Chapter 2

MANETs: AN OVERVIEW

2.1 INTRODUCTION

Mobile ad hoc network is a collection of mobile nodes that are dynamically and arbitrarily located in such a manner that the interconnections between nodes are capable of changing on a continual basis. These networks are built by the wireless nodes which have to do the routing operation as well in addition to the information exchange. The MANET topology changes frequently as mobile hosts migrate, disappear (failure or depletion of battery capacity), or adjust their transmission and reception characteristics.

Mobile ad-hoc network finds tremendous applications in disaster recovery operations, collaborative and distributed computing, hybrid mesh networks and military operations in remote areas [26-29]. Mostly the deployment of MANET is in areas where there is a scarcity of resources. Therefore, it is necessary that the resources in the MANETs be used in an efficient manner. The next section discusses the basic characteristics of these networks.

2.2 CHARACTERISTICS OF MANET

MANET [30-32] is a collection of wireless devices that come together to form a self-organizing network without any support from the existing fixed communication infrastructure. These devices are mobile in nature and are free to move in and out of the network [33-34] with/without notifying the other nodes in the network.

The communication in ad-hoc networks is multi-hop [35] in nature. Each node in the network has to relay data packets of other nodes for the successful communication of MANET as shown in Fig. 2.1.
This strategy of multi-hop communication ensures longer battery life because a lot of transmitting power (proportional to fourth power of the distance between the communicating nodes) is saved. The main characteristics [35] of MANET are as follows:

- **Dynamic topology**: The network topology in MANET is highly dynamic due to movement of nodes; hence an ongoing session suffers frequent path breaks. Disruption occurs either due to the movement of intermediate nodes or due to movement of end nodes. Such situations do not arise in wired networks wherein the nodes are stationary.

- **Bandwidth constraints and variable link capacity**: Abundant bandwidth is available in wired networks due to the advent of fibre optics and due to exploitation of wavelength division multiplexing technologies. But in wireless networks, the radio band is limited, and hence the data rates available are far less in comparison to wired networks.

- **Self-configuring**: MANET has decentralized infrastructure, with all mobile nodes functioning as routers and all wireless devices being interconnected to one another. Intuitively, this means that the MANET is also a self-configuring network in which network activities like the discovery of the topology and delivery of messages, are executed by the nodes themselves.
- **Energy constrained nodes**: Mobile nodes rely on batteries for proper operation. Since an ad hoc network consists of several nodes, depletion of batteries in these nodes will have a great influence on overall network performance. While designing any protocol for MANET this factor is taken into account.

- **Multi-hop communications**: Due to signal propagation characteristics of wireless transceivers, ad hoc networks require the support of multihop communications i.e. mobile nodes that cannot reach the destination node directly will need to relay their messages through other nodes.

- **Limited security**: In MANETs we don’t have centralized devices such as routers and switches which can share major burden of security. The nodes in these networks must have in built capabilities to avoid resource consumption, denial of service, impersonation and similar attacks possible against MANET.

### 2.3. ADVANTAGES OF MANET

Mobile Ad Hoc Network due to its infrastructure-less structure and node mobility posses following advantages:

- **Fast installation**: The level of flexibility for setting up MANET is high, since they do not require any previous installation or infrastructure therefore can be brought up and torn down in very short time.

- **Dynamic topologies**: Nodes can arbitrarily move around the network and can disappear temporarily from the MANET, so the network topology graph can continuously change at undetermined speed.

- **Fault tolerance**: Owing to the limitations of the radio interfaces and the dynamic topology, MANET is robust enough to handle connection failures. The routing and transmission control protocols are designed to manage these situations.
• **Connectivity**: The use of centralized points or gateways is not required for the communication within the MANET and the communication takes place due to the collaboration between the participating nodes.

• **Mobility**: The wireless mobile nodes can move with random speed and can take random direction. Although the routing algorithms deal with this issue, the performance shows that there is a threshold level of node mobility after which the protocol operation begins to fail.

• **Cost**: MANETs are more economical in the case of temporary installation as they eliminate fixed infrastructure costs and reduce power consumption at mobile nodes.

• **Spectrum reuse possibility**: Owing to short communication links (node-to-node instead of node to a central base station) it is possible to use same frequency band in different places of the network.

### 2.4 APPLICATIONS OF MANETS

Mobile ad hoc network, due to their quick and economic deployment find applications in several areas. This set of applications for MANETs [26-29] ranges from small static networks that are constrained by power sources, to large-scale mobile highly dynamic networks. MANETs can be used in large variety of situations, where conventional networking cannot be applied because of difficult terrain, heavy cost or other reasons. MANET is used in following areas:

• **Military Tactical Networks**: MANET can be very useful in establishing communication among a group of soldiers for tactical operations [36]. Setting up a fixed infrastructure for communication among a group of soldiers in enemy territories or in inhospitable terrains may not be possible. In such environment, MANET provides required communication quickly without the support of any fixed infrastructure.

