We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



21

QoS Routing Solutions for Mobile Ad Hoc Network

Jiwa Abdullah
University Tun Hussein Onn Malaysia, Malaysia

1. Introduction

For the past decade, the field of mobile ad hoc networks (MANETs) [1] has been accepted as a legitimate area of research. It avoids the need for base station infrastructure by being able to self-organized and self-configuring. Hence it provides a spontaneous and yet robust wireless communication systems. Initially, MANETs researchers were focused mainly on designing distributed and dynamic communications protocols for shared channel and for route discovery. It offers best-effort protocols to ensure optimum network operation in an unpredictable wireless environment. Additionally it maintained a network topology view and routes in the face of disruption of links, failure mobile devices and short residual connectivity time. Nonetheless one could not actually experienced any successful practical implementation of MANETs in the real world. Entertainment and some other multimedia services usually made an impact on any technological breakthrough but the potential of MANETs are not truly realized. They must be able to deliver such services, for which best-effort protocols are not adequate. This is because multimedia applications often have stringent delay and reliability sensitive service requirements. Subsequently, the research focus has shifted from best-effort services to the provision of better defined QoS in MANET. QoS routing protocols then play an essential role in a QoS mechanism, since it is their task to find which nodes, if any, can serve an application's requirements. It plays a major part in session admission control (SAC), due to its dependence on the route discovery that support the requested QoS. Alternatively, some QoS routing solutions may not attempt to serve applications' requirements directly, rather to improve QoS under a particular metrics. Most of the solutions proposed in the literature, until now have focused on providing QoS based on throughput and delay. Throughput is the most common metric used. This is due to its character as the lowest common denominator requirement. It is noted that, most voice or video applications require some level of guaranteed throughput in addition to their other constraints. However, many other useful metrics are also used to quantify QoS. In this work we cover most of them and provide examples of their use. The remainder of this article is structured as follows. In Section 2 we discuss related work in terms of QoS routing surveys and summarize their main points. Section 3 describes the problem statement of QoS routing. Section 4 explains in detail the existing heuristics of QoS routing protocol. Section 5 describes a brief review of the challenges posed by the provision of QoS on the MANET environment. Section 6 presents the factors that need to be considered in designing a viable QoS routing protocol, QoS routing protocol performance, the network resources consumable

by applications, and some of the trade-offs involved in protocol design. In Section 7 we describes some methods of classifying QoS routing solutions. Section 8 provides some examples of QoS routing protocols that rely on contention-free MAC. Section 9 describes the solutions for operating with a contended MAC. Finally, methods that do not rely on any specific kind of MAC are presented in Section 10. Under each section, we group protocols into different types of approaches, although for some approaches, only one example is provided. In Section 11 we described the QoS routing protocol utilising computational intelligence approach. We discuss our finding and the observed trends in the field of QoS routing in Section 12. Future works are described in Section 13 and finally Section 14 summarizes the chapter.

2. Related work

There are several overviews and surveys of QoS routing issues and solutions. Chen *et al* , provided a fairly comprehensive overview of QoS in networking [2]. Chakrabarti *et al* [3] summarized the important QoS-related issues in MANETs and subsequently produced an updated version [4]. A survey by Zhang *et al* [5], highlighted several significant points: (1) Most of the algorithmic problems, such as multi-constraint routing, have been shown to be NP-complete; (2) QoS and indeed Best Effort routing can only be successfully achieved if the network is stable. This means that the nodes are not moving faster than routing updates can propagate; (3) Techniques of QoS provisioning differ when the network size becomes very large, since QoS state updating mechanism takes longer time to propagate to distant nodes; (4) There is a trade-off between QoS provisioning and minimisation of power utilization. Areas of future work were also identified: (1) Admission control policies and protocols require further attention; (2) QoS robustness; (3) QoS routing protocol security against, for example, denial-of-service attacks. (4) The combination of security and QoS provisioning; (5) Study of QoS preservation under failure conditions; (6) QoS support for multicast applications. Mohapatra et al [56] provides a survey of issues in supporting QoS in MANETs. The paper considered a layered view of QoS provisioning in MANETs. In addition to the basic issues in QoS, it describes the efforts on QoS support at each of the layers, starting from the physical and going up to the application layer.

Al-Karaki et al [6], provided a detailed overview in the field of QoS routing. The following aspects were highlighted: (1) Accommodating multiple classes of traffic, but still allowing the propagation of lower-class traffic with the inclusion of preemptive scheduling; (2) Ensuring QoS guarantees under various failure conditions; (3) usage of localization devices such as GPS; (4) Prioritization of control packets; (5) Using realistic mobility models; (6) Quantifying the impact of cross-layer integration; (7) MANETs and Internet interoperability; (8) Secured QoS routing protocol, preventing malicious and harmful retransmission; (9) Network partitioning in the context of QoS routing; (10) Node heterogeneity in terms of their capacity and capabilities.

In this paper we focus on the essentials of QoS routing, which is the discovery of routes servicing data sessions and admission control. Al-Karaki *et al* [6] also discussed various QoS routing solutions which falls into the following categories: (1) flat, which means that all nodes perform an equal role; (2) hierarchical, where some nodes are group heads; (3) position-based protocol where location information is made available, and (4) power-aware in which battery usage and residual charge are considered. Reddy *et. al* [7] provide a thorough overview of the more widely accepted MAC and routing solutions for providing

better QoS. The author provides two varieties of QoS solution, one is based on the QoS routing employed, while the other one is based on the layer at which they operate in the network protocol stack. The QoS routing employed is further classified into the mechanism which is based on (1) interaction between routing protocol. (2) QoS provisioning mechanism; (3) interaction between network and MAC layers; (4) on the routing information update mechanism employed that is on demand, table driven or hybrid.

3. QoS routing protocol: the problem descriptions

3.1 Goal of QoS routing

The main goal of QoS routing is to select, based on information about the state of the network, the path that is most suitable according to traffic requirements [8]. The maximization of network resource utilization is also an important goal of QoS routing. Hence, the QoS routing schemes must present solutions for metrics distribution mechanisms and path selection algorithm. Generally, QoS routing protocols search for routes with sufficient resources in order to satisfy the QoS requirements of a flow. The information regarding the availability of resources is managed by a resource management function which QoS routing protocol in its search for QoS feasible paths. The QoS routing protocol should find paths that consume minimum resources according to the relevant QoS metrics. Finding an optimal path with multiple constraints may be an NP-complete problem if it involves two or more metrics [9]. For a successful QoS routing operation, the topology information can be maintained at the nodes. The topology information needs to be refreshed frequently by sending link state update messages, which consume precious network resources such as bandwidth and battery power. Otherwise, the dynamically varying network topology may cause the topology information to become imprecise. This trade-off affects the performance of the QoS routing protocol. As path breaks occur frequently in MANET compared to wired networks where a link goes down very rarely, the path satisfying the QoS requirements needs to be recomputed every time the current path gets broken. The QoS routing protocol should respond quickly in case of path breaks and recompute the broken path or bypass the broken link without degrading the level of QoS. In the literature, numerous routing protocols have been proposed for finding QoS paths.

3.2 Mechanism for metrics distribution

The state of the network can be represented by a set of metrics, which includes the available bandwidth, delay, jitter, and congestion level. Traffic requirements can be expressed in several ways, depending on the methodology used for traffic characterization. For instance, in the Integrated Services framework, this can be done using the QoS parameters associated with each data flow during resource reservation [10][11]. In the Differentiated Services framework, traffic requirements are associated with each traffic class [12]. The information about the state of the network must be distributed, and kept updated, to all or some routers in the network. The distribution must be done more frequently than in traditional routing, to reflect the dynamic behavior of the network. However, if this frequency is too high, it will induce too much bandwidth consumption, and it is thus undesirable. In these situations, it is advisable to achieve a compromise between the desired actuality of the state information and the overhead that this introduces. Some approaches to this problem include the distribution of quantified values, instead of instantaneous ones. Associated with this value

quantification, triggers may be used to control the emission of updates and timers to force a minimum interval between the emission of updates. [13]. A problem that relates to the frequency of the distribution of information pertaining to the state of the network is the inaccuracy that a lower frequency can introduce. Other sources of inaccuracy are the propagation delay of routing messages in large networks, the utilization of estimates, the impact of the metrics measurement mechanism used and information aggregation in hierarchical systems. The study of the impact of routing information inaccuracy on the performance of communication systems and the definition of the mechanisms to overcome its problems has been the subject of several research projects [14] [15] [16] [17].

3.3 Path selection algorithm

The path selection algorithm has a degree of complexity that depends on various factors. Since applications generate traffic with very diverse requirements in terms of QoS, the path selection algorithm must select paths that satisfy a set of restrictions. This is however, a problem with high computational complexity, depending on the rule of metrics composition. The value of a metric along a path, based on its value in each hop, depends on the nature of the metric. There is additive, multiplicative and concave metrics. The rule for additive metrics composition is that the value of this metric over a path is the sum of the values of each hop. Delay and number of hops are examples of additive metrics. With a multiplicative metric, the value of the metric over a path is the product of its values in each hop, as it is the case of losses. The value of a concave metric over a path corresponds to the minimum value observed in all hops of that path. Bandwidth is a common example of a concave metric. In these equations, $m(l_i)$ is the value of a metric on link l, and $m(p)$ is the total metric value of the path composed of links l_1 to l_n. The problem of QoS routing when using two additive or multiplicative metrics, or one additive and one multiplicative metrics is a NP-complete problem [9]. This poses a challenge that must be addressed in order to conceive QoS routing strategies that are efficient and scalable.

4. Existing heuristics for QoS routing

The heuristics of QoS routing can be characterized by several aspects, including the metrics, type of path selection algorithm, instant of application of the path selection algorithm and localisation of the routing decision. In this paper we use as the main characterisation feature, the metrics for path selection, because it is an attribute that determines most of the other aspects. Bandwidth is widely used as a metric for QoS routing, alone or associated with other metrics, such as delay [10][9] and number of hops [13]. It is usually coupled with systems where traffic differentiation is done at the flow level, with the specification of path QoS parameters.

4.1 Metric ordering

Metric ordering requires the identification of highest priority metric and then compute the best paths for it. Where more than one best paths, second metric is invoked, to choose the best path. It is a kind of shortest-widest path and widest-shortest path algorithms. In shortest-widest path algorithms, paths with maximum available bandwidth is located. If there are paths of the same available bandwidth, it would then select the path with shortest

number of hops. These algorithms support load balancing, showing top performance with low network loading. However, this approach damages best-effort traffic performance because it contributes to resource consumption. Another shortest-widest path algorithm uses propagation delay, as the second metric. Wang *et al* presented the related path computation algorithms which are based on distance-vector and link-state [9].

4.2 Sequential filtering

In sequential filtering, the network links that do not have enough available bandwidth are excluded from the network graph. The shortest path is then computed. For on-demand path computation, bandwidth value is obtained on request through resource reservation protocol. If paths are pre-computed, bandwidth ranges must be established. On-demand path computation requires parameter specification. For path pre-computation it is necessary to compute and store several pre-computed paths that satisfy the defined range of bandwidth values. Sequential filtering can also be used to find paths subject to more than two constraints. An example is the cheapest-shortest-feasible path algorithm presented by A. Shaikh *et al*[15]. This source routing algorithm aims at finding feasible paths according to a bandwidth constraint, minimizing simultaneously cost and resource consumption.

4.3 Scheduling disciplines

The complexity of route selection algorithms can be overcome by using the relationships among QoS parameters, determined by the nature of scheduling disciplines. In particular, when a Weighted Fair Queuing (WFQ) scheduling mechanism is used, it is possible to find a route, in polynomial time, subject to constraints of delay, jitter and bandwidth [18]. WFQ is a rate proportional scheduling discipline that isolates each guaranteed session from the others. It also has delay bounds that can be mathematically determined.

4.4 Admission control

In some QoS architectures, the admission of new flows in the network is subject to a mechanism of admission control. This mechanism interacts closely with routing. Typically, the routing module can produce information about the network state which contribute to admission control decision [12]. Admission control and QoS routing are connected to resource reservation. The resource reservation protocol can express the flow QoS requirements that are used by the QoS routing protocol to compute suitable paths. The resource reservation protocol can then proceed to flow establishment on the paths produced by the QoS routing algorithm. If this establishment is successful, the flow is accepted; otherwise it is rejected. Rampal in reference [19] presented path computation algorithms that considered QoS requirements and admission control restrictions of multimedia traffic. These algorithms used information associated with the admission control module, the minimum delay and probability of rejection by the admission control module. This information is used for pruning from the network graph the links that do not satisfy admission control restrictions. The remaining graph is then presented to the routing algorithm.

4.5 Control theory approach

Control theory approach offers a successful track record in physical process control. It gives somewhat a performance guarantees in the face of uncertainty, non-linearities and time-variations system. It does not require accurate system models and utilise feedback

mechanism. Performance of software services is governed by queuing dynamics which may be expressed by differential equations akin to those of physical systems. Bao Li *et al* [20] presented a model for QoS mechanism employing feedback control theory. The ideal objectives of the model consist of two fold. First, it could accommodate variable QoS requirements, in a timely fashion. Second, it could accommodate concurrency of resource access among multiple applications sharing the same pool of available resources.

4.6 Computational intelligence approach

QoS routing is a key MANET function for the transmission and distribution of multimedia services. It has two objectives; (1) finding routes that satisfy QoS constraints and, (2) making efficient use of limited resources. The complexity involved in the networks may require the considerations of multiple objectives at the same time, for the routing decision process. In this paper we also introduced the use of Fuzzy Logic [21] and Genetic Algorithm based QoS routing for MANET [22].

