CHAPTER 3

LITERATURE REVIEW

3.1 INTRODUCTION

There are numerous different routing protocols presently proposed for ad hoc networks. Unfortunately, few have been extensively simulated, whereas very few have been implemented in an actual ad hoc environment. We believe via analyzing some of the unknown and famous routing schemes a wider knowledge of the problem could be developed. Moreover, it could also be used to either extend existing schemes or to develop new routing solutions.

Several ad hoc routing schemes have been suggested in the literature that tries to enhance the various performance metrics. Mostly the performance of these schemes is measured in terms of reachability, which is the fraction of the total nodes that receive the broadcast messages, the saved rebroadcast, that is the fraction of the total nodes that does not rebroadcast the messages, and the latency.

Though a very large number of published information in the area of routing in MANETs is available, only selected papers that have close relevance to our proposals, namely reduction in control overhead in route discovery in dense MANETs, route maintenance adapting to the node mobility in dense scenarios are discussed below.
3.2 CONTROL OVERHEAD REDUCTION IN ROUTE DISCOVERY

The literature pertaining to the control of routing overheads generated during route discovery is discussed in this section.

Zhang and Agrawal (2005) have described a dynamic probabilistic broadcast scheme which is a combination of the probabilistic and counter-based approaches. The scheme is implemented for route discovery process using AODV as base routing protocol. The rebroadcast probability \( P \) is dynamically adjusted according to the value of the local packet counter at each mobile node. Therefore, the value of \( P \) changes when the node moves to a different neighbourhood; for example, in sparse areas, the rebroadcast probability is large compared to denser areas. To suppress the effect of using packet counter as density estimates, two constant values \( d \) and \( d1 \) are used to increment or decrement the rebroadcast probability. However, the critical question is how to determine the optimal value of the constants \( d \) and \( d1 \). The value of a packet counter does not necessarily correspond to the exact number of neighbours from the current host, since some of its neighbours might have suppressed their rebroadcasts according to their local rebroadcast probability. On the other hand, the decision to rebroadcast is made after a random delay.

Zhang and Agarwal (2005 a) have implemented an approach that uses the concept of gossip and Connected Dominating Sets (CDS). But the construct minimal dominating set is not required. Instead of that, the authors categorized mobile hosts into four groups according to their neighbourhood information. For each group, there is a specified value of probability so the nodes with more neighbours are given higher probability, while the nodes with fewer neighbours are given lower probability.
Cartigny and Simplot (2003) proposed an algorithm which combines the advantages of both probabilistic and distance method to privilege the retransmission by nodes that are located at the radio border of the sender. The value of probability $P$ is determined by the information collected from the nodes neighbours and the constant value $K$ which is efficiency parameters to achieve high reachability.

Tseng et al (2003) have proposed an adaptive counter based scheme in which each node dynamically adjusts its threshold value $C$ based on local neighbours information. The fixed threshold $C$ is computed based on a function $C(n)$, where $n$ is the number of neighbours of the node. In this approach the value of $n$ can be achieved through periodic exchange of ‘HELLO’ packets among mobile nodes.

Muneer Bani Yassein et al (2010) presents the Smart Probabilistic Broadcasting (SPB) as a new probabilistic method to improve the performance of existing on-demand routing protocols by reducing the RREQ overhead during the route discovery operation. When hearing a broadcast RREQ packet at node $X$ for the first time, the node compares its neighbour by $\text{avg}_{\text{min}}$, avg and $\text{avg}_{\text{max}}$: if the node has a number of neighbours $n$ less than $\text{avg}_{\text{min}}$, this implies that the node is in a high sparse region. Then, the node rebroadcasts the packet according to probability $p_1$. However, the probability $p_2$ is selected if the number of neighbours $n$ satisfies $\text{avg}_{\text{max}} \leq n < \text{avg}$; this implies that the node is in a medium sparse region. Otherwise, the value of probability $p_3$ is chosen if the node is in a medium density region and the number of neighbours $n$ meets the following condition: $\text{avg} \leq n < \text{avg}_{\text{max}}$. Finally, the value of probability $p_4$ is chosen if the number of neighbours $n$ is $n \geq \text{avg}_{\text{max}}$. This implies that the node is in a high density region. Even though the proposed method arrives at a smart probabilistic value to rebroadcast, it involves few time consuming calculation and require a global knowledge.
about the network. This improvement applied over the pure broadcasting, performs better in terms of reachability and saved rebroadcast, it can be further enhanced to work relatively better in dense networks.

Moghim et al (2002) proposed an improvement to AODV which tries to reduce routing overhead by using source routes in route request and reply packets, so that the routing information of AODV node increases and routing tables are widely used instead of route discovery flooding. However, usage of source routes may cause problems, especially in large network with stale routes, due to loss of data packets and also pollution of routing tables in other nodes. The probability of stale route existence will become less since AODV uses sequence numbers, stale routes can be recognized and removed from the routing entries if they don't need to be updated. As source routes are not used in data packets, usage of stale routes will not pollute other routing tables. So the source routing problems is not encountered in AODV. However, using of source routes even in route request and reply packets can still cause the problem of scalability in huge networks.