• **Emergency Operations**: MANETS are very useful in emergency operations such as search and rescue, crowd control and commando operations. The major factors that favors these networks are: self-configuring of the system with
minimal overhead, independent of fixed and centralized infrastructure, the nature of the terrain of such applications, the freedom and flexibility of mobility and the unavailability of conventional communication infrastructure.

- **Personal Area Network:** The concept of personal area networks is about interconnecting different devices used by a single person, e.g. PDA, cellular phone, laptop etc. In this case the PDA or the laptop will connect with the cellular phone in an ad hoc fashion. If both the PDA and the printer were ad hoc enabled the PDA could automatically get access to the printing services.

- **Sensor Networks:** Sensor networks are a special category of ad hoc wireless networks that are used to provide a wireless communication infrastructure among the sensors deployed in a specific application domain. Recent advances in wireless communication technology and research in MANET have made smart sensing a reality.

- **Collaborative Networking:** Another domain in which the MANET finds application is collaborative computing. The requirement of a temporary communication infrastructure for quick communication with minimal configuration among a group of peoples in a conference or gathering necessitates the formation of ad hoc network.

2.5 **LIMITATION OF MANET**

The highly adaptive networking technology in the form of MANET still faces various limitations [37-44]. Some of these limitations are as given below:

- **Bandwidth constraints:** As mentioned above, the capacity of the wireless links is always much lower than in the wired links. Several Gbps are available for wired LAN nowadays while the commercial applications for wireless LANs work typically around 2 Mbps.

- **Processing capability:** Most of the nodes in MANETs are devices without a powerful CPU. The network tasks such as routing and data transmission cannot consume the power resources of the device, as it is intended to play other roles such as sensing functions.
• **Energy constraints**: The power of the batteries is limited, which does not allow infinitive operation time for the nodes. Therefore energy should not be wasted and that is why several energy conserving algorithms have been implemented.

• **High latency**: In an energy conserving design, nodes are sleeping or idle when they do not have to transmit any data. When the data exchange between two nodes goes through nodes that are sleeping, the delay may be higher if the routing algorithm decides that these nodes have to wake up.

• **Transmission errors**: Attenuation and interferences are other effects of the wireless links that increase the error rate.

• **Security**: The authors divide the possible attacks in passive ones, when the attacker only attempts to discover valuable information by listening to the routing traffic; and active attacks, which occur when the attacker injects arbitrary packets into the network with some proposal like disabling the network.

• **Addressing**: The addressing is the another problem for the network layer in MANET, since the information about the IP address used in fixed networks offers some facilities for routing that cannot be applied in MANET.

• **Commercially unavailable**: MANET is yet far from being deployed on large-scale commercial basis.

Due to all the above limitations mentioned above, it is always a challenging task for the researchers to develop a routing protocol for MANET. The next section discusses about the routing strategies present in the protocols available in the literature for MANET.

### 2.6 ROUTING IN MANET

Routing [45] is the process in which a route from a source to a destination node is identified. In order to facilitate communication within MANET, a routing protocol is used to discover routes between nodes. The primary goal of such a routing protocol is to ensure correct and efficient route establishment between a pair of nodes so that messages are delivered in a timely manner. The desirable properties of a routing protocol are as follows:
• **Distributed operation:** The protocol should be distributed. It should not be resident on a particular node. All the nodes in the network have to be aware of the rules embedded in routing protocol and be able to follow them in the true spirit. This is the case even for stationary ad hoc networks.

• **Loop free:** The routing protocol need to ensure that the routes supplied are loop free. Any loop in the route formation can lead to unnecessary wastage of resources on continuous basis.

• **Demand base operation:** Instead of assuming a uniform traffic distribution within the network and maintaining routing between all nodes at all times, the routing protocol should be able to adapt to the traffic pattern. This adaptation should be done in an intelligent manner so as to utilize network energy and bandwidth resources in efficient manner.

• **Unidirectional link support:** The radio environment can cause the formation of unidirectional links. Utilization of these links improves the routing performance.

• **Security:** MANET is especially vulnerable to impersonation attacks, so to ensure the proper working of the routing protocol, there is a need for some sort of preventive security measures. Authentication and encryption is the way to tackle this problem in conventional networks but in case of ad hoc networks the problem lies with distributing keys among the nodes.

• **Sleep period operation:** As a result of energy conservation, or some other need to be inactive, nodes of a MANET may stop transmitting and/or receiving (even receiving requires power) for arbitrary time periods. A routing protocol should be able to accommodate such sleep periods without overly adverse consequences.

• **Multiple routes:** To reduce the number of reactions to topological changes and to control congestion, multiple routes could be used. If one route is invalid, then other stored route is active for communication and thus saving the routing protocol from initiating another route discovery procedure.
• **Quality of service support**: Some sort of quality of service support is probably necessary to incorporate into the routing protocol. This has a lot to do with what these networks will be used for. It could be for instance real time traffic support.

