $F(u)$ may already be covered by node $s$ or node-set $F(s)$ and, thus, $F(u)$ could be further reduced by removing out those nodes.

Two-hop Neighbor Information:

This approach is based on the concept of “2-hop backward information” which is the information of predecessors in order to reduce the number of the retransmitting nodes (RNs) in the network (Le and Choo, 2008). The 2-hop backward information (2HBI) has the following key advantages:

1. 2HBI employs the concept of “2-hop backward information” to minimize the set of RNs at every flooding hop.
2. 2HBI does not require any extra communication overhead other than the exchanging of 1-hop HELLO message.
3. 2HBI has the small time complexity, $O(\text{log}n)$, where $n$ is the number of neighbors.
4. In 2HBI, the sender node decides which node will retransmit the flooding message (sender-based); further, the receiver node can abandon its retransmitting duty (receiver-based).

In 2HBI if the sender node of one node is termed “1-hop backward node,” the sender of the 1-hop backward node is termed the “2-hop backward node”. As shown in Fig. 5, the flooding message propagates from the source node 0 to node 3, then from node 3 to node 4, and so on. Therefore, in the case of node 4, its 1-hop and 2-hop backward nodes are nodes 3 and 0, respectively. The 2-hop backward information is defined as the combination of the forwarding set information of the 1-and 2-hop backward nodes and the 1-hop neighbor information. The forwarding set information includes the node IDs and the geographic location information of all the nodes in that set. The 2-hop backward information is acquired with the aid of the forwarding set information propagation mechanism, in which every retransmitting node (RN) attaches its forwarding set information and that of its sender node to the flooding message before broadcasting it. Upon receiving the message, the next-hop RN node combines this extra information with its 1-hop neighbor information, and this constitutes the entire 2-hop backward information. As shown in Figure 5, node 4 obtains the forwarding set information of nodes 0 and 3 from the flooding message. Then, this information is combined
with the 1-hop neighbor information; therefore, node 4 contains the entire 2-hop backward information. In this manner, every node has to maintain only the 1-hop neighbor information since any extra information is obtained from the incoming flooding message.

Fig. 5: Flooding process in 2HBI

In the 2HBI the selection of the forwarding set is achieved by considering the neighbor list as the input, every node that is assigned as the RN computes the boundary of its neighbors’ coverage area as follows. First, it initiates the set that includes its boundary and that of its neighbors. Then, it merges all the boundaries in this set in pairs until a unique boundary is obtained. We term this boundary as the “boundary of the neighbors’ coverage area.” Hence, only the neighbors that contribute to this boundary are included in the initial forwarding set. For example, the boundary of the neighbors’ coverage area of node 0 is indicated in Figure 5 by the thick line. As shown in this figure, only neighbor nodes 2 and 3 contribute to this boundary; therefore, the initial forwarding set of node 0 includes nodes 2 and 3. As in the merge sort algorithm, the time complexity of the forwarding set selection algorithm is $O(n \log n)$, where $n$ is the number of neighbors.

Connected Dominating Set:

The two methods presented in the previous sections include a forwarding set of neighbors as part of the message. They therefore have message overhead, and the set of retransmitting nodes depends on the source node. The approach presented in this section does not require inclusion of the forwarding set in the message, and has a fixed set of retransmitting nodes, regardless of source choice. Its maintenance does not require more communication overhead, and it offers competitive performance (Wu, 2002). Efficient routing among a set of mobile hosts is one of the most important functions in ad-hoc wireless networks. Routing based on a connected dominating set is a promising approach, where the searching space for a route is reduced to nodes in the set. A set is dominating if all the nodes in the system are either in the set or neighbors of nodes in the set. The efficiency of dominating-based broadcasting or routing mainly depend on the overhead in constructing the dominating set and the size of dominating set (Li et al., 2006).

Authors in (Wu and Li, 1999) consider the wireless ad hoc networks as unweighted graph $G = (V, E)$, where $V$ represents a set of wireless mobile hosts and $E$ represents a set of edges, also it assumes all mobile hosts have same transmission range, in other words an edge between host pairs $\{v, u\}$ indicates that both hosts $v$ and $u$ are within their wireless transmitter ranges. Thus the corresponding graph will be an undirected graph. Routing in ad hoc wireless networks requires fast convergence and low communication overhead. Routing information has to be localized to adapt quickly to network topological changes. Connected-dominating-set-based routing can be a solution to this kind of network environment.