Jian Li and Prasant Mohapatra (2003) proposed a Positional Attribute based Next-hop Determination Approach (PANDA) which uses positional attributes such as location and velocity information to determine the rebroadcast delay time, while aiming at finding a longer-lived route with a smaller number of hops to the destination. The basic idea of PANDA is to discriminate neighbouring nodes as good or bad candidates for the next hop on the basis of positional attributes that are of interest. These attributes can be relative distance and link lifetime estimation, and transmission power consumption. The discrimination is done at the downstream node side. Good candidates will use shorter rebroadcast delay, while bad candidates use longer delay such that the good candidates usually go first. Since good candidates usually go before bad ones, a better route in terms of metrics such as hop
count, delay, power consumption, or residual battery, can be found. It is assumed that each mobile node is equipped with Global Positioning System (GPS) so that it is aware of its geographical location and velocity information. To let the downstream nodes learn the previous-hop node’s location and velocity information, they assume that these information is carried with the route-request message in each hop. Upon receipt of a route request packet, an intermediate node can compare its own location and velocity with that of the previous-hop node and then determine the rebroadcast delay according to the algorithm it uses, namely, PANDA-LO (Location Only), PANDA-LV (Location & Velocity), or PANDA-TP (Transmission Power). This complete protocol depends on GPS, which do not go well with dense networks.

Noishiki et al (2005) proposed an Efficient On-demand Route Establishment Methods for Dense Ad-hoc Networks. In the path discovery mechanism, the source node broadcasts an RREQ packet to adjacent nodes. Each node counts the number of RREQ packets and the number of active routes that it accommodates. When the node that receives an RREQ packet is not the destination node, the node rebroadcasts the RREQ packet with the probability determined by the two metrics, $N_i(j)$ - exponential weighted moving average of the number of RREQ packets that the node receives and $L_i$, the ratio of the path accommodation of Node i. The probability of re-broadcasting is determined with the values of above two parameters. The proposed method decreases the number of RREQs per path discovery to around 20 in 120 nodes scenario without lowering the success ratio of path discoveries regardless of the node density. But the reported values are for static dense networks.

Wisitpongphan and Tonguz (2006) proposed three RREQ broadcast techniques, namely, weighted p-persistence, slotted 1-persistence, and slotted p-persistence schemes. They also show through simulations that the proposed
schemes can achieve up to 75% routing overhead reduction while keeping the route acquisition delay close or equal to the conventional flooding scheme. In Weighted p-Persistence broadcasting, the Rule says that upon receiving a packet from node i, node j checks the packet ID and rebroadcasts with probability $p_{ij}$ if it receives the packet for the first time; otherwise, it discards the packet. Denoting the relative distance between node i and j by $D_{ij}$ and the average transmission range by $R$, the forwarding probability, $p_{ij}$, can be calculated on a per-packet basis using the following simple expression,

$$p_{ij} = \frac{D_{ij}}{R} \tag{3.1}$$

Unlike the p-persistence or the gossip based scheme, the weighted p-persistence assigns higher probability to nodes that are located farther away from the broadcaster given that the GPS information is available and accessible from the packet header. Similarly all the three RREQ rebroadcast techniques depends on GPS.

Abdulai et al (2007) proposed a Two-P Scheme which is named as Adaptive Probabilistic Broadcasting, where the nodes are logically categorized into two groups based on their number of neighbours. A node is categorized as a member of group 1, if the number of neighbours $n$, of this node is less than or equal to $n_{avg}$, otherwise it is categorized as a member of group 2. Nodes in group 1 are in sparse regions of the network and as such are assigned high forwarding probability, $p_1$. Nodes in group 2 are located in dense regions and are assigned a low forwarding probability, $p_2$. To identify and categorize mobile nodes based on their local topological characteristics, they have first conducted extensive simulations to determine the average number of 1-hop neighbours of nodes in the various regions of the network. From the simulation results, they determine $n_{avg}$ that denotes the average number of neighbours over all the nodes in the network. This $n_{avg}$ may not be the right parameter to decide on the rebroadcast probability as the node and
link capacity, network topology and other related network characteristics may vary from network to network and time to time.

There have been a few proposals for establishing a virtual backbone infrastructure over which routing takes place. The design of virtual backbone is either based on Connected Dominating Set (CDS) (Wu and Li 2001 and Wu 2002) or Cluster based algorithms (Chiang 1997 and Jiang et al 1999). A CDS is a set of nodes such that every node in the network is either in the set or is the neighbour of a node in the set. In CDS-base routing, only nodes in the dominating set are privileged to forward the RREQ packet. Undoubtedly, the efficiency of the CDS approach depends on the process of finding and maintaining a CDS and the size of the corresponding sub-network. Unfortunately, the problem of finding minimum CDS is shown to be NP-complete.