While designing any routing protocol all these points are kept in mind by the researchers but achieving simultaneously all the parameters is still a challenging task for them even today. The routing protocols for mobile ad hoc network can be categorized on the basis of how routing information is acquired and maintained by mobile nodes [45-46] into three categories as follows:

- Proactive routing or Table driven routing
- Reactive routing or On demand routing
- Hybrid routing

### 2.6.1 PROACTIVE ROUTING PROTOCOL

A proactive routing protocol is also called "Table driven" routing protocol [45]. Using a proactive routing protocol, nodes in a mobile ad hoc network continuously evaluate routes to all reachable nodes and attempt to maintain consistent, up-to-date routing information. Therefore, a source node can get a routing path immediately if it needs one. In proactive routing protocols, all nodes need to maintain a consistent view of the network topology. When a network topology change occurs, respective updates must be propagated throughout the network to notify the change. Most proactive routing protocols proposed for mobile ad hoc networks have inherited properties from algorithms used in wired networks. To adapt to the dynamic features of mobile ad hoc networks, necessary modifications have been made on traditional wired network routing protocols. Using proactive routing algorithms, mobile nodes proactively update network state and maintain a route regardless of whether data traffic exists or not, the overhead to maintain up-to-date network topology information is very high. Some of the popular routing protocols that falls in this category are as follows:

- Wireless Routing Protocol (WRP)
- Destination Sequence Distance Vector (DSDV)
Fisheye State Routing (FSR)
Clusterhead Gateway Switch Routing (CGSR)

2.6.1.1 Wireless Routing Protocol (WRP)

The Wireless Routing Protocol [47] is a proactive unicast routing protocol for mobile ad hoc networks. WRP uses improved Bellman-Ford Distance Vector routing algorithm. To adapt to the dynamic features of mobile ad hoc networks, mechanisms are introduced to ensure the reliable exchange of update messages and reduces route loops. In an ad hoc network using WRP, each node maintains four tables as given below:

- Distance Table
- Routing Table
- Message Retransmission List
- Link Cost Table

An entry in the routing table contains the distance to a destination node, the predecessor and the successor along the paths to the destination, and a tag to identify its state, i.e., is it a simple path, a loop or invalid. Storing predecessor and successor in the routing table helps to detect routing loops and avoid counting-to-infinity problem, which is the main shortcoming of the original distance vector routing algorithm. A mobile node creates an entry for each neighbor in its link-cost table. The entry contains cost of the link connecting to the neighbor, and the number of timeouts since an error-free message was received from that neighbor.

In WRP, mobile nodes exchange routing tables with their neighbors using update messages. The update messages can be sent either periodically or whenever link state changes happen. The MRL contains information about which neighbor has not acknowledged an update message. If needed, the update message will be retransmitted to the neighbor. Additionally, if there is no change in its routing table since last update, a node is required to send a Hello message to ensure connectivity. On receiving an
update message, the node modifies its distance table and looks for better routing paths according to the updated information.

In WRP, a node checks the consistency of its neighbors after detecting any link change. A consistency check helps to eliminate loops and speed up convergence. One shortcoming of WRP is that it needs large memory storage and computing resource to maintain several tables. Moreover, as a proactive routing protocol, it has a limited scalability and is not suitable for large mobile ad hoc networks.

### 2.6.1.2 Destination Sequence Distance Vector (DSDV)

The Destination Sequence Distance Vector [48] is a proactive unicast mobile ad hoc network routing protocol. Like WRP, DSDV is also based on the traditional Bellman-Ford algorithm. However, their mechanisms to improve routing performance in mobile ad hoc networks are quite different.

In routing tables of DSDV, an entry stores the next hop towards a destination, the cost metric for the routing path to the destination and a destination sequence number that is created by the destination. Sequence numbers are used in DSDV to distinguish stale routes from fresh ones and avoid formation of route loops.

The route updates of DSDV can be either time-driven or event-driven. Every node periodically transmits updates including its routing information to its immediate neighbors. While a significant change occurs from the last update, a node can transmit its changed routing table in an event-triggered style. Moreover, the DSDV has two ways when sending routing table updates. One is "full dump" update type and the full routing table is included inside the update. A “full dump” update could span many packets. An incremental update contains only those entries that with metric have been changed since the last update is sent. Additionally, the incremental update fits in one packet.
2.6.1.3 The Fisheye State Routing (FSR)

The Fisheye State Routing (FSR) [49] is a proactive unicast routing protocol based on Link State routing algorithm with effectively reduced overhead to maintain network topology information. As indicated in its name, FSR utilizes a function similar to a fish eye. The eyes of fishes catch the pixels near the focal with high detail, and the detail decreases as the distance from the focal point increases. Similar to fish eyes, FSR maintains the accurate distance and path quality information about the immediate neighboring nodes, and progressively reduces detail as the distance increases.

In Link State routing algorithm used for wired networks, link state updates are generated and flooded through the network whenever a node detects a topology change. In FSR, however, nodes exchange link state information only with the neighboring nodes to maintain up-to-date topology information. Link state updates are exchanged periodically in FSR, and each node keeps a full topology map of the network. To reduce the size of link state update messages, the key improvement in FSR is to use different update periods for different entries in the routing table. Link state updates corresponding to the nodes within a smaller scope are propagated with higher frequency.