The main advantage of connected-dominating-set-based routing is that it centralizes the whole network into small connected dominating set subnetwork, which means only gateway hosts keep routing information, so that as long as network topological changes do not affect this subnetwork there is no need to recalculate routing tables. Many efforts have been made to find connected dominating set with a minimum size (MCDS), but unfortunately this is NP-complete for most of the graphs. In (Wu and Li, 1999) the authors suggest a distributed approximation algorithm that can quickly determine a connected dominating set in a given connected graph, which represents an ad hoc wireless network. They also discuss the ways to update and recalculate the dominating set when the underlying graph changes with the movement of mobile hosts. Some of the desirable features for that should be available in the formation process of dominating set are:
1. It requires only local information and constant number of iterative rounds of message exchanges among neighboring hosts.
2. The resultant dominating set should be connected and close to minimum.
3. The resultant dominating set should include all intermediate nodes of any shortest path. In this case, an all-pair shortest paths algorithm only needs to be applied to the subnetwork containing the dominating set.

The technique that was used is called marking process that it marks every vertex in a given connected and unweighted graph \( G = (V, E) \), \( m(v) \) is a marker for vertex \( v \), \( V \), which is either T (marked) or F (unmarked), it is assumed that all vertices are unmarked initially. \( N(v) = \{ u | v, u \} \) represents the open neighbor set of vertex \( v \), where \( v \in N(v) \), the marking process consists of the following three steps:

1. Initially assign marker \( F \) to every node \( v \) in \( V \).
2. Every \( v \) exchanges its open neighbor set \( N(v) \) with all its neighbors.
3. Every \( v \) assigns its marker \( m(v) \) to \( T \) if there exist two unconnected neighbors.

As shown in figure in figure 6, \( N(u) = \{ v, y \} \), \( N(v) = \{ u, w, y \} \), \( N(w) = \{ v, x \} \), \( N(y) = \{ u, v \} \), and \( N(z) = \{ w \} \). After Step 2 of the marking process, vertex \( u \) has \( N(v) \) and \( N(w) \), \( v \) has \( N(u) \), \( N(w) \), and \( w \) has \( N(v) \) and \( N(z) \). \( v \) has \( N(u) \) and \( N(v) \), and \( z \) has \( N(w) \). Based on Step 3, only vertices \( v \) and \( w \) are marked \( T \).

![Fig. 6: A sample for dominating nodes in ad hoc wireless network](image)

If \( V' \) is the set of vertices that are marked true in \( V \), that is mean \( V' = \{ vv V, m(v) = T \} \). The induced graph \( G' \) is the subgraph of \( G \) induced by \( V' \). Since the problem of determining a minimum connected dominating set of a given connected graph is NP-complete, the connected dominating set derived from the marking process is normally non minimum. The authors in (Wu and Li, 1999) proposed two rules based on node ID to reduce the size of a connected dominating set generated by the marking process. First of all, a distinct ID, \( id(v) \), is assigned to each vertex \( v \) in \( G \). \( V' = \{ v | id(v) \} \) is the closed neighbor set of \( v \), as oppose to the open one \( N(v) \).

Rule 1: consider two vertices \( v \) and \( u \) in \( G' \). If \( N[v] \) \( N[u] \) in \( G \) and \( id(v) < id(u) \) the marker of \( v \) is changed to \( F \) if vertex \( v \) is marked, that’s mean \( G' \) is changed to \( G' - \{ v \} \). The above rule states that when the closed neighbor set of \( v \) is covered by the one of \( u \), vertex \( v \) can be removed from \( G \) if the ID of \( v \) is smaller than the one of \( u \).

Rule 2: Assume that \( u \) and \( w \) are two marked neighbors of marked vertex \( v \) in \( G' \). if \( N(v) \cap N(u) \cap N(w) \) in \( G \) and \( id(v) = min \{ id(v), id(u), id(w) \} \), then the marker of \( v \) is changed to \( F \). The above rule indicates that when the open neighbor set of \( v \) is covered by the open neighbor sets of two of its marked neighbors, \( u \) and \( w \), if \( v \) has the minimum ID of the three, it can be removed from \( G' \).

Authors in (Wu and Stojmenovic, 2001) propose two additional rules to minimize the size of connected dominating set based on the energy levels (ELs) of the nodes within the set.