Wu (2002) has extended dominating-set-based routing in ad hoc wireless networks with unidirectional links. Wu and Li (2001) have suggested a way to run DSR over a CDS. The cluster-based routing protocol (CBRP) for MANET was initially suggested by Jiang et al (1999) and later extended in Chiang (1997) and Ahn et al (2002). In CBRP, the network topology is divided into several disjoint overlapping clusters. Each cluster elects one node as the cluster-head. The cluster-head of each cluster is responsible for forwarding RREQ packets on behalf of its members. Cluster-heads communicate with each other by gateway nodes. A gateway is a node that has two or more cluster-heads as it neighbours.

Lin and Gerla (1997) have discussed various clustering algorithms where a distributed cluster-head selection process is applied. A node is a cluster-head if it has largest ID (or maximum node degree) in its 1-hop neighbourhood. The process continues until all the nodes are covered. Once
the head selection process completes, gateway nodes are selected in order to connect clusters.

The Gossip-based Ad Hoc Routing (Hass et al 2002) proposed a probabilistic route discovery refer to as GOSSIP3 \((p, k, m)\). The parameter \(k\) is the number of hops at which to start forwarding an RREQ packet with probability \(p\). In this protocol, a node that originally did not broadcast a received RREQ packet, but then did not get the packet from at least \(m\) other nodes within some time interval, broadcast the packet with probability \(p = 1\) immediately after the timeout. The authors have incorporated the GOSSIP algorithm in AODV routing protocol and compared their results with only the conventional AODV implementation in the ns2 simulator (http://www.isi.edu/nsnam/ns/nsbuild). However, in dense and heavy loaded networks the improvement of the gossip based routing algorithm is limited.

Gomez et al (2007) proposed a Neighbour-Assisted Route Discovery (NARD) that is intended for medium to large ad hoc networks where traditional flooding is not a practical solution. In NARD, a source node floods a limited portion of the network looking not only for the destination node, but also for information related to other nodes (called neighbours) that were known to be near the destination node recently. Neighbour nodes can be used as anchor points where a second limited flooding takes place in search for the destination node. Because only two limited portions of the network near the source and destination nodes are flooded by control packets, NARD can significantly reduce the signaling overhead of route discovery compared with blind flooding techniques. In NARD a homogeneous density of nodes in the network had been assumed, however, a well-known feature of ad hoc network is that it creates a greater density of nodes in the center of the network and this issue needs to be addressed.
FRESH (Henri Dubois-Ferrie et al 2007) is an algorithm proposed for efficient route discovery in mobile ad hoc networks. In FRESH, nodes keep a record of their most recent encounter times with all nodes. Instead of searching for the destination node directly, the source node searches in a limited portion of the network for any node that has encountered the destination node more recently than the source node did itself. This procedure continues until the destination is finally reached.

Ramachandran et al (2007) proposed a new adaptive cross layer design, where an adaptive decision making of RREQ forwarding is made in accordance with speed of mobile nodes. The distance between the transmission range boundary of the transmitting node and the known maximum transmission range ($d_b$), the safe distance from transmission range boundary $d_s$ that is calculated using the node’s speed information and minimum route life time as specified by the source node $t_1$, are used by the node to make a decision on forwarding the RREQ. The impact on network connectivity due to signal strength threshold and minimum route life-time enforcement is serious in lightly dense networks.

Abdalla M. Hanashi et al (2009) proposed a route discovery algorithm which is a combination of the dynamic probabilistic and knowledge based approaches. It dynamically adjusts the re-broadcast probability $P$ at every mobile node according to the value of the local number of neighbours. The average number of neighbours for the selection of the value of $P$ is calculated by using equation

$$\text{nbr} \cdot (N \cdot 1) \cdot 0.8 \cdot R^2 \frac{\text{A}}{A}$$

Sangman Moh (2009) proposed a link quality aware routing protocol for MANETs that exploits the strong links by forwarding the RREQ
packet with the highest Signal to Noise Ratio (SINR) among the multiple RREQ packets received during route discovery. In case there are RREQ packets within $\delta$ dB, the threshold set for SINR (where $\delta = 1$) from the highest SINR, the first-arrived one among them is chosen to cope with the dynamic behavior of SINR. Any node that has received an RREQ receives successive RREQ packets until the predetermined RREQ waiting time expires; afterwards, RREQ packets for the route discovery are ignored. Compared to the conventional protocols such as AODV, in which only the first-arrived RREQ is forwarded and the others are ignored, the proposed scheme may not have the minimum hop-count route but the one with more number of hops (links). However, the found route is a reliable path with high data rate because it consists of strong links, resulting in high performance as well as robust routing. The performance of the protocol is not appealing in high mobility scenarios.

Yanbin Yang et al (2009) proposed a hybrid routing protocol called improved AODV (IAODV) which integrates the two features: Multipath and Path Accumulation. It is equipped with source routing characteristics, namely the path accumulation technique, which permits nodes listening to routing messages to acquire knowledge about routes to other nodes without initiating route request discovery themselves. This method decreases the required transmissions, but increases the routing packet size. IAODV has another important feature, Multipath, to keep the alternative paths valid. When a data packet is forwarded, the node updates the timer for the used route, thus keeping the path valid, until it breaks. In this protocol, the alternative routes are not used, unless the primary route fails. In order to keep these routes alive, the time period of the route is extended. Extension of the route is considered valid without any data packet using it. Setting this parameter to a value that exceeds the expected lifetime of the path, which should keep the path valid in
order to use it when the primary path fails and to prevent the accumulation of invalid paths in the routing table is a critical issue.