5. Challenges to QoS routing mechanism

The following is a summary of the major challenges in providing QoS routing mechanism for MANETs.

5.1 Unreliable wireless channel

The wireless channel is prone to bit errors due to interference from other transmissions, thermal noise, shadowing and multi-path fading effects [23]. This makes it impossible to provide hard packet delivery ratio or link longevity guarantees.

5.2 Dynamic topology

The issue of mobility does not exist in fixed wireline networks and in infrastructured wireless networks. The topology of MANET will change dynamically due to mobile host changing their point of connectivity, the rate of node survivability and nodes leaving or joining the network. Saving current knowledge of the network topology and the frequent changes is an important requirement in MANET management system. However the frequent exchanges of topology information may lead to considerable signaling overhead, congesting low bandwidth wireless links, and possibly depleting the limited battery life of the nodes involved. Hence the choice of mechanism used to collect topology information is critical. These complications imposed by mobility in MANETs may severely degrade the network quality. The frequent route breakage is a natural consequence of mobility, which complicates routing. As a result, design of QoS routing protocols in MANETs is challenged by frequent topological changes.

5.3 Node mobility

The nodes in a MANET may move completely independently and randomly as far as the communications protocols are concerned. This means that topology information has a limited lifetime and must be updated frequently to allow data packets to be routed to their destinations. Again, this invalidates any hard packet delivery ratio or link stability guarantees. Furthermore, QoS state which is link-position dependent or node position dependent must be updated with a frequency that increases with node mobility. An important general assumption must also be stated here: for any routing protocol to be able to function properly, the rate of

topology change must not be greater than the rate of state information propagation. Otherwise, the routing information will always be stale and routing will be inefficient or could even fail completely. This applies equally to QoS state and QoS route information. A network that satisfies this condition is said to be *combinatorially stable* [3].

5.4 Lack of centralized control

The major advantage of an ad hoc network is that it may be set up spontaneously, without planning and its members can change dynamically. This makes it difficult to provide any form of centralised control. As such, communications protocols which utilise only locally-available state and operate in a completely distributed manner, are preferred [24]. This generally increases an algorithm's overhead and complexity, as QoS state information must be disseminated efficiently.

5.5 Channel contention

In order to discover network topology, nodes in a MANET must communicate on a common channel. However, this introduces the problems of interference and channel contention. For peer-to-peer data communications these can be avoided in various ways. One way is to attempt global clock synchronization and use a TDMA-based system where each node may transmit at a predefined time. This is difficult to achieve due to the lack of a central controller, node mobility and the complexity and overhead involved [25]. Other ways are to use a different frequency band or spreading code (as in CDMA) for each transmitter. This requires a distributed channel selection mechanism as well as the dissemination of channel information. However data communications take place, without a central controller, some set-up, new neighbour discovery and control operations must take place on a common contended channel. Indeed, avoiding the aforementioned complications, much MANET research, as well as the currently most popular wireless ad hoc networking technology (802.11x) is based on fully-contended access to a common channel with Carrier-Sense Multiple Access with Collision Avoidance (CSMA/CA). However, CSMA/CA greatly complicates the calculation of potential throughput and packet delay, compared to TDMA-based approaches.

5.6 Heterogeneity

MANETs are typically heterogeneous networks with various types of mobile nodes with diverse nature of communication technologies employed. Its diversity comes in the form of different types of nodes, ranging from sensors, palmtops and laptops within an organisation or a result of multiorganisation consortium. In a military application, different military units ranging from soldiers to tanks can come together, hence forming a MANET system. Nodes differ in their energy capacities and computational abilities. Hence, mobile nodes will have different packet generation rates, routing responsibilities, network activities and energy draining rates. Coping with node heterogeneity is a key factor for the successful operation of MANETs.

5.7 Imprecise state information

In most cases, the nodes in a MANET maintain both the link-specific state information and flow-specific state information. The link-specific state information includes bandwidth, delay, delay jitter, loss rate, error rate, stability, cost, and distance values for each link. The flow specific information includes session ID, source address, destination address, and QoS

requirements of the flow (such as maximum bandwidth requirement, minimum bandwidth requirement, maximum delay, and maximum delay jitter). The state information is inherently imprecise due to dynamic changes in network topology and channel characteristics. Hence routing decisions may not be accurate, resulting in some of the real time packets missing their deadlines.

5.8 Hidden terminal problem

The hidden terminal problem is inherent in MANETs. This problem occurs when packets originating from two or more sender nodes, which are not within the direct transmission range of each other, collide at a common receiver node. It necessitates retransmission of packets, which may not be acceptable for flows that have stringent QoS requirements. The RTS/CTS control packet exchange mechanism, proposed in [26] and adopted later in the IEEE 802.11 standard [27], reduces the hidden terminal problem only to a certain extent. BTMA [28] and DBTMA [29] provide two important solutions for this problem.

5.9 Insecure medium

Due to the broadcast nature of the wireless medium, communication through a wireless channel is highly insecure. Hence security is an important issue in MANETs, especially for military and tactical applications. MANETs are susceptible to attacks such as eavesdropping, spoofing, denial of service, message distortion, and impersonation. Without sophisticated security mechanisms, it is very difficult to provide secure communication guarantees.

6. Factors to be considered in designing QoS routing protocol

6.1 QoS requirements specifications

For efficient QoS routing implementation, QoS requirements needs to be specified to the routing protocol. Consequently, they may be used as constraints on route discovery and selection. An application may typically request a particular QOS by specifying its requirements in terms of one or more of the following metrics.

i. Minimum throughput or capacity (bps); which is the desired application data throughput. [30];
ii. Maximum tolerable delay; normally defined as the maximum tolerable source to destination delay for data packets transmission[2];
iii. Maximum tolerable delay jitter, which is the difference between the upper bound on end-to-end delay and the absolute minimum delay [31]. This metric can also be expressed as delay variance [32];
iv. Maximum tolerable packet loss ratio (PLR) (%) which is the acceptable percentage of total packets sent, which are not received by the transport layer agent at the packet's destination node [33];

In most cases, the QoS protocol should only admit this data session into the network if it can provide the requested service. The mechanism by which this decision is made is termed admission control.

6.2 Metrics employed for route selection

This section lists many of the metrics commonly employed by routing protocols for path evaluation and selection in order to improve all-round QoS or to meet the specific requirements of application data sessions.

i. Network Layer Metrics

a. Achievable throughput; which is defined as the achievable data throughput of a path or node. The achievable throughput is often termed as the available bandwidth. [30]
b. End-to-end delay; which refers to the measured end-to-end delay on a route [2];
c. Node buffer space, the number of packets in a node's transmission buffer which determine the amount of delay a packet traveling through a node [34];
d. Delay jitter or variance[31].
e. Packet loss ratio (PLR) (%) which is the percentage of total packets sent, which is not received by the transport or higher layer agent at the packet's final destination node;
f. Route lifetime (s) which is the statistically calculated as expected lifetime of a route. Normally it depends on node mobility as well as node battery. [35];

ii. Medium Access Control(MAC) and Link Layer Metrics:

a. The time taken to transmit a packet between two nodes in a contention-based system is the MAC time delay. It includes the total time deferred and acknowledgement delay [36]. It provides an indication of packet traffic.
b. Frame delivery ratio (%) which is statistically determined, is the probability of a packet successfully being transmitted over a link and correctly decoded at the receiver [37][38].
c. Link stability which can be described as the predicted lifetime of a node pair connection [35]. It indicates the length of time node pairs are connected;
d. Node relative mobility can be measured as the ratio of the number of neighbours that change over a fixed period to the number that remain the same [39].

iii. Physical Layer Metrics:

a. Signal-to-interference ratio (SIR), where the received SIR at a destination node can be used as a routing metric indicating link quality, via cross-layer communication [40].
b. Bit error rate (BER) determines the level of error correction and/or number of retransmissions required over a connection and has major impact on the connection's reliability metric and on energy consumption. [41];
c. Node residual battery charge or cost [42] [43] [37].

QoS metrics such as the above can be classified as either additive, concave or multiplicative metrics, based on their mathematical properties [7].

6.3 Metrics for performance evaluations

The following metrics may be used to evaluate a QoS routing protocol's performance.

1. Transport/Application Layer:

a. Session acceptance/blocking ratio is the percentage of application data sessions (or transport layer connections) that are admitted into or rejected from the network. The value of this metric reflects both the effectiveness of the QoS protocols as well as conditions outside of their control, such as channel quality;
b. Session completion/dropping ratio is the metric represents the percentage of applications that were successfully/ unsuccessfully served after being admitted to the network. For example, if a VoIP session is accepted and the session is completed properly (by the users hanging up) and not aborted due to route failure or any other error, then that counts as a completed session.

2. Network Layer:

a. Throughput measured in bit per second(bps) which is the amount of data traffic the entire network carried to its destination in one second;
b. Node throughput (bps), defined as the average throughput achieved by a single node;
c. Route discovery delay for reactive protocols as a measure of effectiveness of the reactive protocols.
d. In order to measure the operating cost and efficiency of Qos routing protocol, normalised routing load (NRL) could be employed. It is the ratio of routing packets transmitted to data packets received at the destination [44].
e. Network lifetime (s) may be defined as the time until network partitioning occurs due to node failure [42], indicating the energy-efficiency and load balancing ability of the protocol.
f. Average node lifetime (s) [42] shows the effective power usage and optimization.

3. MAC Layer

a. Normalised MAC load is very similar to the normalised routing load (NRL), which represents the ratio of bits sent as MAC control frames to the bits of user data frames transmitted [44].
b. MAC energy efficiency, represent the ratio of energy used for sending data bits to the total energy expended for data plus MAC headers and control frames;

6.4 Factors affecting QoS routing protocol performance

When evaluating the performance of QoS protocols, a number of factors have a major impact on the results. Some of these parameters are a particular manifestation of characteristics of the MANET environment.

1. Node mobility which consists of a number of parameters: the nodes' maximum and minimum velocity, velocity pattern and pause time. The node's velocity pattern determines whether the node moves at uniform velocity at all times or whether it is constantly varying, and also how it accelerates. The pause time determines the length of time, nodes remain stationary between each period of motion. Together with maximum and minimum velocity, this parameter determines how often the network topology changes [44], [45].
2. Network size, the larger the network, the more difficult this becomes in terms of update latency and message overhead [24].
3. Number, type and data rate of traffic sources, a smaller number of traffic sources results in fewer routes being required and vice-versa. Traffic sources can be constant bit rate (CBR) or may generate bits or packets at a rate that varies with time according to the Poisson distribution, or any other mathematical model. The maximum data rate affects the number of packets in the network and hence the network load [44].
4. Node transmission power - some nodes may have the ability to vary their transmission power. This is important, since at a higher power, nodes have more direct neighbours and hence connectivity increases, but the interference between nodes does as well. Transmission power control can also result in unidirectional links between nodes [46], [47], [48].
5. Channel characteristics - as detailed earlier, there are many reasons for the wireless channel being unreliable i.e. many reasons why bits, and hence data packets, may not be delivered correctly. These all affect the network's ability to provide QoS.

6.5 Networking resources utilization for effective QoS routing protocol

1. Node computing time
   Today, mobile devices are manufactured with increasingly powerful processors but are still limited in computing power. It may be that they must not only run the applications, but also protocols necessary to support the network.
2. Node battery charge
   It may be the most critical resource. Node failures due to battery drained, can cause network partitioning, leading to a complete network failure. Hence, power-aware and energy efficient MAC and routing protocols have received a great deal of research attention [42][48].
3. Node buffer space
   Node buffer space must be always available to reduce packet transmission congestion. Data packets must be buffered while awaiting transmission and reception. Furthermore, when the buffers are full, any newly arriving packets must be dropped, contributing to the packet loss rate;
4. Channel capacity
   Channel capacity is measured in bps and affects data throughput. Indirectly, it contribute to the delay, and other metrics too. However, since all nodes must share the transmission medium, we must somehow express the fraction of the medium's total capacity that is granted for each node's use. The way to express this depends on the MAC layer technique employed. In a purely contention-based MAC, transmission opportunities may be envisioned, although no node can be guaranteed channel access, merely granted it with a certain probability.

6.6 A balanced trade-offs design

This section discusses some of the common trade-offs involved in QoS routing protocol design.

1. Route discovery and state dissemination, Proactive, On-Demand or Hybrid
   It refers to two problems under one heading. Firstly, should routes be discovered pro-actively or on-demand? Secondly, how should QoS state required for path selection, be discovered? If both the route and QoS state discovery mechanisms are proactive, then the session establishment time is greatly reduced from an application's point of view. Also, a proactive protocol is largely unaffected by an increase in the fraction of nodes acting as data sources, since routes to all destinations are maintained anyway. However, a large overhead is incurred in keeping routes and state up-to-date, especially in highly mobile scenarios. Additionally, such a mechanism does not scale well with an increasing number of nodes. These are well-known problems of proactive protocols [24]. A major advantage of discovering QoS state proactively surfaces in situations where different applications specify their requirements with different metrics. As long as it is decided which QoS states to keep up to date, a route may be computed from the routing table based on any QoS metric, without the need for a separate discovery process for each metric [49]. A purely reactive routing solution avoids the potential wastage of channel capacity and energy by not discovering routes and QoS state which are not currently needed. However, a discovery delay is incurred when an application requires a route to a destination [39].
2. Between Capacity and Delay
   It has been shown that in MANET, capacity can be traded off with packet delay [50][51]. If delay constraints are relaxed, then the capacity of the network can be

increased by exploiting multiuser diversity [50]. More specifically, if delay is not constrained, a source can split the packets of a session and send them to many different neighbours. These neighbours then forward the packets onto the destination when they move into its transmission range. This scheme has been shown to improve throughput, since far fewer intermediate nodes are transmitting packets and causing interference, but incurs the cost of greatly increased delay [50]. Another strategy is to improve delay by increasing redundancy, at the cost of network capacity utilization efficiency [51]. If multiple copies of a packet are forwarded on multiple paths, it has been shown that the destination receives the packet with less delay on average. On the other hand, more network capacity is consumed in sending duplicate packets [51]. Clearly, increased redundancy also reduces the protocol's energy efficiency.