Rule 1a (energy level): the first rule indicates that when the closed neighbor set of \( v \), is covered by the one of \( u \), vertex \( v \) can be removed from \( G' \) if the EL of \( v \) is smaller than the one of \( u \). ID is used to break a tie when \( el(v) = el(u) \).

Rule 2a (energy level): indicates that if \( v \) is covered by \( u \) and \( w \), where both \( u \) and \( w \) are marked neighbors of the marked vertex \( v \) in \( G \), so if neither \( u \) nor \( w \) are covered by the other two among \( u, v, \) and \( w \), node \( v \) can be removed from \( G' \), also if nodes \( v, u \) are covered by \( u \) and \( w \), \( v \) and \( w \), respectively, but \( w \) is not
covered by $u$ and $v$, node $v$ can be removed from $G'$ if the EL of $v$ is smaller than the one of $u$ or the ID of $v$ is smaller than the one of $u$ when their node degrees are the same. Finally when each of $u$, $v$ and $w$ is covered by the other two among $u$, $v$ and $w$, node $v$ can be removed from $G'$ if one of the following conditions holds:

$v$ has the minimum EL among $u$, $v$ and $w$, the EL of $v$ is the same as the EL of $u$ but it is smaller than the one of $w$ and the ID of $v$ is smaller than the one of $u$, or the EL's of $u$, $v$ and $w$ are the same and $v$ has the minimum ID among $u$, $v$ and $w$.

**Location Awareness:**

Location aided routing protocol (Ko and Vaidya, 1998) is one of location-aware techniques that propose using location information obtained from global positioning system (GPS) to confine the route search to a region where the destination is likely to be found. Figure 7. illustrates the concepts of LAR. By knowing the physical location $L$ and average speed $v$ of the destination at time $t_0$, the source defines at time $t_1$ a circular region of radius $v(t_1 - t_0)$ called “expected zone.” This is the region in which the destination may be found. In addition, the source defines the smallest rectangle that includes the expected zone and itself as the “request zone,” in which only nodes that reside in this zone can forward the RREQ. The source attaches this information on request zone to the RREQ.

The authors in (Wang, 2005) propose an improved location-aided routing (ILAR) scheme to improve the efficiency of location-aided routing (LAR) scheme by using the global positioning system (GPS). In this scheme, they suggest a baseline, which is the line between the source node and the destination node, for route discovery. The request packet is broadcasted in a request zone based on the baseline to determine the next broadcasting node. The neighboring node with the shortest distance to the baseline is chosen as the next broadcasting node. Thus, it is possible to find a better routing path than LAR scheme to reduce the network overhead.

![Fig. 7: LAR’s request and expected zone concepts.](image)

The improved version also proposes a partial reconstruction process that maintains a routing path. When a node on a routing path finds that a link is broken, the node starts the process of routing maintenance.

Chun Yang et al propose in (Yang and Chen, 2002), a reachability-guaranteed approach for reducing broadcast storms in MANET. The approach is based on location awareness of each node, which means each node in the network needs to equip the positioning device like GPS and exchanges location information in the HELLO message with its neighbors. Three mechanisms are included in the proposed approach: Relay Set (RS), Neighbor Coverage (NC), and Transmission Order (TO). The relay set is sender based mechanism in which the sending node of the broadcast message determines the relay set of its neighbors for rebroadcast according to the radio coverage of the neighbors (Yang and Chen, 2003). The first step in the RS algorithm is to sort the neighbors by distance of each neighbor to the sending node. Starting from the farthest neighbor, the sending node examines the radio-transmitting area of each neighbor to identify the neighbors that do not create new radio coverage. These neighbor nodes, which are called exclusive nodes, are actually the redundant nodes of rebroadcast and may not be included in the relay set. Since the overall radio-transmitting area of the nodes in the relay set completely covers the radio transmitting area of the exclusive nodes, the reachability of the RS algorithm is the same as blind flooding (i.e. 100% reachability) (Sun et al., 2001).
Fig. 8: Example Determining Relay Set