Gwalani (2003) proposed a new modified AODV, which was called AODV-PA. AODV-PA combines the route discovery process of AODV and DSR. When the RREQ or RREP messages are generated or forwarded by the nodes in the network, each node appends its own address on these messages, then the RREQ or RREP messages are sent out again. Each node also updates its routing table with all the information contained in the control messages. AODV-PA sets up a lot of reverse routes. AODV-PA’s method has reduced RREQ’s number in the network, and it has lower delay and lower routing load than original AODV. However, it has set up too many reverse routes, this would bring in DSR’s disadvantage, and results in limited performance.

Yongjun Hu et al (2010) proposed an improvement of the Route Discovery Process in AODV. They suggest appending only the second node’s address on RREQ is appropriate (unlike DSR). After the source node generated RREQ messages, the second node received them and appends its own address on these route discovery messages. So the updated RREQ packet contains the source node and the second node’s information. As these updated RREQ messages are broadcasted, every intermediate node that does not have a route to the destination node forwards these RREQ packets. When an intermediate node receives a RREQ packet, it updates the route to the source node and the second node. New entry is made in the routing table for the source node or the second node, if there has no one existed yet. If a route entry for the source node or the second node does exist, and if the hop count to the source node or the second node is less than the previously known hop count to that node, the routing table entry is updated for that node. The major
problem here is, route table are too big and includes too many routes, when
the ad hoc network’s topology was re RREQ and RREP messages to update
route table.

There has been a few other works that aimed at eliminating the
need for a distance metric for the establishment of spatial ordering in routing.
Examples are Destination-Controlled Source Sequenced Labeled Routing
Protocol (DSLR) (Rangarajan et al 2005) and the Sequence-number Window
Routing protocol (SWR) (Garcia-Luna-Aceves 2004). In DSLR, nodes are
ordered based on labels given by source originated sequence numbers carried
in route requests and these cached labels are used to relay route replies that set
up the necessary successor-predecessor relationship. In SWR, nodes are given
some leeway in choosing their sequence numbers for a given destination
within a range of values (a window) dictated by the sequence numbers of its
neighbours in such a way that the sequence numbers assigned to nodes
become larger as the destination is approached and thus mirror the spatial
ordering of hop counts. The advantage of this approach is that it tends to
allow more relays to provide intermediate replies than previous distance-
vector routing protocols.

Sze-Yao Ni (1999) proposed a fixed probabilistic scheme is the
first probabilistic approach and considered as the base for all later dynamic
probabilistic schemes. In which every node receives a broadcast message for
the first time, rebroadcasts it to all nodes in the network with a certain value
of probability P, regardless of the density level of current node. The
researcher Sasson and Schiper (2003) has shown that, the best value of P in
terms of high reachability and saved rebroadcast is approximately equal to
0.07%.

The core of the proposed schemes is the Heading-Direction Angle
Routing Protocol (HARP) (Navid Nikaein 2001) so called because it utilizes
the direction of information of nodes in the network. Heading direction information can be obtained from the node’s own instruments and sensors, such as a compass, which delivers the Heading Direction Angle (HDA) of the mobile device relative to magnetic north. This protocol is used to reduce routing overheads and to increase the lifetime of links between communication nodes. It has been assumed that every node can exchange its information frequently with its neighbours. In this protocol every node classifies its neighbouring nodes into eight different zones according to the heading direction of those neighbours. In theory, the nodes are categorized within at least one of the eight zone ranges, regardless of their location. This protocol is based on an on-demand routing technique. The route request packet is transmitted from a Source Node (ND) to one of the neighbouring nodes (NB) that has an angular heading direction similar or near to the HDA of NB.

Guangxue Wang and Kai Liu (2009) proposed a protocol to avoid excessive energy consumption and high computation complexity; metrics for route decision are piggybacked in the header of route packets such as RREQ and RREP in AODV protocol. The route selection metrics consist of interference level and energy level to evade areas with excessive interference and lack of energy. Due to energy level integrated into the route selection scheme, the XLR protocol has evaded areas lack of energy to transmit traffic packets. Hence, energy of the entire network can be consumed symmetrically. Therefore, the network can avoid draining off their power prematurely and the network lifetime can be prolonged effectively.

Shoba et al (2009) have combined the flooding with relay routing in alternate phases to decrease the amount of overhead in routing. The relay node for carrying out selective flooding is determined on the basis of random mobility of the nodes. The protocol alternates between flooding and relay
routing schemes during Route Discovery and Route Maintenance phase. The data collected from the neighbouring nodes about their mobility during flooding phase is used to choose the relay nodes to which the requests must be sent in the relaying cycle. The two cycles used in the transmission of Route Requests are:

i) **Flooding**: Broadcasting of Route Requests to all the nodes in the network.

ii) **Relaying**: The mobility of the nodes, which have been collected in the flooding cycle, is used in this phase. The mobility of all the one hop neighbours is compared with the mobility of the current node. Only if the mobility of the neighbouring node is greater than the mobility of the current node the Route Request is sent to the neighbouring node. A threshold of $M_{th}$ is set in all the nodes. If $M$ is the mobility of the current node, $M_n$ is the mobility of neighbour node then Route Request is sent to the neighbouring node only if $(M_n - M) > M_{th}$. The number of requests is limited by using the mobility criterion. In this way the route request being flooded to the entire network is prevented.