3. The packet loss rate against the capacity and energy-efficiency
   In a similar way to the trade-off between delay and capacity, PLR can also be traded off against capacity. Increasing the redundancy by sending mueltiple copies of packets over different routes, results in a higher chance of the destination receiving a copy, but reduces the useful capacity of the network. This technique can be more useful in sensor networks where data is often broadcast without a reliable handshaking protocol being employed at the MAC layer. Once again, redundancy also increases the energy expended per packet.
4. Energy consumption vs. responsiveness and accuracy of QoS state information
   Routing can only be accurate if the frequency of neighbour discovery is high enough to reflect frequent topological changes. However, a high-responsiveness to change comes at an increased energy cost [52]. If we consider QoS routing, this tradeoff between accuracy and energy consumption is even more acute, since not only the topology view, but the QoS state information also requires frequent updating, to enable accurate QoS routing decisions to be made.
5. Transmission power control between long and short hops
   Varying the transmission power to adjust the number of hops required to forward a packet to its destination, can yield many advantages and drawbacks. This has often been called the .long hops vs. short hops dilemma [53]. Another question is whether protocol designers should assume the use of transmission power control (TPC) at all. Assuming TPC constrains the type of devices that can be employed, since not all nodes may be equipped with radios with TPC capability. Furthermore, employing TPC can often result in uni-directional links. For example, a node X may be able to transmit to a node Y, but Y cannot reply since it is using a lower transmission power, unless it knows the distance to X and can calculate the transmission power required to reach it.
6. Global goals or individual requirements
   In the eyes of a network designer, the goal is to please, by providing an all-round high QoS. The secondary goal is to increase the network lifetime, by proper management of the battery usage. However, each individual user or data session has its own specific requirements, and to satisfy the user, the network must match their requirements. In more complicated scenarios, an application may specify a variety of QoS constraints. For example, it may specify maximum tolerable values for PLR as well as packet delay. In this case, we desire the routing protocol to find a stable path with a light traffic load. However, from a network lifetime point of view, a path that has the least cost, is preferred. Our goal of low delay matches the aim of load balancing, although the path with the least traffic may not be a stable path. In this case, there's a clear conflict between various requirements. A protocol designer must decide how to address this trade-off.

6.7 Measurement techniques of all QoS routing parameters

1. Bandwidth

   The challenge in wireless ad hoc networks is that neighboring hosts must share the bandwidth, and there is no centralized control for allocating bandwidth among the nodes. Furthermore, intermediate hosts take part in forwarding packets. Therefore, the total effective capacity achievable is not only limited by the raw channel capacity, but is also limited by the interaction and interference among neighboring hosts. Thus, in order to offer bandwidth-guaranteed routing, bandwidth estimation is needed, yet accurately estimating available bandwidth at each host is a challenging problem. Most QoS-aware routing protocols, such as CEDAR, Ticket-based QoS Routing, ADQR and TDR, assume that the available bandwidth is known. However, some routing protocols try to propose an appropriate way to estimate the available bandwidth, such as OLSR-based QoS routing, AQDR, DSDV/TDMA and Reactive/TDMA. Various methods are proposed in these protocols for estimating the available bandwidth at the nodes.

   a. Exploit the carrier-sense capability of IEEE 802.11 and measure the idle and busy time ratio as used in OLSR-based QoS routing protocol and the QOSRGA [22].
   b. Add bandwidth consumption information to AODV routing packets (or Hello messages) and exchange this information with neighbor hosts (used in AQOR).
   c. Monitor and schedule free time slots using a TDMA scheme (used in DSDV / TDMA and Reactive/TDMA).
   d. Broadcast queries with limited hop count to actively contact all neighbors in the carrier-sensing range (used in CACP [25]-Multihop).
   e. Take advantage of power control and send queries to cover the carrier-sensing range (used in CACP-Power).
   f. Approximate the available bandwidth by using a moving average (used in CACP-CS). A drawback of AQOR's bandwidth estimation method is that it assumes that the interference range is same as the transmission range, which is not true in general. Thus AQOR's bandwidth estimation method will not correctly incorporate the bandwidth being used by neighbours in the interference range of the node.[54]

   The available bandwidth depends on the MAC scheduling, and several of the bandwidth estimation techniques currently proposed are associated with the underlying MAC protocols. Therefore, bandwidth estimation should be done with the assistance of the MAC protocol. A cross-layer design between the MAC and routing layers is the key to solve this problem.

2. Delay

   Only two routing protocols incorporate delay estimation: Ticket-based QoS aware routing and AQOR [54]. Ticket based QoS aware routing does not support a specified delay; it only determines the shortest delay route during route discovery. AQOR uses half the round-trip time of the route discovery process as the estimated path delay. These two schemes do not consider that changes in contention levels will impact the end-to-end delay significantly after the flow is started. Also, the effect of intra-flow contention on delay has not been sufficiently studied. Therefore, the second open issue in QoS-aware routing is: how should end-to-end delay be estimated to support delay-constrained real-time data transmission?

3. Jitter

   Jitter is the variation of delay over a period of time. Among the delay components are fixed delay components and variable delay components. Jitter results from the variable

delay components, specifically changes in queuing delays at network switches due to variations in the short term network load. Jitter is the statistical variance of the packet interarrival time. The IETF in RFC 1889 [55] define the jitter to be the mean deviation of the packet spacing change between the sender and the receiver. A node sends packets of identical size at constant intervals which implies that S_j *and* S_i that is sending times of two consecutive packets is constant. The difference of the packet spacing, denoted D, is used to calculate the interarrival jitter. According to the RFC, the interarrival jitter should be calculated continuously as each packet i is received. The jitter of a packet stream is defined as the mean deviation of the difference in packet spacing at the receiver compared to the sender, for a pair of packets. If S_i is the time packet i was sent from the sender, and R_i is the time it was received by the receiver, the jitter sample J_i is given by, $J_1 = (R_{i+1} - R_i) - (S_{i+1} - S_i)$ or $J_1 = (R_{i+1} - S_{i+1}) - (R_i - S_i)$ and the average jitter is the average value over n packets. The jitter is particularly important quantity for time sensitive data such as real-time audio and video, since a large jitter can have a profound effect on the perceived quality.

4. Packet Loss
   The monitor program uses the sequence number, the time stamp, and the local time information to determine the two qos parameters: Packet Loss Rate, L_r,, delay jitter, J_t. During each measuring period, it counts the total number of packet received, N_{total} and the total number of packet loss, n_{lost}. It also record the arrival and send times of the last packet in the measuring period, as $t_{last_arrival}$ and t_{last_send}. The arrival time taken to be the simulation time when the packet arrive to the receiver, while the send time is derived from the packet using the time stamp function. At the end of every period k, the network monitor computes the two qos parameters with the following calculations, $L_r(k) = n_{lost}(k)/n_{total}(k)$.
5. Power
   A lot of work on energy efficient routing in MANET has been done [42]. These efforts tried to maximize the time for network partition and reduces variations in power level of the nodes. The model proposed by Rishiwal *et al* [56] can be used to calculate energy values at different times. Energy consumption of a node after time t is calculated as follows: $E_{c(t)} = N_t * \alpha + N_r * \beta$ where $E_{c(t)}$ is the energy consumed by a node; N_t is the number of packets transmitted by the node after time t; N_r is the number of packets received by the node after time t; α and β are constant factors having a value between 0 and 1. Hence by using this formula, the energy usage can be known.

7. Categorisation of existing QoS routing protocol

In [6], QoS routing protocols are classified mainly by their: (i) treatment of network topology, which is either flat, hierarchical or location-aware, and (ii) approach to route discovery, either proactive, reactive, hybrid, or predictive. Whereas, in [7], they are classified in three ways; (i) the interaction between the route discovery coupled with QoS provisioning mechanism; (ii) decoupled interaction between the route discovery and QoS provisioning, (iii) the interaction with the MAC layer; either independent or dependent, (iv) the approach to route discovery. In Section 4, another form of QoS Routing classification

was presented based on heuristics: (i) admission control; (ii) scheduling discipline; (iii) ordering of QoS metrics; (iv) sequential filtering; (v) control theory approach and (vi) computational intelligence approach.
In this Section, we elaborate on the classification based on MAC protocol interaction, by considering the following classes of QoS routing solutions.

7.1 Contention-free MAC layer

In this classes of the solutions, those that rely on accurately-quantified resources such as channel capacity, availability and resource reservation, definitely requires a contention-free MAC solution. TDMA is one of them providing near hard QoS guarantee, typically afforded by wired network. Solutions employing a contention-free MAC, QoS guarantees are essentially hard, except for when channel fluctuations or node movements occur. This unpredictable conditions result in MANET would be unsuitable for providing truly hard QoS guarantees.

7.2 Contending MAC layer

These QoS routing algorithm rely only on a contended MAC protocol and therefore only on the available resources at that instant. These resources need to be statistically estimated. Such protocols typically use these estimations to provide soft guarantees. Implicit resource reservation may still be performed, by not admitting data sessions which are likely to degrade the QoS of previously admitted ones. However, all guarantees are based on contended and unpredictable channel access or are given only with a certain probability and are thus inherently soft.

7.3 Independent of MAC layer

There are QoS routing that are independent from the MAC protocol. Such protocols cannot offer any type of QoS guarantees. They typically estimate node or link states and attempt to route using those nodes and links in which more favourable conditions may exist. However, the achievable level of performance is usually not quantified or is only relative. The aim of such protocols is typically to foster a better average QoS for all packets according to one or more metrics. This comes often at the cost of trade-offs with other aspects of performance, increased complexity, extra message overhead or limited applicability.

8. Protocols relying on contention-free MAC

8.1 QoS routing using CDMA over TDMA network

The problem of concerned to QoS routing protocol designers was the process of discovering paths that satisfy a session's throughput requirement. This was due to the fact that assured throughput seemed to be the lowest common denominator among multimedia data sessions' requirements. Since throughput depends largely on a node gaining sufficient transmission opportunities at the MAC layer, where the first part of the solution is the ability to measure the channel capacity at a node. Then, a mechanism is required to estimate the achievable throughput on a path, utilising the available channel capacity of the node. Finally, this information can be used to perform session admission control, admitting only data sessions for which a path with adequate throughput has been found. Chen *et al* presented an early channel-capacity estimation scheme for mobile wireless networks [57].

The authors proposed a kind of clustering scheme to group nodes where each cluster employs a different spreading code under a CDMA scheme. Within clusters, the channel was time-slotted to deterministically allocate channel access opportunities for each node. Hence, the channel capacity could be measured in terms of time slots. Additionally, time slots may be reserved as a way of promising channel capacity to individual data sessions.
The ideas in [57] were taken further by Lin *et al* [30], wherein they devised a detailed algorithm for calculating a path's residual traffic capacity, seemingly filling in the gaps in detail left by [57]. Similar to the aforementioned work, they propose using a CDMA over TDMA network. The channel is time-slotted accordingly, but several communicating pairs can share a time slot by employing different spreading codes. A path's capacity is expressed in terms of free time slots. Route discovery is based again on DSDV [58]. Routing updates are used to refresh the *free slots* information in routing tables. The proposed algorithm first calculates the best combination of free slots on the path for maximum throughput and then attempts to reserve them for a particular data session. In brief, the algorithm deals with nodes in groups of three. Below each node we show the time slots that were free prior to a data session being admitted. In this case, the same six slots were free at each node. At a first trivial glance it appears that the path capacity is six slots. This illustrates that nodes must have some common free slots to communicate, but if all nodes have the same set of free slots, the efficiency of utilisation is not very high. Then, the effective path capacity usable by a new session is only two slots, despite six being initially free at each node. Once the available time slots and path capacity have been determined, reservation signaling takes place to reserve the necessary time slots for satisfying the requesting session's throughput requirement. The two described schemes offer a clear-cut definition of path capacity in terms of time slots and allow a routing protocol to provide throughput guarantees to application data sessions by reserving these slots. However, this comes at the cost of many assumptions. First of all, assuming a CDMA network assumes that each group of nodes is assigned a different spreading code. These must either be statically assigned at network start-up, or dynamically assigned. The former mechanism does not deal with nodes/clusters leaving/joining the network, which is one of the most basic characteristics of ad hoc networks. The latter scheme assumes that there is some entity for assigning spreading codes, which is against the ad hoc design principle of not relying on centralized control. A second assumption is that of time-slotting. For each frame to begin at the same time at each node, the network must be globally synchronised. Synchronisation signaling incurs extra overhead, and as stated in previous work [7], [25], in the face of mobility this becomes practically unfeasible. Furthermore, time slot assignments must be continually updated as nodes move, and sessions are admitted or completed. Since these designs were published, new TDMA based MAC protocol designs have come to fruition, such as the IEEE 802.15.3 standard [59]. However, this protocol is designed for use in wireless personal area networks where every node is in range of a controller which provides the time-slot schedule. Thus, it is not suitable for wider-area MANETs. The conclusion is that there is currently no ideal feasible solution for implementing TDMA in a multihop MANET environment.