FSR exhibits a better scalability concerning the network size compared to other link state protocols because it doesn't strive for keeping all nodes in the network on the same knowledge level about link states. Instead, the accuracy of topology information is inversely proportional to the distance. This reduces the traffic overhead caused by exchanging link state information because this information is exchanged more frequently with node nearby than with nodes far away.

2.6.1.4 Clusterhead Gateway Switch Routing

The Clusterhead Gateway Switch Routing (CGSR) protocol differs from the previous protocol in the type of addressing and network organization scheme employed. Instead of a “flat” network, CGSR is a clustered multi-hop mobile wireless network with several heuristic routing schemes [50]. The authors state that by having a cluster head
controlling a group of ad hoc nodes, a framework for code separation (among clusters), channel access, routing, and bandwidth allocation can be achieved. A cluster head selection algorithm is utilized to elect a node as the cluster head using a distributed algorithm within the cluster. The disadvantage of having a cluster head scheme is that frequent cluster head changes can adversely affect routing protocol performance since nodes are busy in cluster head selection rather than packet relaying. Hence, instead of invoking cluster head reselection every time the cluster membership changes, a Least Cluster Change (LCC) clustering algorithm is introduced. Using LCC, cluster heads only change when two cluster heads come into contact, or when a node moves out of contact of all other cluster heads.

Fig. 2.2 CGSR: Routing from Node 1 to 8

- Nodes
- Cluster Heads
- Gateway
CGSR uses DSDV as the underlying routing scheme, and hence has much of the same overhead as DSDV. However, it modifies DSDV by using a hierarchical cluster-head-to-gateway routing approach to route traffic from source to destination. Gateway nodes are nodes that are within communication range of two or more cluster heads. A packet sent by a node is first routed to its cluster head, and then the packet is routed from the cluster head to a gateway to another cluster head, and so on until the cluster head of the destination node is reached. The packet is then transmitted to the destination. Fig. 2.2 illustrates an example of this routing scheme. Using this method, each node must keep a “cluster member table” where it stores the destination cluster head for each mobile node in the network. These cluster member tables are broadcast by each node periodically using the DSDV algorithm. Nodes update their cluster member tables on reception of such a table from a neighbor.

In addition to the cluster member table, each node must also maintain a routing table which is used to determine the next hop in order to reach the destination. On receiving a packet, a node will consult its cluster member table and routing table to determine the nearest cluster head along the route to the destination. Next, the node will check its routing table to determine the next hop used to reach the selected cluster head. It then transmits the packet to this node.

2.6.1.5 Comparison of Table Driven Routing Protocols

Although belonging to the same routing category for mobile ad hoc networks, WRP, DSDV and FSR have distinct features as follows:

- Both WRP and DSDV exploited event-triggered updates to maintain up-to-date and consistent routing information for mobile nodes. In contrast to using event-triggered updates, the updates in FSR are exchanged between neighboring nodes and the update frequency is dependent on the distance between nodes. In this way, update overhead is reduced and the far-reaching effect of Link State routing is restricted.
- Different mechanisms are used in WRP, DSDV and FSR for loop-free guarantee. WPR records the predecessor and the successor along a path in its
routing table and introduces consistence-checking mechanism. In this way, WRP avoids forming temporary route loops but incurs additional overhead. Every node needs to maintain more information and execute more operations. In DSDV, a destination sequence number is introduced to avoid route loops. FSR is a modification of traditional Link State routing and its loop-free property is inherited from Link State routing algorithm.

- WRP has a large storage complexity compared to DSDV because more information is required in WRP to guarantee reliable transmission and loop-free paths. Both periodic and triggered updates are utilized in WRP and DSDV; therefore, their performance is tightly related with the network size and node mobility pattern. As a Link State routing protocol, FSR has high storage complexity, but it has potentiality to support multiple-path routing and QoS routing.

### 2.6.2 REACTIVE ROUTING PROTOCOLS

Reactive routing protocols for mobile ad hoc networks are also called "on-demand" routing protocols [45-46]. In a reactive routing protocol, routing paths are searched only when needed. A route discovery operation invokes a route-determination procedure. The discovery procedure terminates either when a route has been found or no route available after examination for all route permutations.

In a mobile ad hoc network, active routes may be disconnected due to node mobility. Therefore, route maintenance is an important operation of reactive routing protocols. Compared to the proactive routing protocols for mobile ad hoc networks, less control overhead is a distinct advantage of the reactive routing protocols. Thus, reactive routing protocols have better scalability than proactive routing protocols in mobile ad hoc networks. However, when using reactive routing protocols, source nodes may suffer from long delays for route searching before they can forward data packets. Some of the popular routing protocols that fall in this category are as follows:

- Dynamic Source Routing (DSR) Protocol
• Ad Hoc On-demand Distance Vector Routing (AODV) protocol
• Temporally Ordered Routing Algorithm (TORA)
• Associativity-Based Routing (ABR)