In figure 8, since the radio transmitting area of node N3 is totally covered by the radio-transmitting area of sending node S and two farther neighbors N1 and N2; node N3 is not included in the relay set of node S. The next mechanism is neighbor coverage (NC); the basic idea of this mechanism states that if a mobile node receiving a broadcast packet assures that all its neighbors have received the same packet, rebroadcast of the packet is actually redundant. Each node in the NC scheme records the packet ID as well as the neighbors from which one copy of the broadcast packet has been transmitted and calculates the neighbors that are not covered by the broadcast packet. Calculation of the non-covered neighbors is based on the copies of the broadcast packet received by the mobile node and the location of the neighbors that has sent the node one copy of the packet. On receiving a broadcast packet at the first time, the mobile node waits a random number of time slots before rebroadcast the packet (i.e. invoke the underlying CSMA/CA module for broadcasting). Multiple copies of the same broadcast packet may arrive during the waiting time. For the arrival of each copy of the same broadcast packet, the mobile node updates non-covered neighbors for the packet. If all neighbors of the mobile node are covered before the end of the waiting time, rebroadcast of the packet is cancelled. Figure 9. shows that node R has received three copies of a broadcast packet from its neighbors S1; S2; and S3: Since node R knows the locations of all its neighbors, it is easy to know if there are other neighbors of node R that are not covered by the radio-transmitting area of nodes S1; S2; and S3. It’s obvious from the figure that all neighbors are in the coverage of radio transmission of the three senders, node R decides not to rebroadcast the packet (Peng and Lu, 2000).

Fig. 9: Example Neighbor Coverage

The third and last mechanism is transmission order (TO), When all neighbors have received a broadcast packet transmitted by a mobile node, neither the RS scheme nor the NC scheme has a designate rebroadcast order for the neighbors. However, according to the rule of thumb for rebroadcast, a farther neighbor node should rebroadcast the packet earlier than closer neighbor nodes so that the closer nodes have more chance to find that it is redundant to rebroadcast the packet and cancel the rebroadcast operation (Tseng et al., 2003). That is upon receiving upon receiving a broadcast packet from a sending node, a farther neighbor node waits fewer time slots before rebroadcasting the packet than closer nodes. More specifically, when a mobile node R has received a broadcast packet from a sending node S: Node R calculates its transmission order for rebroadcast among the common neighbors of R and S: The number of waiting time slots of R is set the value of its transmission order. According to figure 10. node R identifies itself as the third node (N3 is the first and N1 is the second).
Heuristic Based Approaches:

These approaches aim to reduce the number of rebroadcasts. In their idea upon receiving a flooding packet a node decide whether to relay the packet to its neighbors or not (receiver based) using on of the following heuristics (Tseng et al., 2003):

1. Probabilistic based.
2. Counter based.
3. Distance based.
4. Location based
5. Cluster based.

In the Probabilistic scheme, each node rebroadcasts the message it received for the first time with some fixed probability $P$. Authors in (Thriveni et al., 2007) propose an algorithm to improve flooding performance over the AODV known as Probabilistic-Average-Energy-flooding (PAEF) which periodically performs an averaging algorithm-Calculate-Average-Energy (CAE) to estimate the average energy $E_{avg}$. This algorithm is used in route discovery process to confine the rebroadcast decision by the node. Jie Yang et al (2008) suggest an algorithm to study the connectivity of ad hoc networks and establish the correlations of connectivity with average node degree. Then, based on the simulation result of connectivity, they present a probabilistic flooding algorithm that can tradeoff the efficiency and reliability by using weighting factor. In the counter-based scheme, each node rebroadcasts the message only if the same message has not been heard for more than $C$ times, before it itself can transmit. The assumption made by the node is that if the message has been rebroadcast several times by its neighbors, then the extra coverage contribution from its own rebroadcast is probably too low to be worth transmitting, while in distance based scheme each node rebroadcasts the message only if the physical distance between itself and the node from which it received the message is not less than $d$. The node uses the signal strength of the received message to estimate this distance. Chien Chen et al., (2005) propose an algorithm called “DIS-RAD” which introduces the concept of distance into the counter-based broadcast scheme. The proposed approach gives nodes closer to the border a higher rebroadcast probability since they create better Expected Additional Coverage (EAC), where the distance threshold is adopted to distinguish between interior and border nodes, that two distinct Random Assessment delays (RADs) are applied to the border and interior nodes, with the border nodes having shorter (RADs) than the interior nodes. The location based scheme was reviewed in detail in the location awareness section, finally in the cluster-based scheme only cluster-heads and gateway nodes are able to rebroadcast the message. The nodes may use any of the other schemes to determine whether or not to rebroadcast the message.

As indicated in the table the 2-hop backward neighbor information enjoys the highest deliverability ratio, also did the 1-hop neighbor information scheme but slightly lower than the 2-hop scheme. Heuristics approached and location awareness have a significant improvement over pure flooding while the connected dominating set has a little bit better performance than blind flooding and its deliverability ratio starts to degrade specifically when the number of nodes becomes relatively larger in the ad hoc network.