8.2 Multiple path routing using ticket

Chen *et al* [2] proposed a QoS routing protocol which reduces route discovery overhead while providing guaranteed throughput and delay. The main novelty of their approach was in the method of searching for QoS paths. First of all, a proactive protocol, such as DSDV [58] is assumed to keep routing tables up-to-date, with minimum delay, bottleneck

throughput and minimum hop to each destination. When a QoS-constrained path is required for a data session, probes are issued by the source node, to discover and reserve resources through a path. Each probe is assigned a number of tickets and each ticket represents the permission to search one path. If greater number of tickets are issued, then the delay and throughput requirements are more stringent. Each intermediate node uses its routing table to decide which neighbours to forward the probe to and with how many of the remaining tickets. Neighbours through which a lower delay or higher achievable throughput to the destination is estimated, are assigned more tickets. So, for example, the source sends a probe with three tickets, which splits at the second node. Two tickets are issued to the bottom path since it is deemed to have a higher chance of satisfying the delay requirement. Due to the nature of MANETs, the state information is not assumed to be precise and therefore, each delay and bottleneck channel capacity estimated is assumed to be within a range of the estimate. Eventually all probes reach the destination allowing it to select the most suitable path. It then makes soft reservations by sending a probe back to the source. This probe also sets the incoming and outgoing links for the connection in each node's connections table, setting up a soft connection state. The reservations and states expire when data is not forwarded via that virtual connection for a certain period of time, hence the terms soft reservation/state. Speaking in its favour, this protocol can handle sessions with either a delay or throughput constraint. When such a constrained path is required, flooding is avoided via the ticket mechanism, while at the same time ensuring that more paths are searched when requirements are stringent, increasing the chance of finding a suitable route. Imprecise state information is also tolerated. However, the method has several drawbacks. Firstly, the protocol used to maintain routing tables for guiding the search probes is proactive, requiring periodic updates, thus incurring a large overhead and not scaling well with network size. Secondly, Chen *et al* [2] mentions that a TDMA/CDMA MAC is assumed to take care of channel capacity reservation, which has the drawbacks discussed in the previous section.

8.3 SIR and bandwidth guaranteed routing with additional transmit power

Another TDMA-based QoS routing protocol is presented by Kim *et al* [40] with channel capacity expressed in terms of time slots. Furthermore, this protocol aimed to concurrently satisfy the application's throughput requirement and its BER constraint. For BER constraint, it aims to achieve by assigning adequate transmit power to produce the necessary signal to interference ratio (SIR) between a transmitter and receiver pair, with lower BER. This is in contrast to the previous candidate solutions, which aimed merely to satisfy a single QoS constraint at a particular moment. The protocol is on-demand and in essence, follows a similar reactive route discovery strategy to DSR [61]. An advantage of this protocol is that it gathers multiple routes between a source and destination and allows them to cooperatively satisfy a data stream's throughput requirement. However, only paths that fulfill the SIR requirement on every link qualify as valid routes. However; the maximum achievable SIR is limited by the maximum transmit power. Time is split into frames with a control and data phase, each containing several time slots. In the control phase, each node has a specified slot and uses this to broadcast data phase slot synchronization, slot assignment and power management information. This broadcast is made at a predefined power level. The received power can be measured and knowing the transmit power, the path loss can be calculated. From this, it is possible to calculate the received SIR. This in turn leads to an estimation for

the required link gain and thus the required power at the transmitter, $p_{j-1}^{(i)est}$, where j is the current node in the path and i is the time slot index. When a route is required, a RREQ is broadcast by the source and is received by direct neighbours. As in previous TDMA examples, forwarding nodes must be careful not to transmit in a slot in which their upstream node is receiving contains the number of time slots and SIR requirements. Time slots at the current node must be idle and not used for receiving, to be considered for reservation. Slots for which $p_{j-1}^{(i)est}$ is lower, are preferred. As long as one free slot exists, the node is appended to a list in the RREQ packet, along with the required power estimate for the transmitter for that particular transmission slot. The destination eventually receives multiple RREQs, hence the need for only one free slot on each path, since multiple paths can cooperatively serve the throughput requirement. It returns RREPs to the source along the discovered paths, which deliver the estimated power information so that the correct power can be set in the relevant transmission time slots.

8.4 Node state routing

Most designers wrongly adopted wireline paradigm in designing QoS routing protocols [49]. According to this paradigm, nodes are connected by physical entities called links and routing should be performed based on disseminating the state of these links. It was suggested that the correct wireless paradigm assumed the sharing of a geographical space and the frequency spectrum with other node pairs nearby. It must be asserted that links cannot be considered independently of each other. The author instead proposed the Node State Routing(NSR)[49]. In NSR, each node maintains the state information about itself and the surrounding environment, in a routing table. This includes states such as its IP address, packet queue size and battery charge. However, to avoid relying on link state propagation, NSR requires GPS input. This provides extra states, the node's current location, relative speed and direction of movement. It is assumed that nodes can estimate the path loss to neighbouring nodes, using a pre-programmed propagation model and knowledge of the node positions. In this way, connectivity would be inferred. Using the aforementioned states, it would be possible to predict connectivity between nodes, whereas in most other protocols, links must be discovered. In order to perform routing functions nodes must periodically advertise their states to neighbours. Neighbours should further advertise selected states of their neighbours, for example, only those that have changed beyond a threshold. Using the states of its neighbours, a node may then calculate metrics that may be conceived as link metrics, except that measurements at both ends of the link can be taken into account. Moreover, since node states are readily available, they can be used to calculate QoS routes as required. As opposed to most other QoS routing protocols, the node states allow different QoS metrics to be considered for each requesting session, without re-discovering routes. A route can be calculated from the propagation map at each node, and its lifetime can be estimated. This approach shows huge potential for practical multiconstraint QoS routing in the future. Furthermore, since link states are not used, there is no need to update several link states when a single node moves, as in other protocols. Instead, only that one node's state needs to be updated in neighbours' state tables. Despite its many advantages, NSR also has several drawbacks. First and foremost, it relies on accurate location awareness, which limits its usefulness to devices that are capable of being equipped with GPS receivers or such. Secondly, as described in [49], throughput-

constrained routing depends on a TDMA-based MAC protocol for capacity reservation and throughput guarantees.

9. Protocols based on MAC contention

9.1 Core Extraction Distributed Ad Hoc Routing (CEDAR)

The CEDAR algorithm was proposed by Sivakumar *et al* [60]. Its name is derived by the fact that it is a topology management algorithm with core extraction mechanism as the main function. The core of a network is defined as the minimum dominating set (MDS). It means that all nodes are either part of this set or have a neighbour that is part of the set. The MDS calculation is a known NP-hard problem [60]. Therefore the algorithm only finds an approximation, of it. MDS is calculated in order to set the core nodes, hence be able to provide a routing backbone. It ensures that all nodes are reachable but not every node need to participate in route discovery. Non-core nodes could save energy by not participating and its overhead would also be reduced. Generally, local broadcasts are unreliable due exposed and hidden node problems [60]. Reliable local unicasts may be used to propagate routing and QoS state information. It utilised the uses of RTS-CTS handshaking to avoid hidden and exposed node problems. Additionally it ensures the broadcast packet is delivered to every neighbouring core nodes. This scheme is termed core broadcast. Using [60] only local state for QoS routing incurs little overhead, but far from optimal routes may be computed. Worst still no QoS route may be found, even if one exists. On the other hand, gathering the whole network state at each node results in a very high overhead. Theoretically it allows the computation of optimal routes, although there's a possibility of using stale information. CEDAR compromises, by keeping up to date, information at each core node about its local topology, as well as the link-state information about relatively stable links with relatively high residual capacity further away. This is done via increase and decrease waves. For every link, the nodes at either end are responsible for monitoring the available capacity on it and for notifying their dominators when it increases or decreases by a threshold value. The method of estimating available link capacity is not specified in [60]. However, nodes only have link capacity information from a limited radius due to the wave propagation mechanism. Thus, the QoS core path is determined in stages with each node routing as far as it can see capacity information, then delegating the rest of the routing to the furthest .seen. node on the core path. This process iterates until the final destination is reached and all links satisfy the achievable throughput requirement. The greatest novelties of this technique were the core broadcast and link capacity dissemination mechanisms. These ensure efficient use of network resources and relatively accurate and up-to-date knowledge of the QoS state, where it is required. Furthermore, this protocol does not rely on a TDMA network, as the protocols discussed in the previous section do. However, the problem of estimating available link capacities was left open.

9.2 Interference- aware QoS routing

In [62] the authours consider throughput-constrained QoS routing based on knowledge of the interference between links. The so-called clique graphs are established, reflecting the links that interfere with each other, hence preventing occurrence of simultaneous transmission. It operates by first recording the channel usage in *bps* of each existing data session on each link. It was noted that the total channel usage of the sessions occupying the links within the same clique should not exceed the channel capacity. A link's residual capacity is then calculated by

subtracting the channel usage of all sessions on links in the same clique from the link's nominal capacity. This link capacity information may be utilised to solve the throughput-constrained MANET routing problem. Additionally, Yang *et al* [25] published and discussed the problems of achievable throughput estimation in a contended-access network which depend on the node's transmission range, ***R.*** Nodes within the Carrier-Sense rang are termed as CS-neighbours, and this set of nodes is the CS-neighbourhood. The CS-range which is equivalent to ***2R*** model simulates the physical layer characteristics of network adapters which are able to sense the presence of a signal at a much greater range than that at which they are able to decode the information it carries. In a contention-based MAC protocol such as the 802.11 distributed coordination function(DCF)[63], a node may only transmit when it senses the channel idle. Therefore, any nodes transmitting within its CS-range may cause the channel to be busy and are thus in direct contention for channel access. This is one of the key realizations in [25] such that all nodes in the CS-range (CS-neighbours) must be considered when estimating a node's achievable throughput. More specifically, in 802.11, the channel is deemed idle if both the transmit and receive states are idle and no node within ***R*** has reserved the channel via the network. The major advantage of this protocol is that no extra control packets are introduced, since bandwidth information is piggybacked on the existing HELLO packets. While the approaches discussed in this section represent significant progress in achievable throughput estimation and admission control, and hence throughput constrained QoS routing, there are still shortcomings. It is well-known that as a network nears saturation, ready-to-send and data packet collisions (in a multihop network) become more frequent, wasting capacity. Additional capacity is wasted due to the 802.11 backoff algorithm, as the level of contention for the channel increases. The protocols discussed in this section do not consider these sources of wastage when calculating the residual capacity at each node.

9.3 Cross-layer multi-constraint QoS routing

Fan *et al* [36] proposed MAC delay metric, which was defined as the time between a packet being received by the MAC protocol from the higher layers, and an ACK being received for it, after it is transmitted. This includes the time deferred when awaiting channel access and is thus a useful metric for avoiding busy links. Link reliability and throughput constraints are also considered in [36], but they use pre-existing definitions and methods of calculation. The focus of the paper is on performing multiconstraint QoS routing with the aforementioned three metrics. The authour reiterates the fact that the multi-constraint QoS routing problem is NP-complete [2] when a combination of additive and multiplicative metrics are considered. Among the above metrics, delay is additive, link reliability is multiplicative and achievable throughput is concave. However, methods have been proposed for reducing this NP complete problem to one that can be solved in polynomial time. In one such method, all QoS metrics, except one, take bounded integer values. Then, the task of finding a path to satisfy all constraints can be performed by a modified Dijkstra's algorithm. The multiplicative metric is reduced to an additive one by taking the logarithm of the reliability percentage of a link. Also, the delay metric is reduced such that each link is represented by the percentage of the allowable total delay it introduces. The resulting problem in the new metric space can be solved in polynomial time. Then, a modified Bellman-Ford or Dijkstra's algorithm with the new reliability metric for link weights can be used to find an approximation to the optimal path. In each iteration, the total MAC delay along a path is checked and also paths which do not satisfy the channel capacity constraint are eliminated. An obvious advantage of this approach is the concurrent consideration of

several important QoS metrics in path selection. However, QoS state for all paths must be discovered and kept fresh. This incurs extra overhead. Furthermore, such a protocol requires the participation of other mechanisms which could measure the link reliability, MAC delay and available channel capacity at each node.

9.4 On-demand delay-constrained unicast routing protocol

Zhang *et al* proposed [5] a protocol with delay constrained routes for data sessions. The operation of the protocol are as follows: firstly, a proactive distance vector algorithm is employed to establish and maintain routing tables consists the distance and next hop along the shortest path to each destination node. When a delay constrained path is required, this information is used to send a probe to the destination along the shortest path to test its suitability. If this path satisfies the maximum delay constraint, the destination returns an ACK packet to the source, which reserves resources. For this purpose a resource reserving MAC protocol is assumed. If the minimum hop path does not satisfy the delay constraint, the destination initiates a directed and limited flood search by broadcasting a RREQ packet. Intermediate nodes forward the RREQ if the total of their respective distances from the destination and source is below a set threshold and also the path delay is below the delay constraint value. When a copy of the RREQ reaches the source with a path that meets the delay constraint, the route discovery process is complete. While this protocol aims to minimize the hop-distance between source and destination and discovers paths that satisfy a session's delay constraint, extra overhead is incurred by the proactive distance-vector protocol which maintains the routing tables.