2.6.2.1 The Dynamic Source Routing (DSR) Protocol

The Dynamic Source Routing (DSR) [51] is a reactive unicast routing protocol that utilizes source routing algorithm. In source routing algorithm, each data packet contains complete routing information to reach its dissemination. Additionally, in DSR each node uses caching technology to maintain route information that it has learnt. There are two major phases in DSR,

_The Route Discovery phase_

In this phase when a source node wants to send a packet, it firstly consults its route cache. If the required route is available, the source node includes the routing information inside the data packet before sending it. Otherwise, the source node initiates a route discovery operation by broadcasting route request packets. A route request packet contains addresses of both the source and the destination and a unique number to identify the request. Receiving a route request packet, a node checks its route cache. If the node doesn’t have routing information for the requested destination, it appends its own address to the route record field of the route request packet. Then, the request packet is forwarded to its neighbors. To limit the communication overhead of route request packets, a node processes route request packets that both it has not seen before and its address is not presented in the route record field. If the route request packet reaches the destination or an intermediate node has routing information to the destination, a route reply packet is generated. When the route reply packet is generated by the destination, it comprises addresses of nodes that have been traversed by the route request packet. Otherwise, the route reply packet comprises the addresses of nodes the route request packet has traversed concatenated with the route in the intermediate node’s route cache.
After being created, either by the destination or an intermediate node, a route reply packet needs a route back to the source. There are three possibilities to get a backward route. The first one is that the node already has a route to the source. The second possibility is that the network has symmetric (bi-directional) links. The route reply packet is sent using the collected routing information in the route record field, but in a reverse order as shown in Fig.2.3. In the last case, there exists asymmetric (uni-directional) links and a new route discovery procedure is initiated to the source. The discovered route is piggybacked in the route request packet.

![Route Reply with Route Record in DSR](image)

**Fig. 2.3 Route Reply with Route Record in DSR**

The Route Maintenance Phase

In this phase, when the data link layer detects a link disconnection, a ROUTE_ERROR packet is sent backward to the source. After receiving the ROUTE_ERROR packet, the source node initiates another route discovery operation. Additionally, all routes containing the broken link should be removed from the route caches of the immediate nodes when the ROUTE_ERROR packet is transmitted to the source.

DSR has increased traffic overhead by containing complete routing information into each data packet, which degrades its routing performance.
2.6.2.2 Ad Hoc On-demand Distance Vector Routing (AODV) protocol

The Ad Hoc On-demand Distance Vector Routing [52] protocol is a reactive unicast routing protocol for mobile ad hoc networks. As a reactive routing protocol, AODV only needs to maintain the routing information about the active paths. In AODV, routing information is maintained in routing tables at nodes. Every mobile node keeps a next-hop routing table, which contains the destinations to which it currently has a route. A routing table entry expires if it has not been used or reactivated for a pre-specified expiration time. Moreover, AODV adopts the destination sequence number technique used by DSDV in an on-demand way.

Fig. 2.4 Flooding RREQ in AODV

In AODV, when a source node wants to send packets to the destination but no route is available, it initiates a route discovery operation. In the route discovery operation, the source node floods route request (RREQ) packets as shown in Fig. 2.4. A RREQ includes addresses of the source and the destination, the broadcast ID, which is used as its identifier, the last seen sequence number of the destination as well as the source node’s sequence number. Sequence numbers are important to ensure loop-free and up-to-date routes. To reduce the flooding overhead, a node discards RREQs that it has seen before and the expanding ring search algorithm is used in route discovery operation. The RREQ starts with a small TTL (Time-To-Live) value. If the destination is not found, the TTL is increased in following RREQs.
In AODV, each node maintains a cache to keep track of RREQs it has received. The cache also stores the path back to each RREQ originator. When the destination or a node that has a route to the destination receives the RREQ, it checks the destination sequence numbers it currently knows and the one specified in the RREQ. To guarantee the freshness of the routing information, a route reply (RREP) packet is created and forwarded back to the source (see Fig. 2.5) only if the destination sequence number is equal to or greater than the one specified in RREQ. AODV uses only symmetric links and a RREP follows the reverse path of the respective RREP. Upon receiving the RREP packet, each intermediate node along the route updates its next-hop table entries with respect to the destination node. The redundant RREP packets or RREP packets with lower destination sequence number will be dropped.

![Fig. 2.5 Route Reply in AODV](image)

In AODV, a node uses hello messages to notify its existence to its neighbors. Therefore, the link status to the next hop in an active route can be monitored. When a node discovers a link disconnection, it broadcasts a route error (RERR) packet to its neighbors, which in turn propagates the RERR packet towards nodes whose routes may be affected by the disconnected link. Then, the affected source can re-initiate a route discovery operation if the route is still needed.
2.6.2.3 Temporally Ordered Routing Algorithm (TORA)

The Temporally Ordered Routing Algorithm [53-54] is a reactive routing algorithm based on the concept of link reversal. TORA improves the partial link reversal method by detecting partitions and stopping non-productive link reversals. TORA can be used for highly dynamic mobile ad hoc networks.