The second column shows how each scheme participates in reducing the search space at the route discovery phase in the routing process, both the 2-hop and 1-hop schemes have a good performance in this issue because there will be a considerable excluding for nodes, that will not participate in the flooding process. The CDS-based and location awareness (RS-NC-TO) schemes make a significant improvement in this domain,
Protocol Comparison:

Table 1: Flooding Schemes Features:

<table>
<thead>
<tr>
<th>Method</th>
<th>Deliverability ratio</th>
<th>Redaction in search space</th>
<th>Number of collisions</th>
<th>Ratio of forwarding nodes</th>
<th>Network information</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-hop neighbor Information</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Local</td>
<td>Yes</td>
</tr>
<tr>
<td>2-hop neighbor information</td>
<td>Very high</td>
<td>Very high</td>
<td>Medium</td>
<td>Medium</td>
<td>Local</td>
<td>Yes</td>
</tr>
<tr>
<td>CDS</td>
<td>Medium</td>
<td>Medium</td>
<td>Low</td>
<td>Medium</td>
<td>Local</td>
<td>Yes</td>
</tr>
<tr>
<td>Location awareness</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Local</td>
<td>No</td>
</tr>
<tr>
<td>Heuristics approaches</td>
<td>High</td>
<td>medium</td>
<td>High</td>
<td>High</td>
<td>Global (on-demand)</td>
<td>Yes</td>
</tr>
<tr>
<td>AODV</td>
<td>High</td>
<td>medium</td>
<td>High</td>
<td>High</td>
<td>Global (on-demand)</td>
<td>Yes</td>
</tr>
<tr>
<td>DSR</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>Global (on-demand)</td>
<td>Yes</td>
</tr>
</tbody>
</table>

which is the major advantage of the dominating set is the reduction of the search space within only the nodes included in the set, also in the location awareness schemes two mechanisms are used to reduce the search space “relay set” and “neighbor coverage” which are sender-based and receiver-based schemes respectively, they are merged together in one scheme “hybrid approach” for further reduction with the aid of location information provided by GPS devices. In protocols like counter, distance, probabilistic based all the space will be searched to find a route between the source and destination nodes. Third column shows the number of collisions that could occur with each scheme, the 1-hop and 2-hop schemes provide the lowest rate of collisions, location awareness and heuristic approaches could optimize the number of collisions than pure flooding , while CDS performs well just when the network size is small or medium but it starts to suffer from high percentage of collisions when the network size tend to be large. The ratio of forwarding nodes are minimized at both the 1-hop and 2-hop schemes as indicated in the fourth column, location awareness makes a significant improvement due to the its hybrid approach while in CDS the number of forwarding nodes increases linearly with network size (number of nodes) because with the expansion of network size, members of the dominating set increase also and this is according to the principle of dominating that is a set is dominating if all nodes in the system are either in the set or neighbors of nodes in the set. The network information that needed to reduce the number of rebroadcasts is local for the 1-hop neighbor scheme, location awareness, CDS and heuristic-based approaches, while its quazi-local for the two hop neighbor scheme. The latest column indicates the reliability of each scheme, where all of the schemes are considered as reliable except the heuristic ones due to their dependency on threshold values which are fixed and this may result in some messages not being broadcast to the destinations under certain conditions i.e. when the network is sparse.

Conclusion:

In this paper we review and compare between the most prominent flooding schemes that were recently proposed to improve the flooding activity within on-demand routing protocols over mobile ad hoc wireless networks. The key objective of the optimization techniques that were proposed is to minimize the amount of control traffic generated in a route discovery, but some times this will effect other aspects of performance in ways that are not always desired. The early quenching of route requests by intermediate nodes may result in fewer packets and shorter query time. However, the routes obtained can be obsolete or non-optimal, which results in both increased packet loss and latency. Inherently, there exists a tradeoff between overhead of route discovery and other performance areas. therefore the future work should focus on whether such tradeoffs in performance can be averted by exploring some ideas that include adaptive methods for optimizing route discovery For example, if the message to be sent is urgent, or the current network utilization is low, then flooding may be used to discover a better route in a shorter time. If not, better resource conserving techniques such as one of the efficient flooding schemes may be employed for route discovery.

REFERENCES