9.5 QoS greedy perimeter stateless routing for ultra-wideband MANETs

A proposal by Abdrabou *et al* [33] highlights new direction for MANETs, that of employing an ultra-wideband (UWB) signal. Using UWB, a node's position can easily be estimated via triangulation techniques. This provides location information, without having to rely on GPS, for enabling a position-based routing protocol. The proposed algorithm extends to another protocol, Greedy Perimeter Stateless Routing (GPSR) for QoS routing, referring as QoS of GPSR for UWB MANETs (QGUM). Each node broadcasts beacons containing its ID and position to all of its neighbour nodes. The destination's position is learnt at the same time as its ID. When a route is required, the source node sends a RREQ to the neighbour node which is closest to the destination. The RREQ specifies, among other information, the requesting data session's total delay bound, its PLR constraint and the accumulated PLR so far. A node receiving the RREQ factors in its own PLR and compares the result with the PLR bound. If it is unacceptable, a <*Route Failure*> is sent back to the source node. In this case, the source node begins route discovery again, starting with a different node in its neighbour list. If the PLR bound is not exceeded, the intermediate node appends its ID to the RREQ, in a manner akin to other source-routing protocols. It also adds its location before performing the same procedure as the source to find the next node to forward the RREQ to. Each intermediate node performs the PLR checks and passes the RREQ to the neighbour closest to the destination, until the destination receives the RREQ. The above procedure describes route discovery. The methods for ensuring QoS on routes are as follows. QGUM can operate[33] with a contended MAC protocol, similar to the 802.11 DCF. After a route to the destination is discovered as detailed above, the session admission control procedure begins. Owing to the available position information, the destination can calculate which nodes on

the route are inside each other's CS-ranges and thus can transmit simultaneously. The destination then calculates the channel capacity required at each node for the data session to be admitted. It then sends an admission request (AdReq) back along the route. Each intermediate node checks its locally available capacity and the capacity of its csneighbours by flooding an AdReq. If the intermediate node and all its CS-neighbours have sufficient capacity, they temporarily reserve the necessary capacity for the session and the AdReq is forwarded to the next hop in the route back towards the source node. If any nodes or their CS-neighbours on the route have insufficient capacity, they generate an admission refused message, towards the source, which then invokes a route repair mechanism. However, the advantages of QGUM, must be balanced against the typically shorter range offered by UWB radios, which is only 10m at 110Mbps [64]. Hence, current standardisation efforts involving UWB radio technologies for wireless networks are targeted at personal area networks [65] [54] and not larger-scale ad hoc WLANs as 802.11x is. This limits the applicability of protocols based on a UWB physical layer.

10. Protocols independent of the type of MAC

10.1 QoS optimized link state routing

A QoS routing protocol based on Optimized Link State Routing(OLSR) is presented by Badis et al [65]. OLSR is a pro-active protocol in which information about 1-hop and 2-hop neighbours is maintained in each node's routing table. This information is disseminated via periodically broadcast HELLO messages. OLSR minimises the control overhead involved in flooding routing information by employing only a subset of nodes, termed multi-point relays (MPRs), to rebroadcast it. As a consequence, only MPRs are discovered during route discovery and are used as intermediate nodes on routes. Since only a subset of nodes are MPRs, the best links may not be utilised for routing. In QoS-OLSR (QOLSR) [65], this problem is solved by proposing new heuristics for building nodes' MPR sets in order to enable QoS routing to take place. QOLSR employs both a variation on the MAC delay metric and the achievable throughput metric for QoS routing. In contrast to many of the protocols discussed so far, although the analysis in [65] is based on the 802.11 MAC, QOLSR does not rely on the MAC protocol to provide residual channel capacity. These values are estimated statistically, using the periodic HELLO messages. The total expected MAC delay of a packet is a product of the average estimated delay or expected service time (EST) of one packet and the total number of packets awaiting transmission. The value of EST in turn depends on packets' transmission times and the expected number of retransmissions the MAC layer will have to perform. The FER (Frame Error Ratio) is approximated by taking the ratio of the number of HELLO messages received during a monitoring window to the number expected, which is calculated from the known HELLO sending rate. The FER provides an estimate of the number of retransmissions required for successful delivery of a data packet. The transmission delay of a packet depends on the amount of time a node spends backing off and resolving collisions. A detailed analysis in [65] shows that this is a function of the average backoff window size and the FER. Using these, the derived formulae yield an estimation for the EST of each packet and therefore the total MAC delay of a link between a node and its neighbour. The achievable throughput of a link is also calculated statistically. The MAC delay or EST of a packet is estimated as described above. Using this, and knowledge of the overhead posed by packet headers and MAC control frames, the throughput experienced by packets can be estimated.

10.2 Link stability-based routing

Rubin *et al* [35], considered the link stability as an important QoS metric. Stability is defined as the expected lifetime of a link, which is largely dependent on the node movement pattern. The paper describes the probability distribution functions (PDF) of link lifetimes under various node mobility models. The remaining link lifetime is estimated as the area under the PDF for a given mobility model, taken between the link's measured lifetime so far, and the infinity. For example, in the random destination mobility model, nodes do not change direction after selecting a destination, until they reach it. This mobility model was found to produce a link lifetime PDF similar to a Rayleigh distribution [35]. To find the probability that a link's remaining lifetime is greater than a time t, the PDF of the link lifetime is integrated between ($t + L_p$) and infinity, where L_p is the link's past lifetime. A link lifetime model such as the one above is proposed for each of a selection of mobility models. An application may specify a lower limit for acceptable path failure probability, P_{fail}. This value can be calculated based on a data session's delay, delay jitter and packet loss rate requirements. It is proposed [35] that this mechanism is combined with AODV for QoS routing. The value P_{fail} is inserted into RREQ packets. Intermediate nodes test that the cumulative failure probability of links up to that point (also stored in the RREQ and updated by each node), is not greater than P_{fail} . Therefore, using an appropriate model such as the above and given the data session's duration, it is possible to calculate the probability of a path remaining intact for the duration of the data session, $P_{survive}$. If this is unacceptable i.e. $P_{survive} < P_{\text{fail}}$, the session is not admitted. This simple mechanism could be useful for statistically predicting link lifetimes and therefore avoiding links and paths that have a high probability of failure while a session is active. An obvious difficulty with this approach is that the node mobility pattern must be known and must be modeled accurately for the lifetime estimation to be useful. However, combined with other stability metrics, as shall be discussed later, this could be a useful component of a more sophisticated QoS provisioning mechanism. Another approach that considers link and path stability as an important QoS metric, is presented in [66]. A new variation on the stability metric is introduced in the form of the entropy metric. This is defined for a link as a function of the relative positions and velocities, and the transmission ranges of the link's two end nodes. A path's entropy is defined as the product of the link entropies along it. The lower the entropy, the higher the path stability. This scheme is incorporated into a source-routed scheme somewhat akin to DSR, and during route discovery, the path entropy (among other metrics) is calculated. A destination receives RREQs over multiple paths and waits a specified interval after receiving the first one, before selecting the path with the lowest entropy i.e. highest stability. This route is returned to the source in the RRep, thereby completing the route discovery. This approach has the potential to be more accurate than that in [35], since it considers nodes' relative positions and velocities for calculating the probability of link failure, rather than just a general PDF for a given mobility model. However, this comes at the price of assuming that each node is capable of determining its position via GPS or some similar system [42].

10.3 Hybrid Ad hoc Routing Protocol

The Hybrid Ad hoc Routing Protocol (HARP) is introduced in [39]. It uses the notion of quality of connectivity (QoC) as its routing metric. This is defined as a function of two nodes states: residual buffer space and relative stability. The latter is defined for node x over a chosen period of time, t_1-t_0, as $stab(x) = \left| \frac{N_{t0} \cap N_{t1}}{N_{t0} \cup N_{t1}} \right|$, where N_{t0} and N_{t1} are the set of neighbours of x at

times t_0 and t_1 respectively. Thus, stability is greater, the fewer the number of neighbour nodes that change between t_0 and t_1. The higher a node's residual buffer space and relative stability, the better the QoC to it is. The QoC of each node is used in a logical topology construction algorithm. Each node periodically broadcasts a beacon to all of its neighbours, which contains its address and QoC. Then, each node selects as its preferred neighbour (PN) the neighbour node with the highest QoC. A link between a node and its PN is called a preferred link. A logical tree is constructed by connecting nodes together using only preferred links. A tree's growth terminates where a node's preferred link is with a node that is already part of the tree. This heuristic has been proven to yield a forest of trees [39]. In brief, each tree is then considered a routing zone, within which proactive routing occurs. Inter-zone routing is performed on-demand, and hence the hybrid route discovery of this protocol. In inter-zone routing, other zones may be abstracted as nodes, thus a packet can be routed to another zone, and on arrival, the intra-zone routing mechanism can direct the packet to its final destination. HARP also includes route discovery optimizations which reduce overhead. Firstly, the forest structure can be used to avoid having to flood route request (RREQ) packets used in inter-zone routing. This is done by forwarding RREQs only via gateway nodes; a node is considered to be a gateway, if it is the neighbour of a leaf node, but it is in another zone. Secondly, features of the Relative Distance Microdiscovery (RDM) routing protocol (RDMAR) [67] are incorporated into HARP. RDMAR does not limit the number of neighbours propagating a flooded packet, but limits the scope of the flooding instead. Thus, RREQs do not propagate to areas of the network where they will be useless, thereby wasting resources. The time-to-live (TTL) field in a RREQ is set based on an estimation of the relative distance of the destination in terms of hops. However, the estimation can only be made if there is some previous knowledge of the destination, and a replacement path to it is sought. In this case, the relative stabilities of each node on the path, combined with the time elapsed since the stabilities were recorded, yields an estimation for the total maximum change in the positions of the nodes on the path. This is added to the previous known distance in metres of the destination. The sum is divided by the radio range to obtain an estimated upper bound on the distance of the destination in number of hops. This value is used for the TTL.

10.4 Delay-Sensitive Adaptive Routing Protocol

The Delay-Sensitive Adaptive Routing Protocol (DSARP) [34] employs reactive route discovery, is completely decoupled from the MAC protocol and provides delay guarantees for time-sensitive data sessions. Its basic operation is very similar to classical reactive MANET routing protocols such as DSR. However, when a path is required for delay-sensitive traffic, a different algorithm is employed. The source node sends a route request (RREQ), as usual. This is allowed to propagate to the destination, which sends a route reply (RRep). When forwarding the RRep, each intermediate node on the path attaches the number of packets awaiting transmission in its buffer. Multiple RReps may be received by the source node, which then selects several shortest paths, if there are multiple. Alternatively, the shortest path plus the next shortest path are selected. Using the information about buffer usage at each node, the source calculates the total number of packets on each selected path. Finally, the traffic flow on each path is adjusted such that the new traffic allocated to it is greater if the existing traffic on it is lower and the number of packets on other paths is greater. This algorithm pushes the network towards a state where each path has an equal flow of traffic on it and thus is likely to produce the same packet delay. Essentially, this implements a form of load balancing,

ensuring that the energy usage of nodes is also distributed evenly. After adjusting the traffic on each path, a statistical guarantee can be made about the delay on that path. DSARP is simple to implement and provides delay guarantees without relying on the MAC protocol, but has the following disadvantages. The number of buffered packets on each path must be rediscovered each time a new session begins, regardless of whether the route has failed or not. This incurs extra overhead. Also, the delay guarantees may fail in the face of mobility, if other nodes move into contention range and cause greater channel access delays for nodes on a session's path.

10.5 Application-aware QoS routing

A rather unique approach to QoS routing is presented in [32]. It is unique because instead of using lower layer (MAC) information, it is based on the aid of the transport layer. The proposal, referred to as Application Aware QoS Routing (AAQR) in the literature, assumes the use of the real-time transport protocol (RTP) [68]. The delay between two nodes is estimated statistically by examining the difference between time stamps on transmission and receipt of RTP packets between those two nodes. The delay variance is also calculated. Furthermore, each node records the throughput requirement of RTP sessions which are flowing through it. Subtracting the total of these throughput values from the raw channel capacity gives an estimate for the total remaining capacity at that node. When a QoS-route is required, applications may specify throughput and delay constraints. In [32] delay is considered the most important constraint for multimedia applications. Routes are discovered on-demand, although the details of the route-discovery procedure are not discussed. A subset of the discovered routes is selected, such that all paths satisfy the delay constraint of the application. From this subset a further subset of routes is selected, which also satisfy the application's throughput constraint. Finally, from this second subset, the route with the lowest variance in RTP packet transmission delays, is chosen. If there are no routes that meet the throughput requirement, the route with the highest available channel capacity, which satisfies the delay constraint, is selected. A major advantage of AAQR is that no extra overhead is incurred for QoS routing, since the existing transport layer packets are used for QoS metric estimation. Additionally, both delay and throughput constraints may be considered. However, the use of RTP is assumed, and therefore the range of application scenarios for this protocol is obviously limited.

10.6 Ad hoc QoS On-demand Routing (AQOR)

AQOR [54] is a QoS-aware routing protocol with the following features: (1) available bandwidth estimation and end-to-end delay measurement, (2) bandwidth reservation, with the optimal bandwidth path is the path with the largest bottleneck bandwidth among all possible paths and (3) adaptive route recovery. AQOR is an on-demand QoS-aware routing protocol. When a route is needed, the source host initiates a route request, in which the bandwidth and delay requirements are specified. The intermediate hosts check their available bandwidth and perform bandwidth admission hop-by-hop. If the bandwidth at the intermediate host is sufficient to support the request, an entry will be created in the routing table with an expiration time. If the reply packet does not arrive in the allotted time, the entry will be deleted. Using this approach, a reply packet whose delay exceeds the requirement will be deleted immediately in order to reduce overhead. To estimate available bandwidth for assisting in call admission, each node puts its reserved bandwidth in periodic

Hello messages that are sent to their neighbors. AQOR uses the sum of a node's neighbors' traffic as the estimated total traffic affecting the node. This estimated traffic can be larger than the real overall traffic. This overestimation imposes a stringent bandwidth admission control threshold. The available bandwidth is thus a lower bound on the real available bandwidth. End-to-end one way downstream delay is approximated by using half the round trip delay. With the knowledge of available bandwidth and end-to-end delay, the smallest delay path with sufficient bandwidth is chosen as the QoS route. Temporary reservation is used to free the reserved resources efficiently at each node when the existing routes are broken. If a node does not receive data packets in a certain interval, the node immediately invalidates the reservation. This avoids using explicit resource release control packets upon route changes. The adaptive route recovery procedure includes detection of broken links and triggered route recovery at the destination, which occurs when the destination node detects a QoS violation or a time-out of the destination's resource reservation.