In TORA, the network topology is regarded as a directed graph. A Directional Acyclical Graph (DAG) is accomplished for the network by assigning each node i a height metric hi. A link directional from i to j means hi > hj. In TORA, the height of a node is defined as a quintuple, which includes the logical time of a link failure, the unique ID of the node that defines the new reference level, a reflection indicator bit, a propagation ordering parameter and a unique ID of the node. The first three elements collectively represent the reference level. The last two values define an offset with respect to the reference level. Like water flowing, a packet goes from upstream to downstream according the height difference between nodes. DAG provides TORA the capability that many nodes can send packets to a given destination and guarantees that all routes are loop-free.

TORA has three basic operations: route creation, route maintenance and route erasure. A route creation operation starts with setting the height (propagation ordering parameter in the quintuple) of the destination to 0 and heights of all other nodes to NULL (i.e., undefined). The source broadcasts a QRY packet containing the destination’s ID. A node with a non-NULL height responds by broadcasting a UPD packet containing the height of its own. On receiving a UPD packet, a node sets its height to one more than that of the UPD generator. A node with higher height is considered as upstream and the node with lower height is considered as downstream. In this way, a directed acyclic graph is constructed from the source to the destination and multiple paths route may exist. In this way, a directed acyclic graph is constructed from the source to the destination and multiple paths route may exist.
The DAG in TORA may be disconnected because of node mobility. So, route maintenance operation is an important part of TORA. TORA has the unique feature that control messages are localized into a small set of nodes near the occurrence of topology changes. After a node loses its last downstream link, it generates a new reference level and broadcasts the reference to its neighbors. Therefore, links are reversed to reflect the topology change and adapt to the new reference level. The erase operation in TORA floods CLR packets through the network and erase invalid routes.

2.6.2.4 Associativity-Based Routing (ABR)

A totally different approach in mobile routing is proposed in [55]. The Associativity-Based Routing (ABR) protocol is free from loops, deadlock, and packet duplicates, and defines a new routing metric for ad hoc mobile networks. This metric is known as the degree of association stability. In ABR, a route is selected based on the degree of association stability of mobile nodes. Each node periodically generates a beacon to signify its existence. When received by neighboring nodes, this beacon causes their associativity tables to be updated. For each beacon received, the associativity tick of the current node with respect to the beaconing node is incremented. Association stability is defined by connection stability of one node with respect to another node over time and space. A high degree of association stability may indicate a low state of node mobility, while a low degree may indicate a high state of node mobility. Associativity ticks are reset when the neighbors of a node or the node itself move out of proximity. A fundamental objective of ABR is to derive longer-lived routes for ad hoc mobile networks. The three phases of ABR are:

- Route discovery
- Route reconstruction (RRC)
- Route deletion

The route discovery phase is accomplished by a broadcast query and await-reply (BQ-REPLY) cycle. A node desiring a route broadcasts a BQ message in search of mobiles that have a route to the destination. All nodes receiving the query (that are not the destination) append their addresses and their associativity ticks with their neighbors along with QoS information to the query packet. A successor node erases its upstream
node neighbors’ associativity tick entries and retains only the entry concerned with
itself and its upstream node. In this way, each resultant packet arriving at the
destination will contain the associativity ticks of the nodes along the route to the
destination. The destination is then able to select the best route by examining the
associativity ticks along each of the paths. When multiple paths have the same overall
degree of association stability, the route with the minimum number of hops is selected.
The destination then sends a REPLY packet back to the source along this path. Nodes
propagating the REPLY mark their routes as valid. All other routes remain inactive,
and the possibility of duplicate packets arriving at the destination is avoided. RRC may
 consist of partial route discovery, invalid route erasure, valid route updates, and new
route discovery, depending on which node(s) along the route move. Whenever a source
node is displaced from its position, message is sent to the downstream nodes for route
notification and the downstream neighbors erases its route entries. Similarly when the
destination moves, the immediate upstream node erases its route and determines if the
node is still reachable by a localized query (LQ[H]) process, where H refers to the
hopcount from the upstream node to the destination. If the destination receives the LQ
packet, it REPLYs with the best partial route; otherwise, the initiating node times out
and the process backtracks to the next upstream node. Here an RN[0] message is sent to
the next upstream node to erase the invalid route and inform this node that it should
invoke the LQ[H] process. If this process results in backtracking more than halfway to
the source, the LQ process is discontinued and a new BQ process is initiated at the
source.

When a discovered route is no longer desired, the source node initiates a route delete
(RD) broadcast so that all nodes along the route update their routing tables. The RD
message is propagated by a full broadcast, as opposed to a directed broadcast, because
the source node may not be aware of any route node changes that occurred during
RRCs next upstream node to erase the invalid route and inform this node that it should
invoke the LQ[H] process. If this process results in backtracking more than halfway to
the source, the LQ process is discontinued and a new BQ process is initiated at the
source.
2.6.2.5 Comparison of Reactive Routing Protocols

The main differences between DSR, TORA and AODV are as follows:

- DSR exploits source routing and routing information caching. A data packet in DSR carries the routing information needed in its route record field. DSR uses flooding in the route discovery phase. AODV adopts the similar route discovery mechanism used in DSR, but stores the next hop routing information in the routing tables at nodes along active routes. Therefore, AODV has less traffic overhead and is more scalable because of the size limitation of route record field in DSR data packets.