### 3.3 ROUTE MAINTENANCE

Some of the researchers argue that the underlying problem for packet loss is the interaction between routing and MAC protocols, since apparent link failures are largely a result of network congestion rather than physical link breakages (Ashraf et al 2008, Nahm et al 2008). These failures are frequent but transient, and therefore should be treated differently to longer-term link failures. One way in which this can be achieved is to introduce a Bulk Trigger (BT) policy, which increases the link failure
threshold by allowing small amount of packet losses before announcing link failure. Hence the route re-discovery process only takes place after a certain number of consecutive packet losses. The drawback of BT policy is that the network assumes a fixed link failure threshold on all stations, whereas the collisions mostly occur at specific nodes (i.e. bottleneck nodes). The choice of link failure threshold is thus a tradeoff between network dynamism and stability. On one hand, if the threshold is too low, it may be insufficient to accommodate the level of network congestion. If the threshold is too high, the network becomes stuck in a static configuration and is unable to react quickly to physical topology changes (e.g. the physical loss or addition of a node). Therefore, given the distributed nature of multi-hop networks, link failure thresholds should be determined by each station based on the local network conditions.

More recent studies have offered an alternative explanation, and raised the issue of interaction between reactive routing protocols and the underlying MAC protocols. Ng and Liew (2004) demonstrated that the instability problem is not restricted to TCP - it also occurs in UDP traffic. Their work looked at IEEE 802.11 ad hoc networks using the AODV reactive routing protocol, and showed that the large throughput fluctuation is the result of frequent route re-discovery processes triggered by the loss of data packets. The ongoing attempts at data transmission are blocked until the route is recovered, resulting in a sudden drop in throughput. Hence, the problem is redefined as a “re-routing instability problem”. Other authors re-examined the re-routing instability problem with TCP traffic (Ashraf et al 2008, Nahm et al 2008) and confirmed that the excessive data traffic disrupts the routing dynamic of reactive routing protocols, leading to network instability.

Moreover, the interaction between reactive routing and MAC protocols creates potential “instability loops” in the network, particularly
under high traffic load. Reactive routing protocols rely heavily upon broadcast transmission to collect and distribute routing information. However, the basic 802.11 DCF only offers a minimal service quality for broadcast transmissions, as the stations do not acknowledge, received broadcast frames, nor do they have the ability to re-transmit in the event of packet loss (IEEE std. 2007). Therefore, when competing against data traffic (which typically is dominated by unicast data), the routing packets are prone to loss (Wang et al 2008, Oliveira 2009). Hence, extended involuntary disconnections occur in the network. In response to these route breakages, reactive routing protocols generate yet more routing (broadcast) packets to flood the network, further exacerbating the problem.

Jerry Chun-Ping Wang et al (2009) discussed the instability which may arise when reactive routing protocols interact with the IEEE 802.11 MAC protocol is investigated. In particular, several erratic behaviours of AODV routing protocol in a congested IEEE 802.11 ad hoc network are demonstrated. A cross-layer solution is proposed based on an Adaptive Bulk Trigger policy and a Dynamic Window Selection scheme. The authors proposed a solution to enhance the link-failure tolerability of reactive routing protocols and provide prioritized channel access based on routing demands

The Signal Stability Adaptive protocol (SSA) (Dibe et al 1997) tries to discover longer-lived routes based on signal strength and location stability. Each link is differentiated as strong and weak according to the average signal strength at which packets are heard. The location stability criteria further biases the protocol toward choosing a path which has existed for a longer period of time. Beacons are sent periodically by each host for its neighbours to measure these criteria. Each host maintains a signal stability table. On an active link failure, SSA broadcast a “local” ROUTE REQ packet with a small hop limit, in hope of rebuilding the route with little effort. Also,
the initiator of ROUTE REQ should set up a timer, so that if the broken route cannot be rebuilt within the timeout period, it can send a normal route error packet to the source node so that a “global” ROUTE REQ packet can be sent.

ABR (Toh 1997) uses pilot signal to determine link stability. Every node sends pilot signal periodically. When a node receives this pilot signal from its neighbour nodes, it records pilot signal received. If it receives these pilot signals from a neighbour continuously and the number of continuous pilot signals beyond certain thresholds hold, it considers the link between them as stable link. Route search in ABR is different from SSA. In ABR, it searches all possible routes to find a route that contains more strong links.

ASBM is proposed in by Lim et al (2002) by enhancing SBM. Link stability in SBM is decided only with signal strength. In ASBM, Differentiated Signal Strength is added (DSS) as a parameter. DSS indicates the signal strength is going stronger or weaker. If it becomes stronger, it means that two nodes will be closer and the link between them would have longer lifetime. In SSA, only a link, the signal strength of that exceeds certain limit, are considered as a stable link. Whereas ASBM, considers both strong signal links and weak signal links that are coming closer as stable links.

Shin-Lin et al (2000) implemented route maintenance with a local recovery where they took the distance as the only metric to find a substitute node without checking its physical condition. Thus, it is difficult to guarantee the robustness of the newly recovered path. Tauchi et al (2005) proposed a route maintenance algorithm to avoid route breaks based on AODV. In their algorithm, each node broadcasts a one-hop Hello packet periodically for local connectivity. Whenever a node fails to receive any reply from a neighbour for a set period of time, it assumes a link break to that neighbour. And route-maintenance is realized differently according to the position of the node that has detected a link break to its neighbours. If it is near the destination, this
node will try to establish a path form itself to the destination node with route discovery. Otherwise, this node will report the link break to the source node, where traditional route maintenance will be triggered.

Chang et al (2002) presented an active routing maintenance protocol to prevent the current route from disconnecting. They implement their protocol on SSA and employ the signal strength as the metric to select suitable substitute nodes. After an unreliable link is detected, the two end nodes shall have a local session to decide which will be replaced. And after a substitute node is sought, the node to be replaced will broadcast packets about the route change to its neighbours.

Fei Dai et al (2005) proposed a proactive route maintenance, in which routing information is disseminated along active routes and advertised only by active nodes that forward data packets. Alternative paths are dynamically discovered and maintained by active nodes and their 1-hop neighbours.

Goff et al (2001) checked the link reliability with signal strength. And each time a link failure is detected, instead of repairing this link, they will try to find an alternative path for the one including this broken link. Qi et al (2005) proposed a routing protocol, taking the node power as the major consideration when selecting paths. In their protocol, either link failure or a too low value in node power may trigger route maintenance. Also, they check the path quality periodically for unreliable links. Sathyaraj et al (2005) tried to implement their route maintenance by making prediction on each node. The link expiration time is computed to see whether the route maintenance is necessary.

Dong Shi et al (2007) related the link reliability to the physical condition of the two end nodes and we provide a new method to evaluate the
link reliability. Based on that, they proposed a new Link Reliability-aware Route Maintenance mechanism (LRRM). LRRM checks the link reliability for every hop and actively. If necessary, LRRM will replace those week nodes with more robust ones, which consequently keep the link healthy and contribute in traffic balancing too. This local recovery, with only a few nodes involved, is always triggered before a path breaks in order to avoid costly path reestablishment.

Sung-Ju Lee and Mario Gerla (2001) proposed Dynamic Load-Aware Routing (DLAR) protocol that uses the routing load of the intermediate nodes as the main route selection criteria. In the route construction phase, each intermediate node records in the control packet the number of packets queued at the interface and the destination uses that information when selecting the route. Three different route selection algorithms were described. Scheme 1 uses the total number of packets buffered at the intermediate nodes and scheme 2 uses the average number of queued packets at each node. Scheme 3 defines a load threshold and selects the route that has the least number of intermediate nodes that have packets buffered more than the threshold value. To avoid producing bottlenecks and to use the most up-to-date route information when discovering routes, DLAR does not allow intermediate nodes to reply from cache. DLAR periodically monitors the congestion status of active data sessions and dynamically reconfigures the routes that are being congested. Using the least-loaded routes helps balance the load of the network nodes and utilize the network resources efficiently.

Neighbour Stability Routing (NSR) (Chen and Lee 2005) algorithm selects the most historically and accumulatively stable mobile nodes to form a path between the source node and destination node. The relative stability is then propagated from the collective data by all the nodes along a path. The
cumulative collective data, or stability factor, reflects the historical neighbourhood stability among neighbours. When a node or segment on the path is down, NSR will dynamically find an alternative most stable path. Yang and Huang (2008), Taleb et al (2007) and Wang et al (2007) used Link Expiration Time (LET) (William and Mario 1999) to assess the stability of link which is based on GPS information, and the route expiration time is the minimum LET on the route. These routing protocols integrate the evaluation of LET into on-demand routing algorithms, such as DSR or AODV, for discovering stable route. For every node can obtain GPS information by itself, no periodic message exchange needed, which can reduce much control overhead.

Sakhaee et al (2008) proposed a self-adaptive and mobility-aware path selection in mobile ad-hoc networks. To be aware of a mobility of node, Doppler value is calculated based on the Doppler shift which can be obtained through the forwarding of route request packet like DSR for assessing the stability. The protocol proposed by Wu et al (2008) uses the Newton interpolation polynomial to gain all the received signal strength within the whole sample domain based on only several sampling points, and then finds appropriate reference points to complete the whole prediction of stability.