10.7 Adaptive QoS Routing algorithm (ADQR)

Hwang *et al* proposed an adaptive QoS routing algorithm (ADQR) to find multiple disjoint paths with long lifetimes [41]. ADQR differs from other QoS routing protocols by using signal strength to predict the route breaks and initiate a fast reroute of data. Three levels of signal strength, $Th1$, $Th2$, and Sr ($Th1 > Th2 > Sr$), are defined. Sr is the minimal signal strength to receive a data packet. Three different classes are also defined for nodes, links and routes. If the received signal strength from a neighbor node is higher than $Th1$, that neighbor node is in the first node class. If the received signal strength from the neighbor is between $Th1$ and $Th2$, that neighbor node is in the second node class. If the signal strength is between $Th2$ and Sr, that neighbor node is in the third node class. Links between the first node class nodes are in the first link class; links between the second node class nodes are in the second link class; and links between the third node class nodes are in the third link class. Also, three route classes are defined, where the bottleneck link determines the path class. Each node keeps a neighbor table, which records the node's neighbors and their corresponding cumulative signal strength, defined as:

$$SSnew-cummulative = \delta \times SSold-cummulative + (1 - \delta) \times SSnew-measured$$

where δ is adjusted according to network conditions and is the current received signal strength. Also, two symbols are used to indicate the relative motion of the two nodes: "+" indicates that the two nodes are moving away from each other; while "-" indicates that the distance between the two nodes is shrinking. Each node also keeps a routing table, of the form <*source, destination, next hop, hop count, available bw, reserved bw, active, route class, first class link, second class link, third class link*>. The source node sends a *Route Request* packet, which carries the information <*source, destination, request id, hop cnt, QoS metric, route class, int nodes, first class link, second class link, third class link*>. Intermediate nodes append their own address in the *int nodes* field, update the parameters *QoS metric*, *route class*, and *hop cnt*, and forward the *Route Request* to their neighbors. The destination node checks whether this path is disjoint from other paths already found and whether *route class* is anything but "+3" . If the first condition is true and the second is false, the destination node does the same procedure as an intermediate node, creates a *Route Reply* packet, and inserts the route information into its routing table. When an intermediate node receives a *Route Reply* packet,

the node inserts the route into its local routing table, if there is no corresponding route entry; or the node updates its routing table, if the route already exists. When the source node receives multiple routes, the choice of which route to use is based on the *route class* information. The *first route class* routes obtain higher priority than the *second route class* and the *third route class* routes. Similarly, the *second route class* routes obtain higher priority than the *third route class* routes. After selecting the desired route(s), bandwidth is reserved by sending a *QoS Reserve* packet from the source to the destination along the selected route(s). ADQR uses a fast route maintenance scheme, called two-phase monitored rerouting, which is composed of *Pre Rerouting* and Rerouting. The *Pre Rerouting* phase occurs when the route changes from *first route class* to *second route class*, and the Rerouting phase is invoked when the route changes from *second route class* to *third route class*. In *Pre Rerouting*, the source node finds alternate paths in advance, before the current path becomes unavailable, and in Rerouting, the source node switches to one of these alternate paths in advance of the current path becoming unavailable.

11. Computational intelligence approach

11.1 Genetic Algorithm-based QoS routing

In [38], a Genetic Algorithm-based source-routing protocol for MANETs (GAMAN) is proposed, which uses end-to-end delay and transmission success rate for QoS metrics. Genetic Algorithms (GAs) may be employed for heuristically approximating an optimal solution to a problem, in this case finding the optimal route based on the two QoS constraints mentioned. The first stage of the process involves encoding routes so that a GA can be applied; this is termed gene coding. For this purpose, paths are discovered on-demand and then a network topology view is constructed in a logical tree-like structure. Each node stores a tree routed at itself with its neighbour nodes as child nodes and in turn their neighbour nodes as their children. Tree reductions are used to avoid duplicate subtrees. Each tree junction is considered a gene and multiple genes make up a chromosome which represents a path. The route discovery algorithm is assumed to collect locally computed metrics such as average delay over a link and the link reliability for the links on each path. After the gene encoding stage, the fitness, T of each path, is calculated as follows:

$$T = \frac{\sum_{i=1}^{n} D_i}{\prod_{i=1}^{n} R_i}$$

where D_i and R_i are the delay and reliability of link i respectively. The fitness values are used to select paths for cross-over breeding and mutation operations. The fittest path (with the smallest T) and the offspring from the genetic operations are carried forward into the next generation. While this method is a useful heuristic for approximating the optimal value over the delay and link reliability metrics at the same time, it requires many paths to be searched in order to collect enough .genetic information for the GA operations to be meaningful. This means that the method is not suited to large networks, as the authors themselves admit [38]. The methods of calculating Di and Ri are not detailed, but we assume they can be calculated statistically by the end nodes of each link. Collecting and maintaining sufficient route and

QoS state information to make a GA useful for QoS routing is costly in terms of both overhead and energy consumption. However, heuristic methods are often the only feasible way of solving NP-complete multi-constraint multihop QoS routing problems. Thus, while their general applicability to MANETs is limited, GAs may play a niche role in finding near-optimal routes, while satisfying multiple QoS constraints in certain environments. For example, MANETs which are less power-constrained and experience lower levels of mobility, and/or MANETs having topologies where a relatively small number of nodes can be combined in a relatively large number of ways to construct valid routes. The GAMAN protocol discussed in this section provides an exploratory example of how GAs may possibly be applied in such networks.

Another QoS routing algorithm was proposed by Peng et al [69]. The authors proposed route discovery technique, RLGAMAN. It tries to increase the probability of success in finding QoS feasible routes and integrates a distributed route discovery scheme with a reinforcement learning (RL) method. It utilizes the local information for the dynamic network environment; and the route expand scheme based on genetic algorithms (GA) method to find more new feasible paths and avoid the problem of local optimize. The performance of the RLGAMAN was investigated by simulation experiment using NS2. The authours claimed that when compared to traditional method, the experiment results showed the network with RLGAMAN had improved its efficient and effectiveness.

11.2 QOSRGA

QOSRGA (QoS Routing Using GA) was designed to select QoS route based on QoS metrics such as bandwidth, delay and node connectivity index (nci) [70]. QoS Routing for MANET posses several challenges that must be addressed. In selecting the most optimal route from source to destination, one has to choose from a set of routes with the corresponding quality of connectivity and resources. Due to the nature of node mobility the protocol demands an exceptional performance. It needs to select a single route with the longest residual node-pair connectivity time simultaneously. The proposed QOSRGA is based on source routing which effectively select the most viable routes in terms of bandwidth availability, end-to-end delay, media access delay and the sum of nci. The NDMRD protocol [22] initially determined a number of potential routes by calculating the number of returning Route Reply(RREP) packet from destination. The returning RREP packets extract the QoS parameters from each node along the routes. Genetic Algorithm (GA), then operates on the accumulated set of routes and the corresponding set of QoS parameters. A genetic algorithm for this particular problem must have these five issues resolved before the application of the generic GA framework: (1) a genetic representation for potential solutions to the problem called *chromosomes*. (2) a methodology to create an initial population of potential solutions. (3) an evolution function that plays the role of the environment, rating solutions in terms of their *fitness*. (4) GA operators that alter the structure of chromosomes. (5) values for various parameters that the genetic algorithm uses such as population size and probabilities of applying genetic operators.

11.3 Fuzzy logic approach

Gomathy and Shanmugavel [21] have shown how to integrate the techniques of fuzzy logic and scheduling principles to produce a fuzzy-based priority scheduler. The paper analyzed the performance of the novel fuzzy-based priority scheduler, for data traffic and evaluated

the effect of inclusion of this scheduler with different underlying multicast routing protocols, like NTPMR, CAMP, and ODMRP, run over IEEE 802.11 as the MAC protocol. Queuing dynamics with different degrees of mobility and routing protocols show that the composition of packets in the queue determines the effect of giving priority to control packets or setting priorities among data packets, for the average delay. During low mobility, the average delay is dominated by network congestion due to data traffic. During high mobility, it is dominated by route changes. We have addressed a fuzzy-based priority scheduler for data packets, which improves the QoS parameters in MANETs. The fuzzy scheduler attaches a priority index to each packet in the queue of the node. Unlike the normal sorting procedure for scheduling packet, a crisp priority index is calculated based on the inputs such as queue length, data rate, and expiry time of packets, which are derived from the network. The membership functions and rule bases of the fuzzy scheduler are carefully designed.

Sun *et al* [71] proposed QoS routing algorithm based on fuzzy logic. They proposed Fuzzy controller based QoS Routing Algorithm with a multiclass scheme (FQRA) for mobile ad hoc networks. In FQRA, a routing table is maintained to manage the lifetime of the active routes. Then FQRA applies a fuzzy logic system to dynamically evaluate the route expiry time. The fuzzy logic is chosen because there are uncertainties associated with node mobility and the estimation of link crash; moreover, there exist a mathematical model capable of estimating the node mobility. In addition, FQRA is able to take some controlling factors into consideration. The performance of the FQRA is studied using NS2 and evaluated in terms of quantitative measures such as improved path success ratio, reduced average end-to-end delay and increased packet delivery ratio. Generally it shows a promising approach.

11.4 Biologically inspired algorithm

In this paper, we propose a new version of the self organized Emergent Ad hoc Routing Algorithm with QoS provisioning (EARA-QoS). This QoS routing algorithm uses information from not only the network layer but also the MAC layer to compute routes and selects different paths to a destination depending on the packet characteristics. The underlying routing infrastructure, EARA originally proposed in [72], is a probabilistic multi-path algorithm inspired by the foraging behaviour of biological ants. The biological concept of *stigmergy* in an ant colony is used for the interaction of local nodes to reduce the amount of control traffic. Local wireless medium information from the MAC layer is used as the artificial *pheromone* (a chemical used in ant communications) to reinforce optimal/sub-optimal paths without the knowledge of the global topology. One of the optimisations of EARA-QoS over EARA is the use of metrics from different layers to make routing decisions. This algorithm design concept is termed as the *cross-layer* design approach. Research [73] has shown the importance of cross-layer optimisations in MANETs, as the optimisation at a particular single layer might produce non-intuitive side-effects that will degrade the overall system performance. Moreover, the multiple-criteria routing decisions allow for the better usage of network characteristics in selecting best routes among multiple available routes to avoid forwarding additional data traffic through the congested areas, since the wireless medium over those *hotspots* is already very busy. The parameters for measuring wireless medium around a node depend largely on the MAC layer. In this paper, we focus on the IEEE 802.11 DCF mode [74], since it is the most widely used in both cellular wireless networks and in MANETs. This cross-layer technique of using MAC layer information can

be applied easily to other MAC protocols. In addition to the basic routing functionality, EARA-QoS supports an integrated lightweight QoS provision scheme. In this scheme, traffic flows are classified into different service classes. The classification is based on their relative delay bounds. Therefore, the delay sensitive traffic is given a higher priority than other insensitive traffic flows. The core technique of the QoS provision scheme is a *token bucket* queuing scheme, which is used to provide the high priority to the real-time traffic, and also to protect the lower-priority traffic from starvation. Experimental results from simulation of mobile ad hoc networks show that this QoS routing algorithm performs well over a variety of environmental conditions, such as network size, nodal mobility and traffic loads.

11.5 Energy- and reliability-aware routing

The Maximum Residual Packet Capacity (MRPC) protocol is proposed in [37], which considers battery charge as well as link reliability during route selection. Admittedly, MRPC is not intended to be a QoS routing protocol, but we consider it here since it utilizes some QoS-related metrics to improve all-round QoS. Routing based on residual battery charge is considered extensively in the literature [48]. However, in our view, protocols that consider only this state are not useful for QoS routing, since they do not improve the QoS experienced by individual data sessions or packets. On the other hand, MRPC also considers link reliability, as detailed below. In [37] a node-link metric is introduced to capture the energy-lifetime of a link between nodes ***i***(transmitter) and ***j***, which is defined as:

$$L_{i,j} = \frac{R_i}{E_{i,j}}$$

where R_i is the residual battery charge at node i and $E_{i,j}$ is the energy required to transmit a data packet of a given size over the link (i, j). A suggested formulation for $E_{i,j}$ is as follows

$$E_{i,j} = \frac{T_{i,j}}{\left(1 - p_{i,j}\right)^H}$$

where $T_{i,j}$ is the energy required for one transmission attempt of the aforementioned data packet with a fixed transmission power. Also, $p_{i,j}$ is the packet error probability of the link (***i***, ***j***) and H = 1 if hop by hop retransmissions are performed by the link layer. From the above formulae, it is clear that the lifetime of a link is higher when greater battery charge remains at the transmitter node, and when the reliability of the link is high, resulting in a low energy cost for correctly transmitting a packet. These formulae give an estimation for the expected number of data packets that can be transmitted over a link before the battery of the transmitter fails [37]. Then, if a route failure is said to occur when any single link on it fails, the lifetime of path ***p*** in number of packets is simply:

$$Life_p = \min_{(i,j)\in p} \left\{ L_{i,j} \right\}$$

MRPC considers the best route to be the one with the greatest residual lifetime. The authours[23], suggests that the MRPC algorithm may be implemented in AODV [75] for application in MANETs. As routes are discovered, the lifetime of the path is accumulated by

calculating the lifetime of each link. The next hop to a destination is always selected to be the neighbour which results in the greatest possible value for $Life_p$. This protocol results not only in load balancing, increasing the life of the network and avoiding congestion, but also yields closer-to-optimal energy consumption per packet, as well as lower packet delay and packet loss probability, due to the preference for more reliable links. It can also be implemented in an on-demand fully distributed routing protocol, such as AODV. However, link reliabilities must somehow be estimated, which may not be a trivial problem. Furthermore, like HARP, MRPC does not cater to particular sessions' requirements, only fosters better all-round QoS, and hence may be unsuitable for many applications. On the other hand, as mentioned above, MRPC is not primarily intended to be a QoS routing protocol, rather an energy-efficient best effort protocol.