- Both DSR and TORA support unidirectional links and multiple routing paths, but AODV doesn’t. In contrast to DSR and TORA, nodes using AODV periodically exchange hello messages with their neighbors to monitor link disconnections. This incurs extra control traffic overhead. In TORA, utilizing the "link reversal" algorithm, DAG constructs routing paths from multiple sources to one destination and supports multiple routes and multicast [37].

- In AODV and DSR, a node notifies the source to re-initiate a new route discovery operation when a routing path disconnection is detected. In TORA, a node re-constructs DAG when it lost all downstream links. Both AODV and DSR use flooding to inform nodes that are affected by a link failure. However, TORA localizes the effect in a set of node near the occurrence of the link failure.

- AODV uses sequence numbers to avoid formation of route loops. Because DSR is based on source routing, a loop can be avoided by checking addresses in route record field of data packets. In TORA, each node in an active route has a unique height and packets are forwarded from a node with higher height to a lower one. So, a loop-free property can be guaranteed in TORA. However, TORA has an extra requirement that all nodes must have synchronized clocks. In TORA, oscillations may occur when coordinating nodes currently execute the same operation.
- ABR routing protocol requires beacons packet to be generated periodically whereas it is not so with the other protocols.

2.6.3 HYBRID ROUTING PROTOCOLS

Hybrid routing protocols are proposed to combine the merits of both proactive and reactive routing protocols and overcome their shortcomings. Normally, hybrid routing protocols for mobile ad hoc networks exploit hierarchical network architectures. Proper proactive routing approach and reactive routing approach are exploited in different hierarchical levels, respectively. Some of the popular routing protocols that fall in this category are as follows:

- Zone Routing Protocol (ZRP)
- Zone-based Hierarchical Link State routing (ZHLS)
- Hybrid Ad hoc Routing Protocol (HARP)
- Distributed Dynamic Routing (DDR)

2.6.3.1 The Zone Routing Protocol (ZRP)

The Zone Routing Protocol [56-57] is a hybrid routing protocol for mobile ad hoc networks. The hybrid protocols are proposed to reduce the control overhead of proactive routing approaches and decrease the latency caused by route search operations in reactive routing approaches.

In ZRP, the network is divided into routing zones according to distances between mobile nodes. Given a hop distance d and a node N, all nodes within hop distance at most d from N belong to the routing zone of N. Peripheral nodes of N are N’s neighboring nodes in its routing zone which are exactly d hops away from N.

In ZRP, different routing approaches are exploited for inter-zone and intra-zone packets. The proactive routing approach, i.e., the Intra-zone Routing protocol (IARP), is used inside routing zones and the reactive Inter-zone Routing Protocol (IERP) is used between routing zones, respectively. The IARP maintains link state information for
nodes within specified distance \(d\). Therefore, if the source and destination nodes are in the same routing zone, a route can be available immediately. Most of the existing proactive routing schemes can be used as the IARP for ZRP. The IERP reactively initiates a route discovery when the source node and the destination are residing in different zones. The route discovery in IERP is similar to DSR with the exception that route requests are propagated via peripheral nodes.

### 2.6.3.2 Zone-based Hierarchical Link State routing (ZHLS)

The Zone-based Hierarchical Link State routing (ZHLS) [58] is a hybrid routing protocol. In ZHLS, mobile nodes are assumed to know their physical locations with assistance from a locating system like GPS. The network is divided into non-overlapping zones based on geographical information.

ZHLS uses a hierarchical addressing scheme that contains zone ID and node ID. A node determines its zone ID according to its location and the pre-defined zone map is well known to all nodes in the network. It is assumed that a virtual link connects two zones if there exists at least one physical link between the zones. A two-level network topology structure is defined in ZHLS, the node level topology and the zone level topology. Respectively, there are two kinds of link state updates, the node level LSP (Link State Packet) and the zone level LSP. A node level LSP contains the node IDs of its neighbors in the same zone and the zone IDs of all other zones. A node periodically broadcast its node level LSP to all other nodes in the same zone. Therefore, through periodic node level LSP exchanges, all nodes in a zone keep identical node level link state information. In ZHLS, gateway nodes broadcast the zone LSP throughout the network whenever a virtual link is broken or created. Consequently, every node knows the current zone level topology of the network.

Before sending packets, a source firstly checks its intra-zone routing table. If the destination is in the same zone as the source, the routing information is already there. Otherwise, the source sends a location request to all other zones through gateway nodes. After a gateway node of the zone, in which the destination node resides, receives
the location request, it replies with a location response containing the zone ID of the destination. The zone ID and the node ID of the destination node will be specified in the header of the data packets originated from the source. During the packet forwarding procedure, intermediate nodes except nodes in the destination zone will use inter-zone routing table, and when the packet arrives the destination zone, an intra-zone routing table will be used.