Xi Hu et al (2011) introduced a routing algorithm which enhances the stability and the continuity of communications in MANETs. Communication stability is ensured by choosing the most stable route which bases on the computation of the Link Expiration Time (LET). The route with the longest LET is considered as the most stable. Then the reactive calculation of LET, asynchronous mobility information and LET update, and alternative route pre-discovery based on the Critical LET (CL) zone are proposed to further enhance the adaptability of stability-oriented routing to the dynamic of network and ensure the continuity of communications.
Rei-Heng Cheng (2010) proposed integrating the Dynamic Link Break Avoidance (DLBA) and Dynamic Path Shortening (DPS) mechanisms into a modified protocol by developing a pair of parameters to determine the timing for activating DLBA or DPS so that the two algorithms can work cooperatively and complementarily together. The DLBA mechanism operates by monitoring the quality of communication links and inserts an intermediate node in between the two end nodes of some breaking link. The DPS mechanism operates by recognizing potential redundant intermediate nodes and either removes or replaces the nodes from the established path. There are two types of path shortening scenarios. In type I scenario, some redundant nodes are directly removed from the active route and in type II scenario, some redundant nodes in the active route is replaced with a node that is not on the active route.

Sengul and Kravets (2006) proposed a local recovery algorithm and combined that with the DSR protocol. In case of link breakage and no alternative route in the route cache, the protocol issues a local recovery request looking for potential node for path detour. If any packet successfully reaches the destination through the newly discovered route, the destination node would then notify the source node for such a route change by returning a notification message. Castañeda et al (2002) proposed two heuristics that utilize prior routing histories to localize the query floods to a limited region of the old routes and the dynamically collected information to repair the broken route.

Yu et al (2007) proposed a dynamic route repairing protocol that repairs a broken route using information provided by nodes overhearing the main route communication. When links go down, their protocol intelligently replaces these failed links or nodes with backup ones that are adjacent to the main route. Soliman and Al-Otaibi (2009) proposed an algorithm to enhance
the Local Repair (LR) phase of AODV by using a preemptive mechanism to detect potential link failures and to find in advance some alternative links. By eliminating the use of regular and costly control messages such as RREQ, RREP, and RERR, their protocol is able to reduce substantially the control overhead generated by the repair process. A Neighbour Activity Table (NAT) that records the information about the nodes located two-hop away is maintained to reduce the overheads. The maintenance of such a NAT depends on nodes of an active route monitoring activities of the other active routes, which make their protocol sensitive to the number of conversations. To be more specific, in cases of fewer conversation-pairs, the NAT table may not be able to populate enough alternative links for future path detouring. On the other hand, the protocol may perform well as compared to similar protocols in case of higher loading condition.

Crisostomo et al (2004) proposed to use nodes’ position and mobility information, and apply such information to predict vulnerable links and potential backup nodes. Special hardware like GPS is required to assist the implementation of such protocol. Srinath et al (2002), on the other hand, proposed to maintain the so called neighbour power list and power difference table so as to initiate backup route switching mechanism when signal strength between two adjacent nodes in an active link is weakening. However, maintaining two tables requires periodical HELLO messages interchange, and thus imposes a considerable burden on the network.

Tsai et al (2006) proposed keeping track of the Signal to Noise Ratio (SNR) of links so that a new route request would be initiated once the SNR is below some pre-defined threshold to reduce the possibility of link breaking. However, in mistaken prediction cases, their method may result in additional route reestablishment. Even if the prediction is correct, local route repair may well be sufficient instead of complete new route construction. Two
independently proposed protocols Boukerche and Zhang (2004) and Goff et al (2003) make use of link status information of routes that is piggybacked in the RREP packets to notify the source so that it can plan for the right timing to route reconstruction. In the meantime, nodes on the route also need to perform link quality monitoring and confirm that with PING-PONG process for possible link breakage. As soon as a potential link breakage is identified, the source will be warned so that it can preemptively initiate route reconstruction procedure. Note that there always exists the possibility of false identifying link breakage which may lead to flooding.

3.4 ROUTE MAINTENANCE WITH BACKUP ROUTES

Here we discuss few of the protocols which use the backup routing strategy.

Wang et al (2005) proposed Dynamic Backup Route Routing Protocol (DBR$^2$P) that enhances DSR by adding route backup in case of a link/route failure. DBR$^2$P “armed” intermediate nodes with alternative/backup path to reach destination node. DBR$^2$P includes three phases, route discovery, backup node setup and route maintenance. When a destination node receives route request messages (in DBR$^2$P it is called RD-request), it will build backup setup packet (BS-Packet). BS-Packet contains backup routes to reach the destination node. A node (including the source node) which has more than one possible route to a destination node is called backup node. BS-Packet is sent by destination node to the backup nodes. When the backup nodes receive BS-Packet, they will store the routes to their local cache. In case of route/link failure, the detector will try to replace the routes with backup routes, taken from the backup cache. This replacement is done on the spot. If that node does not have backup routes, it will send route error packet to their upstream. Then, the upstream nodes will check whether they have backup routes or not,
if a backup node is available they will replace the routes and continue sending the data packets, otherwise a route error packet is sent to the upstream.

Park and Van Voorst (2001) proposed Anticipated Route Maintenance (ARM). ARM is a distributed algorithm that anticipates route failure and performs preventative route maintenance using location information to increase a route lifespan by expanding the routes. The term route lifespan refers to the amount of time the route can function without failing. ARM determines the position when a node becomes unsafe by calculating the Time-to-Failure (TTF). If TTF is less than or equals to some pre-defined value (T), then the link is called unsafe. ARM itself depends on another MANET routing protocol to perform route discovery or searching path. ARM can be embedded into any reactive routing protocol.

The combination of DSR and ARM can be seen in the work done by Al-Shurman et al (2004). Unfortunately, computing TTF value needs a complex calculation, because each node needs to know their next-hop position, as well as the velocity and the angle of the movement. This means that each node have to plan their movement precisely before informing the other nodes of its current location and plans to move somewhere. BS-Packet contains backup routes to reach the destination node. A node (including the source node) which has more than one possible route to a destination node is called backup node. BS-Packet is sent by destination node to the backup nodes. When the backup nodes receive BS-Packet, they will store the routes to their local cache. In case of route/link failure, the detector will try to replace the routes with backup routes, taken from the backup cache. This replacement is done on the spot. If that node does not have backup routes, it will send route error packet to their upstream. Then, the upstream nodes will check whether they have backup routes or not, if a backup node is available.
they will replace the routes and continue sending the data packets, otherwise a route error packet is sent to the upstream.

AOMDV is the enhanced AODV that each node has multiple routes to a destination node (Marina and Das 2002). When a route breaks down, the node alternates to another route if there are valid routes to the destination node in the routing table. When all routes break down, it re-establishes new routes. But this may not be an efficient protocol because the route established at a certain time T is not always valid at time T + t in ad-hoc networks.

ABR (Toh et al 1997) uses partial route discovery through backtracking pivoting nodes. In WAR (Aron and Gupta 1999) after the witness node detects a broken link, it rebroadcasts the data packet until the packet reaches its original route. In NSR (Spohn and García-Luna-Aceves 2001) each relay node uses periodic topology updates to maintain its two-hop neighbourhood information for route repairing. In CHAMP (Valera et al 2002), packets are cached at relay nodes for salvaging when a route error packet is received from downstream nodes. However, these protocols carry their own respective limitations. For example, ABR can cause significant latency if its pivoting nodes fail. WAR requires witness nodes to be dedicated to each link and causes new overhead for broadcasting undeliverable packets. NSR may overload the links with periodical link-state updates, and CHAMP requires additional storage for packet recovery. Fault-resilient routing protocols are demonstrated by using redundant routers in wired networks (Zhao et al 2001). For MANET, multipath routing (Nasipuri and Das 1999) (Lee and Gerla 2001) can provide limited fault-resilience with low node mobility.

Tsung-Chuan (2010) had proposed an AODV-based backup routing scheme for MANETs. The proposed scheme utilizes 2-hop neighbour knowledge to establish backup paths during the route discovery phase and
maintain updated knowledge of such links. These backup paths are geographically close to the primary path that can repair disrupted links locally without activating a route re-discovery procedure. Additionally, the proposed scheme selects the shortest path containing the largest number of backup paths to provide efficient recovery from route failure and maintain an adequate routing length.

Lai et al (2007) modified the AODV-BR algorithm (Lee and Gerla 2000) and proposed two improved algorithms, AODV-ABR and AODV-ABL, to enhance the adaptability of AODV protocol to the topology-changing environment. Overhearing function also needs to be turned on so that nodes around the main route may be able to learn whichever nodes on the main routes those are within their power range. In case a link A-B were broken, the upstream node A would send a one-hop Backup Route ReQuest (BRRQ) and await for neighbouring node’s Backup Route RePly (BRRP) regarding their hop counts to the destination. The upstream node can then select an appropriate detouring node with the least hop count. However, the protocols do not work preemptively. In addition, the average hop count of routes in the two protocols may add up as the protocols proceed.

Chung et al (2001) proposed the Ad hoc Backup node setup Routing Protocol (ABRP) that is similar to the DSR. ABRP saves backup route information in certain on-the-route node. Similarly, ABRP does not update its backup route information to reflect the network topology change. Agarwal and Jain (2004) proposed a modified AODV protocol in which a node records information from its two upstream unnecessary route reconstruction and may possibly cause the broadcast storm in heavy traffic condition nodes. After some evaluation, one of the upstream nodes will be included in the main route with the other being used for the backup route.
Chen and Lee (2005) proposed a 2HBR protocol that extends the idea in (Lee and Gerla 2000) further by including nodes that are two hops from the main route as the potential backup nodes. However, it is questionable whether the pre-established backup routes would ever be used, not to mention that these backup routes may not be available when they are actually needed due to the mobile nature of nodes in ad hoc networks.

3.5 CONCLUSION

The research papers that presented various ideas on path finding and maintaining routes in MANETs have been studied exhaustively. The lessons learnt were the following:

i. A blind flooding and other proposed controlled flooding algorithms result in excessive redundancy, contention, and collision when the MANET is massively dense. These may lead to lower reachability (to the potential receiving hosts) and longer latency (for the broadcast to complete).

ii. The protocol must adapt seamlessly to the changing network conditions in order to achieve better performance.

The following chapter discusses the new protocol that is proposed to overcome the difficulties experienced on using the traditional protocols in Dense MANETS.