12. Progressive trends in this area

As we discussed in Section 6, many of the earlier QoS routing proposals for MANETs were based on contention-free MAC protocols and relied on either TDMA or TDMA/CDMA channel access mechanisms. This was probably due to their well-understood nature from the field of cellular communications. A TDMA approach offers a straightforward method of quantifying channel capacity and access opportunities, as well as allowing such opportunities to be deterministically reserved for particular application data sessions. This enables throughput guarantees to be made, provided that the network dynamics do not invalidate them. Due to mobility, as well as the unpredictable nature of the wireless channel, truly hard guarantees can never be made in a MANET. Even though some newer proposals continue to assume TDMA, it is believed that non-hierarchical TDMA-based methods are highly unfeasible in MANETs[25], since time slotting requires global clock synchronisation, which is difficult to achieve in a mobile environment. A further drawback of this approach is the high signaling overhead incurred by slot scheduling and the potential complexities thereof [57]. Newer MAC protocols such as that specified by 802.15.3 [59] offer feasible TDMA solutions for MANETs by introducing node hierarchies whereby a group of nodes in a piconet is synchronised by a central controller node. However, this protocol is designed only for personal area networks and not for largescale multi-hop MANETs. On the other hand, CDMA based methods introduce the problem of code allocation in a dynamic mobile environment. In light of these conclusions, QoS routing methods that rely on such channel access methods are not the solution for general and especially larger-scale MANETs. This is reflected in the literature, since the majority of later solutions, are based on contended MAC protocols (generally 802.11). In Section 9 we discussed several proposals relying on a contended MAC protocol, such as 802.11. Many less mature solutions in this category did not consider the nature of contention between neighbouring nodes sufficiently accurately and thus reliable QoS provisioning did not become a reality for MANETs. It was through key works such as [25], [76], that the nature of contention and its effect on (primarily throughput-constrained) QoS routing, begun to be well-understood. Other newer proposals take this understanding as a basis for further QoS routing designs. Some proposals greatly further the field of QoS session admission control. Many solutions continue to be based upon 802.11x and its CSMA/CA-based channel access mechanism. Even though 802.11 is an aging standard, the CSMA/CA mechanism has survived into its most recent versions and therefore proposals based on the 802.11 MAC protocol continue to be very relevant. On the other hand, QoS routing proposals based on an ultra-wideband physical layer [33] are

emerging. As we discussed though, UWB radios have a limited shorter range compared to 802.11x. Accordingly, current UWB standardisation efforts are all aimed at personal area networks, meaning that UWB-based QoS routing proposals have limited applicability to small-scale MANETs only. Statistical QoS Protocols that make no assumptions about the MAC layer have also received greater attention in the last few years. Such protocols allow a simpler modular network stack design, without the complications of cross-layer issues. However, no guaranteed level of service is provided, as we saw in the proposals discussed in Section 10. Instead, such protocols generally improve the all-round average QoS experienced by packets under some metrics, at the expense of other performance metrics or increased complexity or overhead. Such protocols may not be sufficient for supporting applications with stringent QoS requirements. By contrast, protocols in this category have done much to improve QoS robustness to failures, which was another area identified as future work in previous surveys. The link and node stability-based techniques that were summarised in Section 10 can find longer-lasting routes and thus improve the robustness of QoS solutions against failures caused by mobility. In summary we can say that there is a trend for QoS routing solutions to move away from contention free MAC dependence and towards contended-MAC dependence for throughput-constrained applications. To cater for many other metrics, such as delay and PLR, numerous statistical protocols which are independent of the MAC layer, have been proposed. Another aspect of development considers the metrics themselves. Again, in the earlier proposals, the focus was on providing an assured throughput service only, since throughput was deemed the most important requirement. Some earlier protocols could serve, for example, either a throughput or a delay requirement, but not both simultaneously. In this context, the trend we observe has been to move from single-constraint routing to multi-constraint routing, as demonstrated by the later proposals we have discussed. However, multiconstraint routing remains an NP-complete problem [2], [77] and thus most of the described solutions do not aim to find optimal routes. Instead, they simply apply multiple metrics to route filtering, removing all that do not satisfy a particular constraint. One exception was described in Section 11.2, in which a genetic algorithm is employed as an heuristic to finding the optimal route based on more than one metric.

13. Future works

Following on from this survey, we believe that there is still some way to go in the area of throughput-constrained routing, before perfect QoS Routing protocol is achieved, even in a low-mobility scenario. Works such as [25], [75] consider channel contention, as well as MAC overheads in achievable throughput estimation, but the time wasted due to deferring transmission, random back-off and collisions has not been considered. The wastage due to collisions is especially difficult to calculate in a multi-hop environment. This is important future work, if accurate residual channel capacity estimation is to be realised with contented MAC. The understanding of contention among nodes also needs to be transferred to considerations of other QoS metrics, such as end-to-end packet delay, which is affected by the queues of all nodes within contention range [49]. Delay jitter and energy consumption (due to collisions), are also affected. Quantifying the impact on these metrics and more, in the light of contention awareness and collisions, designing routing protocols that incorporate this knowledge and evaluating them with realistic application layer models, is all future work. A further trend that we have observed, is that many designers place great

emphasis on the session admission (QoS route finding) capability of their protocol, which is admittedly very important. In contrast, they often neglect or downplay the importance of session completion i.e. maintaining the routes and the QoS for as long as an application data session requires. An aspect of this, QoS robustness, was highlighted by earlier survey writers. However, more work on the evaluation of QoS sensitive session completion performance with realistic application layers, would be useful. Ultimately, session completion is more important from a user perspective, than session admission. This is because the perceived QoS is better when some sessions are blocked but none are dropped mid-session, rather than all sessions being admitted, but some failing. Furthermore, fast local QoS route-repairing schemes require additional investigation to improve QoS session completion rates and protocols' robustness against mobility. In Section III we reiterated that one of the major challenges to the provision of QoS in MANETS is the unreliable wireless channel. However, we have found that the majority of QoS routing protocol evaluation studies assume a perfect physical channel, ignoring the effects of shadowing and multi-path fading. Therefore, studying the impact of a more realistic physical layer model on QoS routing protocol performance is another interesting area of future work.

As mentioned in the previous section, while simple multi-constraint QoS routing proposals are numerous, there are few that attempt to optimise multi-constraint routing. One example was based on genetic algorithms [38]. However, such methods have limited applicability due to the overhead and energy cost of collecting enough state information. Accurate studies are required to establish, with various networking environments and topologies, whether or not it is feasible to collect and maintain sufficient state information to apply methods such as GAs. For the cases where it is, more research is required on different types of heuristic algorithms for calculating near-optimal paths with multiple QoS constraints. Comparative studies on the performance and impact of the heuristics, are additional future work. Moreover, there is a distinct lack of protocol frameworks for incorporating such methods into practically-realisable systems. One promising, but perhaps not yet mature or feasible approach is that of Node State Routing [49]. Such a solution would provide the mechanism by which to disseminate the information to enable multi-constraint QoS routing.

14. Summary

In this paper we reviewed the challenges to and basic concepts behind QoS routing in MANETs and provided a thorough overview of QoS routing metrics and design considerations. We then classified many of the major contributions to the QoS routing solutions pool published in recent years. The protocols were selected in such a way as to highlight many different approaches to QoS routing in MANETs, while simultaneously covering most of the important advances in the field since the last such survey was published. We summarised the operation, strengths and drawbacks of these protocols in order to enunciate the variety of approaches proposed and to expose the trends in designers' thinking. The protocols' interactions with the MAC layer were also described. Finally, we provided an overview of the areas and trends of progress in the field and identified topics for future research.

15. References

[1] I. Chlamtac, M. Conti, J.J.-N. Liu, "*Mobile ad hoc networking: imperatives and challenges*," Adhoc, vol. 1, pp. 13-64, 2003.

[2] S. Chen, K.Nahrstedt; Distributed Quality-os-Service Routing in Ad Hoc Networks, IEEE JSAC V17, N8, Aug99, pp 1-18.

[3] S. Chakrabarti and A. Mishra, .QoS issues in ad hoc wireless networks,. *IEEE Commun. Mag.*, vol. 39, pp. 142.148, Feb. 2001.

[4] S. Chakrabarti and A. Mishra, .Quality of service challenges for wireless mobile ad hoc networks,. *Wiley J. Wireless Commun.and Mobile Comput.*, vol. 4, pp. 129.153, Mar 2004.

[5] B. Zhang and H. T. Mouftah, .QoS routing for wireless ad hoc networks: problems, algorithms and protocols,. *IEEE Commun. Mag.*, vol. 43, pp. 110.117, Oct. 2005.

[6] J. N. Al-Karaki and A. E. Kamal, .Quality of service routing in mobile ad hoc networks: Current and future trends,. in *Mobile Computing Handbook* (I. Mahgoub and M. IIays, eds.), CRC Publishers, 2004.

[7] T. B. Reddy, I. Karthigeyan, B. Manoj, and C. S. R. Murthy, .Quality of service provisioning in ad hoc wireless networks: a survey of issues and solutions. available online: http://www.sciencedirect.com, Apr 2004.

[8] E. Crawley, R. Nair, B. Rajagopalan, H. Sandick, *A Framework for QoS-based Routing in the Internet*, IETF, Network Working Group, RFC 2386, August 1998.

[9] Z. Wang, J. Crowcroft, "Quality of Service Routing for Supporting Multimedia Applications", *IEEEJSAC*, September 1996.

[10] Z. Zhang, C. Sanchez, W. Salkwicz, E. Crawley, *Quality of Service Extensions to OSPF or Quality of Service Path First Routing (QOSPF)*, IETF, Internet Draft, September 1997.

[11]R. Guérin, S. Kamat, A. Orda, T. Przygienda, D. Williams, *QoS Routing Mechanisms and OSPF Extensions*, IETF, RFC 2676, August 1999.

[12]M. Oliveira, J. Brito, B. Melo, G. Quadros, E. Monteiro, "Quality of Service Routing in the Differentiated Services Framework", *Proceedings of SPIE's International Symposium on Voice, Video, and Data Communications (Internet III: Qu ality of Service and Future Directions)*, Boston, Massachusetts, USA, November 5-8, 2000.

[13]G.Apostolopoulos, R. Guérin, S. Kamat, and S. Tripathi, "Quality of service Based Routing: A Performance Perspective", *Proceedings of ACM SIGCOMM'98*, Vancouver, BC, Canada, August 31-September 4, 1998.

[14] R. Guérin, A. Orda, "QoS-based Routing in Networks with Inaccurate Information: Theory and Algorithms'", *Proceedings of IEEE INFOCOM'97*, Kobe, Japan, April 1997.

[15]A. Shaikh, J. Rexford, and K. Shin, *Dynamics of Quality-of-Service Routing with Inaccurate Link-StateInformation*, University of Michigan Technical Report CSE-TR-350-97, November 1997.

[16]S. Chen, K. Nahrstedt, "Distributed QoS Routing with Imprecise State Information", Proceedings of the International Conference on Computer, Communications and Networks (ICCCN'98), Lafayette, LA, October 1998.

[17] G. Apostolopoulos, R. Guérin, S. Kamat, and S. Tripathi, "Improving QoS Routing Performance Under Inaccurate Link State Information", *Proceedings of the16th International Teletraffic Congress (ITC-16)*, Edinburgh, UK, June 1999.

[18] Q. Ma, P. Steenkiste, "Quality-of-Service Routing for Traffic with Performance Guarantees" *Proceedings of IFIP Fifth International Workshop on Quality of Service*, Columbia University, New York, May 1997.

[19] S. Rampal, *Routing and End-to-end Quality of Service in Multimedia Networks*, PhD Thesis, Department of Electrical and Computer Engineering, North Caroline State University, August 1995.

[20] B. Li and K. Nahrstedt, "A control theoretical model for quality of service adaptations " in *Proceedings of Sixth IEEE International Workshop on Quality of Service,* 1998, pp. 145-153.

[21] C. Gomathy and S. Shanmugavel, Supporting QoS in MANET by a Fuzzy Priority Scheduler and Performance Analysis with Multicast Routing ProtocolsEURASIP Journal onWireless Communications and Networking 2005:3, 426–436

[22] J.Abdullah and D.J.Parish, "QOSRGA Protocol Using Non-Disjoint Multiple Routes in Mobile AD Hoc Networks." The MMU International Symposium on Information and Communication Technologies (M2USIC2007), 19-20 Nov 2007, Petaling Jaya, Malaysia

[23] S. Saunders, *Antennas and Propagation for Wireless Communication Systems Concept and Design*. New York, USA: John Wiley and Sons, 1999.