2.6.3.3 Hybrid Ad hoc Routing Protocol (HARP)

The Hybrid Ad hoc Routing Protocol [59] is a hybrid routing scheme, which exploits a two-level zone based hierarchical network structure. Different routing approaches are utilized in two levels, for intra-zone routing and inter-zone routing, respectively.

The Distributed Dynamic Routing (DDR) [60] algorithm is exploited by HARP to provide underlying supports. In DDR, nodes periodically exchange topology messages with their neighbors. A forest is constructed from the network topology by DDR in a distributed way. Each tree of the forest forms a zone. Therefore, the network is divided into a set of non-overlapping dynamic zones.

A mobile node keeps routing information for all other nodes in the same zone. The nodes belonging to different zones but are within the direct transmission range are defined as gateway nodes. Gateway nodes have the responsibility forwarding packets to neighboring zones. In addition to routing information for nodes in the local zone, each node also maintains those of neighboring zones.

As in ZRP, the intra-zone routing of HARP relies on an existing proactive scheme and a reactive scheme is used for inter-zone communication. Depending on whether the forwarding and the destination node are inside the same zone, the respective routing scheme will be applied.
2.6.3.4 Distributed Dynamic Routing (DDR)

DDR [60] is also a tree-based routing protocol. However, unlike DST, in DDR the trees do not require a root node. In this strategy tree are constructed using periodic beaconing messages which are exchanged by neighboring nodes only. The trees in the network form a forest, which is connected together via gateway nodes (i.e. nodes which are in transmission range but belong to different trees). Each tree in the forest is forms a zone which is assigned a zone ID by running a zone naming algorithm. Furthermore since each node can only belong to a single zone (or tree), then the network can be also seen as a number of non-overlapping zones. The DDR algorithm consists of six phases: preferred neighbor election, forest construction, intra-tree clustering, inter-tree clustering, zone naming and zone partitioning. Each of these phases is executed based on information received in the beacon messages. During the initialization phase, each node starts in the preferred neighbor election phase. The preferred neighbor of a node is a node that has the most number of neighbors. After this, a forest is constructed by connecting each node to their preferred neighbor. Next, the intra-tree clustering algorithm is initiated to determine the structure of the of the zone 5 (or the tree) and to build up the intra-zone routing table. This is then followed by the execution of the inter-tree algorithm to determine the connectivity with the neighboring zones. Each zone is then assigned a name by running the zone naming algorithm and the network is partitioned into the a number of non-overlapping zones. To determine routes, hybrid ad hoc routing protocols (HARP) [23] to work on top of DDR. HARP uses the intra-zone and inter-zone routing tables created by DDR to determine a stable path between the source and the destination. The advantage of DDR is that unlike ZHLS, it does not rely on a static zone map to perform routing and it does not require a root node or a cluster head to coordinate data and control packet transmission between different nodes and zones. However, the nodes that have been selected as preferred neighbors may become performance bottlenecks. This is because; they would transmit more routing and data packets than every other nodes. This means that these nodes would require more recharging as they will have less sleep time than other nodes. Furthermore, if a node is a preferred neighbor for many of its neighbors, many nodes may want to communicate
with it. This means that channel contention would increase around the preferred neighbor, which would result in larger delays experienced by all neighboring nodes before they can reserve the medium. In networks with high traffic, this may also result in significant reduction in throughput, due to packets being dropped when buffers become full.

2.6.3.5 Comparison of Hybrid Routing Protocols

In ZRP, the network is divided into overlapping zones according to the topology knowledge for neighboring nodes of each node. In HARP, the network is divided into non-overlapping zones dynamically by DDR through mapping the network topology to a forest. For each node in HARP, the topology knowledge for neighboring nodes is also needed and the zone level stability is used as a QoS parameter to select more stable route. ZHLS assumes that each node has a location system such as GPS and the geographical information is well known, and the network is geographically divided into non-overlapping zones. The performance of a zone based routing protocol is tightly related to the dynamics and size of the network and parameters for zone construction. However, because zones heavily overlap, ZRP in general will incur more overhead than ZHLS and HARP. The DDR routing protocol have a disadvantage that the preferred neighboring nodes may be overused.

2.7 NEED FOR SECURITY

Several categories of routing protocols as discussed above have been proposed in literature. The major concern addressed in these protocols was to develop a route between a pair of node by simply identifying the intermediate nodes with minimum hop count in an environment where the nodes can be mobile. All these protocols assumed that the nodes in the ad hoc network are benevolent and will faithfully forward the data/control packets of other nodes. While designing these protocols the possible selfish/malicious behavior of the nodes was not taken into consideration. However with the time, the realization was made that there is a need to develop protocols which are able to handle the problem of malicious and selfish nodes. The problem of selfish nodes is
specific to the MANETs wherein a power starved node is not willing to forward the data/ control packets of other nodes. Due to their distributed operations, mobility of nodes and absence of centralized routers the security issues are much complex in ad hoc networks than their other wireless counterparts. The next chapter deals with the security issues that are relevant to the ad hoc networks and the strategies that are being used to tackle these issues.