[24] C. E. Perkins, ed., *Ad Hoc Networking*, ch. 3. Addison Wesley, 2001.

[25] Y. Yang and R. Kravets, .Contention-aware admission control for ad hoc networks,. *IEEE Trans. Mobile Comput.*, vol. 4, pp. 363. 377, Aug 2005.

[26] P. Karn, MACA: A new channel access method for packet radio, in: Proceedings of ARRL/CRRL Amateur Radio 9th Computer Networking Conference, September 1990, pp. 134–140.

[27] IEEE Standards Board, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, The Institute of Electrical and Electronics Engineers Inc., 1997.

[28] F.A. Tobagi, L. Kleinrock, Packet switching in radio channels: Part II––The hidden terminal problem in carrier sense multiple-access and the busy-tone solution, IEEE Transactions on Communications 23 (12) (1975) 1417–1433.

[29] J. Deng, Z.J. Haas, Dual busy tone multiple access (DBTMA): a new medium access control for packet radio networks, in: Proceedings of IEEE ICUPC 1998, vol. 1, October 1998, pp. 973–977.

[30] C. R. Lin and J.-S. Liu, .Qos routing in ad hoc wireless networks,. *IEEE J. Select. Areas Commun.*, vol. 17, pp. 1426. 1438, Aug. 1999.

[31] A. R. Bashandy, E. K. P. Chong, and A. Ghafoor, .Generalized quality-of-service routing with resource allocation,. *IEEE J. Select. Areas Commun.*, vol. 23, pp. 450.463, Feb 2005.

[32] M. Wang and G.-S. Kuo, .An application-aware QoS routing scheme with improved stability for multimedia applications in mobile ad hoc networks,. in *Proc. IEEE Vehicular Technology Conf.*, pp. 1901.1905, Sep. 2005.

[33] A. Abdrabou and W. Zhuang, .A position-based qos routing scheme for UWB mobile ad hoc networks,. *IEEE J. Select. Areas Commun.*, vol. 24, pp. 850.856, Apr. 2006.

[34] M. Sheng, J. Li, and Y. Shi, .Routing protocol with QoS guarantees for ad-hoc network,. *Electronics Letters*, vol. 39, pp. 143.145, Jan. 2003.

[35] I. Rubin and Y.-C. Liu, .Link stability models for QoS ad hoc routing algorithms,. in *Proc. 58th IEEE Vehicular Technology Conf.*, vol. 5, pp. 3084.3088, Oct. 2003.

[36] Z. Fan, .QoS routing using lower layer information in ad hoc networks,. in *Proc. Personal, Indoor and Mobile Radio Communications Conf.*, pp. 135.139, Sep. 2004.

[37] A. Misra and S. Banerjee, .MRPC: Maximising network lifetime for reliable routing in wireless environments,. in *Proc. IEEE Wireless Communications and Networking Conf.*, (Orlando, Florida), March 2002.

[38] L. Barolli, A. Koyama, and N. Shiratori, .A QoS routing method for ad-hoc networks based on genetic algorithm,. in *Proc. 14th Int. Wksp. Database and Expert Systems Applications*, pp. 175. 179, Sep. 2003.

[39] N. Nikaein, C. Bonnet, and N. Nikaein, .Hybrid ad hoc routing protocol - HARP,. in *Proc. Int. Symp. Telecommunications*, 2001.

[40] D. Kim, C.-H. Min, and S. Kim, .On-demand SIR and bandwidth-guaranteed routing with transmit power assignment in ad hoc mobile networks,. *IEEE Trans. Veh. Technol.*, vol. 53, pp. 1215.1223, July 2004.

[41] N. Wisitpongphan, G. Ferrari, S. Panichpapiboon, J. Parikh, and O. Tonguz, .Qos provisioning using BER-based routing in ad hoc wireless networks,. in *Proc. Vehicular Technology Conf.*, vol. 4, pp. 2483.2487, May 2005.

[42] S. Singh, M. Woo, and C. S. Raghavendra, .Power-aware routing in mobile ad hoc networks,. in *Proc. Int. Conf. Mobile Computing and Networking*, pp. 181.190, 1998

[43] C.-K. Toh, .Maximum battery life routing to support ubiquitous mobile computing in wireless ad hoc networks,. *IEEE Trans. Commun.*, vol. 39, no. 6, pp. 138.147, 2001.

[44] C. E. Perkins, E. M. Royer, S. R. Das, and M. K. Marina, Performance comparison of two on-demand routing protocols for ad hoc networks,. *IEEE Personal Commun. Mag.*, vol. 8, pp. 16.28, Feb. 2001.

[45] J. Broch, D. A. Maltz, D. B. Johnson, Y.-C. Hu, and J. Jetcheva, .A performance comparison of multi-hop wireless ad hoc network routing protocols,. in *Proc. Int. Conf. on Mobile Computing and Networking*, Oct. 1998.

[46] J.-H. Chang and L. Tassiulas, .Energy-conserving routing in wireless ad-hoc networks,. in *Proc. IEEE INFOCOM*, vol. 1, pp. 22.31, 2000.

[47] S. Doshi, S. Bhandare, and T. Brown, .An on-demand minimum energy routing protocol for a wireless ad-hoc network,. *Mobile Computing and Communications Review*, vol. 6, no. 2, pp. 50.66, 2002.

[48] C. Yu, B. Lee, and H.-Y. Youn, .Energy-ef_cient routing protocols for mobile ad-hoc networks,. *Wiley J. Wireless Commun. and Mobile Comput.*, pp. 959.973, December 2003.

[49] J. Stine and G. de Veciana, .A paradigm for quality of service in wireless ad hoc networks using synchronous signalling and node states,. *IEEE J. Select. Areas Commun.*, vol. 22, pp. 1301.1321, Sep. 2004.

[50] M. Grossglauser and D. Tse, .Mobility increases the capacity of ad hoc wireless networks,. *IEEE/ACM Trans. Networking*, 2002.

[51] E. Neely, M.J.and Modiano, .Capacity and delay tradeoffs for ad hoc mobile networks,. *IEEE Trans. Inform. Theory*, 2005.

[52] L. Galluccio and S. Morabito, G.and Palazzo, .Analytical evaluation of a tradeoff between energy efficiency and responsiveness of neighbor discovery in self-organizing ad hoc network,. *IEEE J. Select. Areas Commun.*, vol. 22, pp. 1167.1182, Sep. 2004.

[53] D. Haenggi, M.and Puccinelli, .Routing in ad hoc networks: a case for long hops,. *IEEE Commun. Mag.*, vol. 43, pp. 93.101, Oct. 2005.

[54] Qi Xue and Aura Ganz, Ad Hoc Qos On-Demand Routing(AQOR) in Mobile Ad Hoc Networks, Journal of Parallel and Distributed Computing, V63, 2003, Pg 154-165.

[55] RTP: A Transport Protocol for Real Time Applications, IETF RFC1889,
Available: www.ietf.org/rfc/rfc1889.txt

[56] Prasant Mohapatra, Jian Li and Chao Gui, 'QoS In Mobile Ad Hoc Networks ', IEEE Wireless Communications, June 2003.

[57] T.-W. Chen, J. T. Tsai, and M. Gerta, .QoS routing performance in multihop, multimedia, wireless networks,. in *Proc. IEEE 6th Int. Conf. Universal Personal Communications*, vol. 2, pp. 557. 561, Oct 1997.

[58] C. E. Perkins and P. Bragwat, .Highly dynamic destination sequenced distance-vector routing (DSDV) for mobile computers, in *Proc. ACM SIGCOMM '94*, pp. 234.244, 1994.

[59] IEEE Computer Society, *Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for High-Rate Wireless Personal Area Networks (WPANs)*, 2003. IEEE Std. 802.15.3-2003.

[60] R. Sivakumar, P. Sinha, and V. Bharghavan, .CEDAR: a core extraction distributed ad hoc routing algorithm,. *IEEE J. Select. Areas Commun.*, vol. 17, pp. 1454.1465, Aug. 1999.

[61] R. Gupta, Z. Jia, T. Tung, and J. Walrand, .Interference-aware qos routing (IQRouting) for ad-hoc networks,. in *Proc. Global Telecommunications Conf.*, vol. 5, pp. 2599.2604, Nov. 2005.

[62] D. Johnson, D. Maltz, and J. Broch, *DSR: The Dynamic Source Routing Protocol for Multihop Wireless Ad Hoc Networks in Ad Hoc Networking*, ch. 5, pp. 139.172. Addison-Wesley, 2001.

[63] IEEE Computer Society, *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, 1999. ANSI/IEEE Std. 802.11, 1999 Ed.

[64] D. Porcino and W. Hirt, .Ultra-wideband radio technology: Potential and challenges ahead,. *IEEE Commun. Mag.*, vol. 41, pp. 66.74, July 2003.

[65] H.Badis and K.Al Agha, QOLSR, "QoS routing for Ad Hoc Wireless Networks Using OSLR", Wiley European Transactions on Telecommunications, Vol 15, no 4: 427-442, 2005.

[66] H. Shen, B. Shi, L. zou, and H. Gong, .A distributed entropy based long-life qos routing algorithm in ad hoc network,. in *Proc.IEEE Canadian Conf. on Electrical and Computer Engineering*, vol. 3, pp. 1535.1538, May 2003.

[67] G. Aggelou and R. Tafazolli, .RDMAR: A bandwidth-efficient routing protocol for mobile ad hoc networks,. in *Proc. 2nd ACM Int. Wksp. Wireless mobile multimedia*, pp. 26.33, 1999.

[68] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, .RTP: A transport protocol for real-time applications (rfc 3550). IETF RFC, July 2003.

[69] Fu Peng and Zhang Deyun, Hybrid Optimize Strategy based QoS Route Algorithm for Mobile Ad hoc Networks, Journal of Computer Science 2 (2): 160-165, 2006, ISSN 1549-3636 © 2006 Science Publications

[70] J.Abdullah and D.J.Parish, Node Connectivity Index as Mobility Metric for GA-Based QoS Routing in MANET, Proc of ACM/IEEE, Intl Conf on Mobile Technology, Application and Systems, 10-12 Sept 2007, Singapore.

[71] B. Sun, C. Gui, Q. Zhang, H. Chen, Fuzzy Controller Based QoS Routing Algorithm with a Multiclass Scheme for MANET, Int. J. of Computers, Communications & Control, ISSN 1841-9836, E-ISSN 1841-9844 Vol. IV (2009), No. 4, pp. 427-438

[72] Z. Liu, M. Z. Kwiatkowska, and C. Constantinou, "A swarm intelligence routing algorithm for manets," in *Proceedings of the IASTED International Conference on Communications Internet and Information Technology*, no. 433, November 2004, pp. 484–489.

[73] E. M. Royer, S. J. Lee, and C. E. Perkins, "The effects of mac protocols on ad hoc networks communication," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, vol. 2, 2000, pp. 543–548.

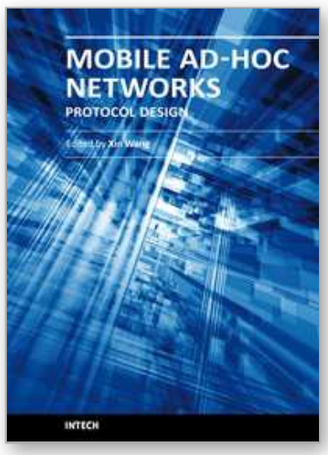
[74] IEEE Computer Society LAN MAN Standards Committee, *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, Institute of Electrical and Electronics Engineers Std., 1997.

[75] C. E. Perkins and E. M. Royer, .Ad hoc on-demand distance vector routing,. in *Proc. 2nd IEEE Wksp. Mobile Computing Systems and Applications*, (New Orleans, LA), pp. 90.100, Feb. 1999.

[76] L. Chen and W. Heinzelman, .QoS-aware routing based on bandwidth estimation for mobile ad hoc networks,. *IEEE J.Select. Areas Commun.*, vol. 23, pp. 561.572, Mar. 2005.

[77] F. Kuipers and P. Van Mieghem, .Conditions that impact the complexity of QoS routing,. *IEEE/ACM Trans. Networking*, vol. 13, no. 4, pp. 717.730, 2005.

IntechOpen



Mobile Ad-Hoc Networks: Protocol Design

Edited by Prof. Xin Wang

ISBN 978-953-307-402-3
Hard cover, 656 pages
Publisher InTech
Published online 30, January, 2011
Published in print edition January, 2011

Being infrastructure-less and without central administration control, wireless ad-hoc networking is playing a more and more important role in extending the coverage of traditional wireless infrastructure (cellular networks, wireless LAN, etc). This book includes state-of-the-art techniques and solutions for wireless ad-hoc networks. It focuses on the following topics in ad-hoc networks: quality-of-service and video communication, routing protocol and cross-layer design. A few interesting problems about security and delay-tolerant networks are also discussed. This book is targeted to provide network engineers and researchers with design guidelines for large scale wireless ad hoc networks.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Jiwa Abdullah (2011). QoS Routing Solutions for Mobile Ad Hoc Network, Mobile Ad-Hoc Networks: Protocol Design, Prof. Xin Wang (Ed.), ISBN: 978-953-307-402-3, InTech, Available from: http://www.intechopen.com/books/mobile-ad-hoc-networks-protocol-design/qos-routing-solutions-for-mobile-ad-hoc-network

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2011 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License, which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen