QUALITY OF SERVICE IN AD HOC WIRELESS NETWORKS

Issues and Challenges in Providing QoS in MANET- Classification of QoS Solutions - MAC Layer and Network Layer Solutions - QoS Frameworks - Need for Energy Management - Classification, Battery Management Scheme - Transmission Power and System Power Management Scheme.

INTRODUCTION

Quality of service (QoS) is the performance level of a service offered by the network to the user. The goal of QoS provisioning is to achieve a more deterministic network behavior, so that information carried by the network can be better delivered and network resources can be better utilized. A network or a service provider can offer different kinds of services to the users. Here, a service can be characterized by a set of measurable pre specified service requirements such as minimum bandwidth, maximum delay, maximum delay variance (jitter), and maximum packet loss rate. After accepting a service request from the user, the network has to ensure that the service requirements of the user's flow are met, as per the agreement, throughout the duration of the flow (a packet stream from the source to the destination). In other words, the network has to provide a set of service guarantees while transporting a flow. After receiving a service request from the user, the first task is to find a suitable loop-free path from the source to the destination that will have the necessary resources available to meet the QoS requirements of the desired service. This process is known as QoS routing. After finding a suitable path, a resource reservation protocol is employed to reserve necessary resources along that path. QoS guarantees can be provided only with appropriate resource reservation techniques. For example, consider the network shown in Figure 5.1. The attributes of each link are shown in a tuple < *BW, D* >, where *BW* and *D* represent available bandwidth in Mbps and delay1 in milliseconds. Suppose a packet-flow from node *B* to node *G* requires a bandwidth guarantee of 4 Mbps. Throughout the chapter, the terms "node" and "station" are used interchangeably. QoS routing searches for a path that has sufficient bandwidth to meet the bandwidth requirement of the flow. Here, six paths are available between nodes *B* and *G* as shown in Table 5.1.

QoS routing selects path 3 (*i.e., B → C → F →* G) because, out of the available paths, path 3 alone meets the bandwidth constraint of 4 Mbps for the flow. The end-to-end bandwidth of a path is equal to the bandwidth of the bottleneck link (*i.e.,* the link having minimum bandwidth among all the links of a path). The end-to-end delay of a path is equal to the sum of delays of all the links of a path. Clearly, path 3 is not optimal in terms of hop count and/or end-to-end delay parameters, while path 1 is optimal in terms of both hop count and end-to-end delay parameters. Hence, QoS routing has to select a suitable path that meets the QoS constraints specified in the service request made by the user. QoS routing has been described in detail in Section 10.5.1. Delay includes transmission delay, propagation delay, and queuing delay.

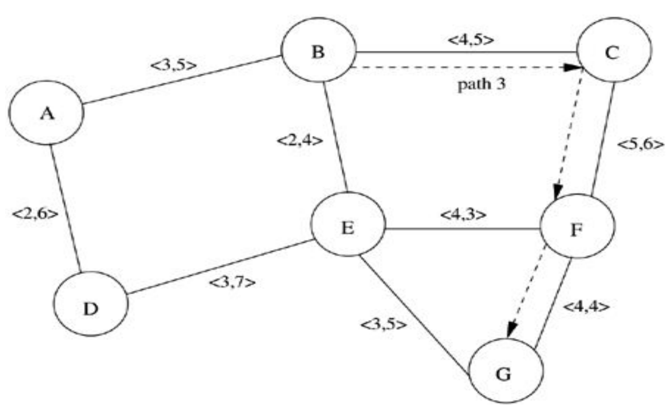

Figure 5.1. An example of QoS routing in ad hoc wireless network.

Table 5.1. Available paths from node *B* to node *G*

No.	Path	Hop Count	End-to-end Bandwidth (Mbps)	End-to-end Delay (milliseconds)
1	$B \rightarrow E \rightarrow G$	2	2	9
2	$B \rightarrow E \rightarrow F \rightarrow G$	3	2	11
3	$B \rightarrow C \rightarrow F \rightarrow G$	3	4	15
4	$B \rightarrow C \rightarrow F \rightarrow E \rightarrow G$	4	3	19
5	$B \rightarrow A \rightarrow D \rightarrow E \rightarrow G$	4	2	23
6	$B \rightarrow A \rightarrow D \rightarrow E \rightarrow F \rightarrow G$	5	2	25

QoS provisioning often requires negotiation between host and network, call admission control, resource reservation, and priority scheduling of packets. QoS can be rendered in ad hoc wireless networks through several ways, namely, per flow, per link, or per node. In ad hoc wireless networks, the boundary between the service provider (network) and the user (host) is not defined clearly, thus making it essential to have better coordination among the hosts to achieve QoS. Characteristics of ad hoc wireless networks such as lack of central coordination, mobility of hosts, and limited availability of resources make QoS provisioning very challenging.

ISSUES AND CHALLENGES IN PROVIDING QOS IN AD HOC WIRELESS NETWORKS

Providing QoS support in ad hoc wireless networks is an active research area. Ad hoc wireless networks have certain unique characteristics that pose several difficulties in provisioning QoS. Some of the characteristics are dynamically varying network topology, lack of precise state information, lack of a central controller, error-prone shared radio channel, limited resource availability, hidden terminal problem, and insecure medium. A detailed discussion on how each of the above-mentioned characteristics affects QoS provisioning in ad hoc wireless networks is given below.

Dynamically varying network topology: Since the nodes in an ad hoc wireless network do not have any restriction on mobility, the network topology changes dynamically. Hence, the admitted QoS sessions may suffer due to frequent path breaks, thereby requiring such sessions to be reestablished over new paths. The delay incurred in reestablishing a QoS session may cause some of the packets belonging to that session to miss their delay targets/deadlines, which is not acceptable for applications that have stringent QoS requirements.

• **Imprecise state information:** In most cases, the nodes in an ad hoc wireless network 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.

• **Lack of central coordination:** Unlike wireless LANs and cellular networks, ad hoc wireless networks do not have central controllers to coordinate the activity of nodes. This further complicates QoS provisioning in ad hoc wireless networks.

• **Error-prone shared radio channel:** The radio channel is a broadcast medium by nature. During propagation through the wireless medium, the radio waves suffer from several impairments such as attenuation, multipath propagation, and interference (from other wireless devices operating in the vicinity).

• **Hidden terminal problem:** The hidden terminal problem is inherent in ad hoc wireless networks. 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 the retransmission of the packets, which may not be acceptable for flows that have stringent QoS requirements. The RTS/CTS control packet exchange mechanism, proposed in and adopted later in the IEEE 802.11 standard, reduces the hidden terminal problem only to a certain extent.

• **Limited resource availability:** Resources such as bandwidth, battery life, storage space, and processing capability are limited in ad hoc wireless networks. Out of these, bandwidth and battery life are critical resources, the availability of which significantly affects the performance of the QoS provisioning mechanism. Hence, efficient resource management mechanisms are required for optimal utilization of these scarce resources.

• **Insecure medium:** Due to the broadcast nature of the wireless medium, communication through a wireless channel is highly insecure. Therefore, security is an important issue in ad hoc wireless networks, especially for military and tactical applications. Ad hoc wireless networks 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. Some of the design choices for providing QoS support are described below.

• **Hard state versus soft state resource reservation:** QoS resource reservation is one of the very important components of any QoS framework (a QoS framework is a complete system that provides required/promised services to each user or application). It is responsible for reserving resources at all intermediate nodes along the path from the source to the destination, as requested by the QoS session. QoS resource reservation mechanisms can be broadly classified into two categories: *hard state* and *soft state* reservation mechanisms. In hard state resource reservation schemes, resources are reserved at all intermediate nodes along the path from the source to the destination throughout the duration of the QoS session. If such a path is broken due to network dynamics, these reserved resources have to be explicitly released by a de-allocation mechanism. Such a mechanism not only introduces additional control overhead, but may also fail to release resources completely in case a node previously belonging to the session becomes unreachable. Due to these problems, soft state resource reservation mechanisms, which maintain reservations only for small time intervals, are used. These reservations get refreshed if packets belonging to the same flow are received before the timeout period. The soft state reservation timeout period can be equal to packet inter- arrival time or a multiple of the packet inter-arrival time. If no data packets are received for the specified time interval, the resources are deallocated in a decentralized manner without incurring any additional control overhead. Thus no explicit teardown is required for a flow. The hard state schemes reserve resources explicitly and hence, at high network loads, the call blocking ratio will be high, whereas soft state schemes provide high call acceptance at a gracefully degraded fashion.

• **Stateful versus stateless approach:** In the stateful approach, each node maintains either *global state* information or only *local state* information, while in the case of a stateless approach, no such information is maintained at the nodes. State information includes both the topology information and the flow specific information. If global state information is available, the source node can use a centralized routing algorithm to route packets to the destination. The performance of the routing protocol depends on the accuracy of the global state information maintained at the nodes. Significant control overhead is incurred in gathering and maintaining global state information. On the other hand, if mobile nodes maintain only local state information (which is more accurate), distributed routing algorithms can be used. Even though control overhead incurred in maintaining local state information is low, care must be taken to obtain loop-free routes. In the case of the neither stateless approach, neither flow specific nor link specific state information is maintained at the nodes. Though the stateless approach solves the scalability problem permanently and reduces the burden (storage and computation) on nodes, providing QoS guarantees becomes extremely difficult.

• **Hard QoS versus soft QoS approach:** The QoS provisioning approaches can be broadly classified into two categories: *hard QoS* and *soft QoS* approaches. If QoS requirements of a connection are guaranteed to be met for the whole duration of the session, the QoS approach is termed a hard QoS approach. If the QoS requirements are not guaranteed for the entire session, the QoS approach is termed a soft QoS approach. Keeping network dynamics of ad hoc wireless networks in mind, it is very difficult to provide hard QoS guarantees to user applications. Thus, QoS guarantees can be given only within certain statistical bounds. Almost all QoS approaches available in the literature provide only soft QoS guarantees.

CLASSIFICATIONS OF QOS SOLUTIONS

The QoS solutions can be classified in two ways. One classification is based on the QoS approach employed, while the other one classifies QoS solutions based on the layer at which they operate in the network protocol stack.

Classifications of QoS Approaches

As shown in Figure 5.2, several criteria are used for classifying QoS approaches. The QoS approaches can be classified based on the interaction between the routing protocol and the QoS provisioning mechanism, based on the interaction between the network and the MAC layers, or based on the routing information update mechanism. Based on the interaction between the routing protocol and the QoS provisioning mechanism, QoS approaches can be classified into two categories: *coupled* and *decoupled* QoS approaches. In the case of the coupled QoS approach, the routing protocol and the QoS provisioning mechanism closely interact with each other for delivering QoS guarantees. If the routing protocol changes, it may fail to ensure QoS guarantees. But in the case of the decoupled approach, the QoS provisioning mechanism does not depend on any specific routing protocol to ensure QoS guarantees.

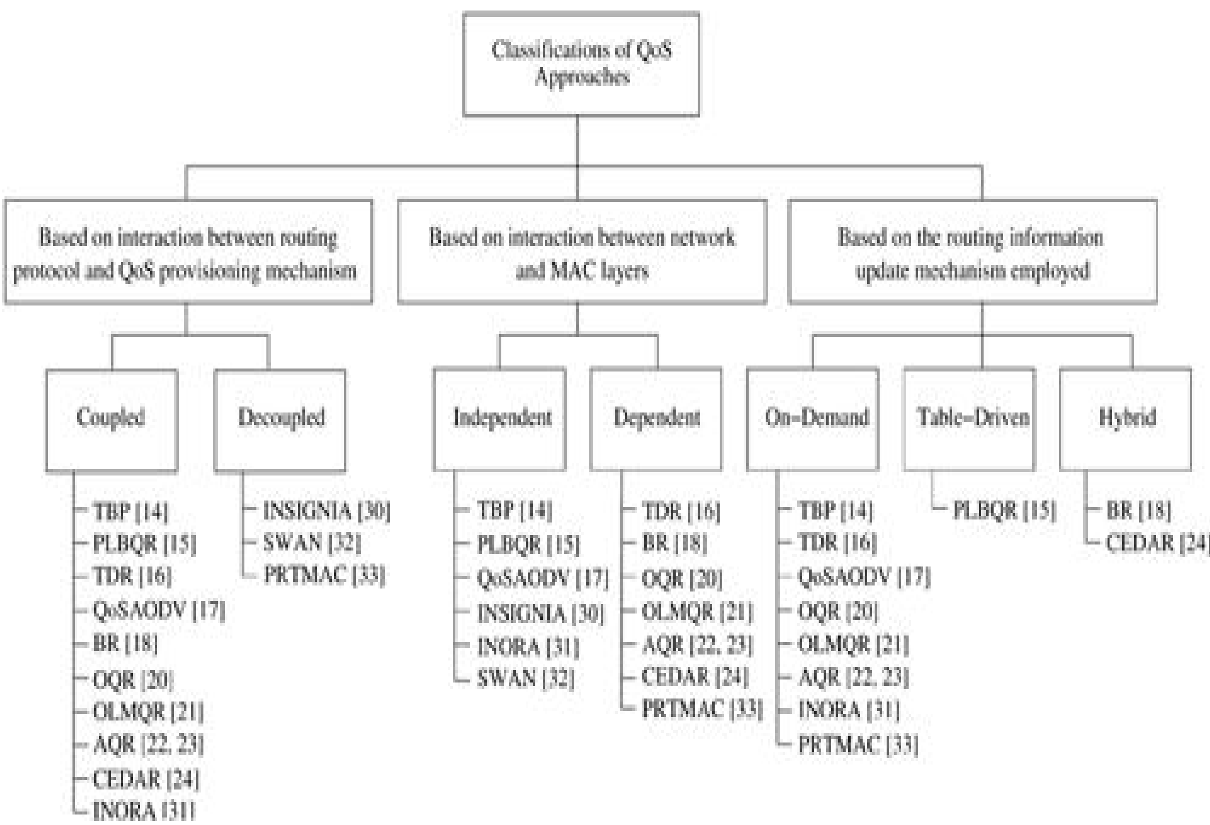

Figure 5.2. Classifications of QoS approaches

Similarly, based on the interaction between the routing protocol and the MAC protocol, QoS approaches can be classified into two categories: *independent* and *dependent* QoS approaches. In the independent QoS approach, the network layer is not dependent on the MAC layer for QoS provisioning. The dependent QoS approach requires the MAC layer to assist the routing protocol for QoS provisioning. Finally, based on the routing information update mechanism employed, QoS approaches can be classified into three categories, namely, *table-driven, on demand,* and *hybrid* QoS approaches. In the table-driven approach, each node in the network maintains a routing table which aids in forwarding packets. In the on-demand approach, no such tables are maintained at the nodes, and hence the source node has to discover the route on the fly. The hybrid approach incorporates features of both the table-driven and the on demand approaches.

Layer-Wise Classification of Existing QoS Solutions

The existing QoS solutions can also be classified based on which layer in the network protocol stack they operate in. Figure 5.3 gives a layer-wise classification of QoS solutions. The figure also shows some of the cross-layer QoS solutions proposed for adhoc wireless networks. The following sections describe the various QoS solutions listed in Figure 5.3.

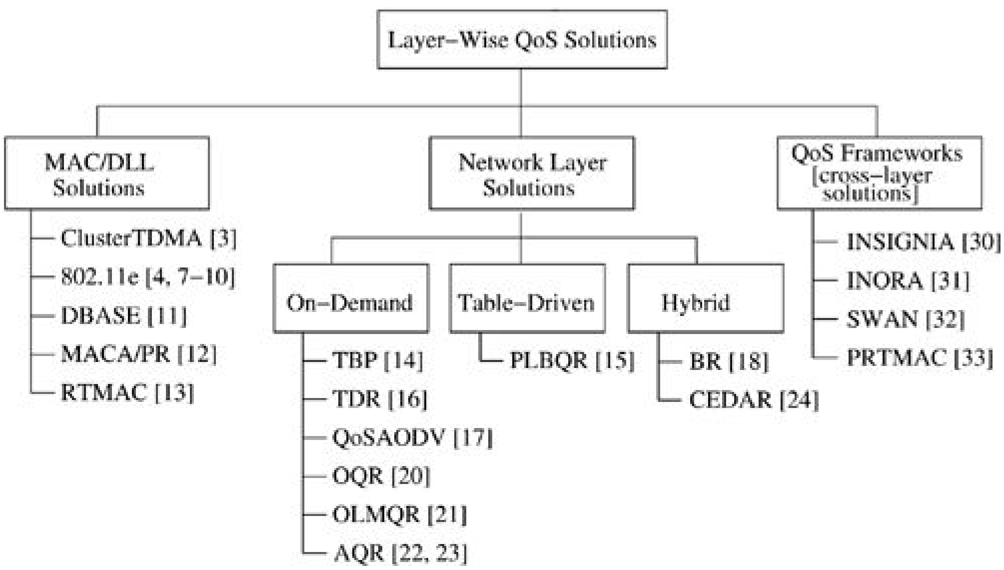

Figure 5.3 Layer-wise classification of QoS solutions

MAC LAYER SOLUTIONS

The MAC protocol determines which node should transmit next on the broadcast channel when several nodes are competing for transmission on that channel. The existing MAC protocols for ad hoc wireless networks use channel sensing and random back-off schemes, making them suitable for best-effort data traffic. Real-time traffic (such as voice and video) requires bandwidth guarantees. Supporting real-time traffic in these networks is a very challenging task.

In most cases, ad hoc wireless networks share a common radio channel operating in the ISM band2 or in military bands. The most widely deployed medium access technology is the IEEE 802.11 standard [2]. The 802.11 standard has two modes of operation: a distributed coordination function (DCF) mode and a point coordination function (PCF) mode. The DCF mode provides best-effort service, while the PCF mode has been designed to provide real-time traffic support in infrastructure-based wireless network configurations. Due to lack of fixed infrastructure support, the PCF mode of operation is ruled out in ad hoc wireless networks. Currently, the IEEE 802.11 Task Group e (TGe) is enhancing the legacy 802.11 standard to support real-time traffic. The upcoming 802.11e standard has two other modes of operation, namely, enhanced DCF (EDCF) and hybrid coordination function (HCF) to support QoS in both infrastructure-based and infrastructure-less network configurations. These two modes of operation are discussed later in this section. In addition to these standardized MAC protocols, several other MAC protocols that provide QoS support for applications in ad hoc wireless networks have been proposed. Some of these protocols are described below. 2 ISM refers to the industrial, scientific, and medical band. The frequencies in this band (from 2.4 GHz to 2.4835 GHz) are unlicensed.

Cluster TDMA

Gerla and Tsai proposed Cluster TDMA for supporting real-time traffic in ad hoc wireless networks. In bandwidth-constrained ad hoc wireless networks, the limited resources available need to be managed efficiently. To achieve this goal, a dynamic clustering scheme is used in Cluster TDMA. In this clustering approach, nodes are split into different groups. Each group has a cluster-head (elected by members of that group), which acts as a regional broadcast node and as a local coordinator to enhance the channel throughput. Every node within a cluster is one hop away from the cluster-head. The formation of clusters and selection of cluster-heads are done in a distributed manner. Clustering algorithms split the nodes into clusters so that they are interconnected and cover all the nodes. Three such algorithms used are lowest-ID algorithm, highest- degree (degree refers to the number of neighbors which are within transmission range of a node) algorithm, and least cluster change (LCC) algorithm. In the lowest-ID algorithm, a node becomes a cluster-head if it has the lowest ID among all its neighbors. In the highest-degree algorithm, a node with a degree greater than the degrees of all its neighbors becomes the cluster-head. In the LCC algorithm, cluster-head change occurs only if a change in network causes two cluster-heads to come into one cluster or one of the nodes moves out of the range of all the cluster-heads.

The time division multiple access (TDMA) scheme is used within a cluster for controlling access to the channel. Further, it is possible for multiple sessions to share a given TDMA slot via code division multiple access (CDMA). Across clusters, either spatial reuse of the timeslots or different spreading codes can be used to reduce the effect of inter-cluster interference. A synchronous time division frame is defined to support TDMA access within a cluster and to exchange control information. Each synchronous time division frame is divided into slots. Slots and frames are synchronized throughout the network. A frame is split into a control phase and a data phase. In the control phase, control functions such as frame and slot synchronization, routing, clustering, power management, code assignment, and virtual circuit (VC) setup are done. The cluster-head does the reservation for the VC by assigning the slot(s) and code(s) to be used for that connection. The number of slots per frame to be assigned to a VC is determined by the bandwidth requirement of the VC. Each station broadcasts the routing information it has, the ID of its cluster-head, the power gain3 list (the power gain list consists of the power gain values corresponding to each of the single-hop neighbors of the node concerned) it maintains, reservation status of the slots present in its data phase, and ACKs for frames that are received in the last data phase. Upon receiving this information, a node updates its routing table, calculates power gains for its neighbors, updates the power gain matrix, selects its cluster-head, records the slot reservation status of its neighbors, obtains ACKs for frames that are transmitted in the last data phase, and reserves slot(s). In each cluster, the corresponding cluster-head maintains a power gain matrix. The power gain matrix contains the power gain lists of all the nodes that belong to a particular cluster. It is useful for controlling the transmission power and the code division within a cluster.

IEEE 802.11e

In this section, the IEEE 802.11 MAC protocol is first described. Then, the recently proposed mechanisms for QoS support, namely, enhanced distributed coordination function (EDCF) and hybrid coordination function (HCF), defined in the IEEE 802.11e draft, are discussed.

IEEE 802.11 MAC Protocol

The 802.11 MAC protocol, describes how a station present in a WLAN should access the broadcast channel for transmitting data to other stations. It supports two modes of operation, namely, distributed coordination function (DCF) and point coordination function (PCF). The DCF mode does not use any kind of centralized control, while the PCF mode requires an access point (AP, *i.e.,* central controller) to coordinate the activity of all nodes in its coverage area. All implementations of the 802.11 standard for WLANs must provide the DCF mode of operation, while the PCF mode of operation is optional.

The time interval between the transmission of two consecutive frames is called the inter frame space (IFS). There are four IFSs defined in the IEEE 802.11 standard, namely, short IFS (SIFS), PCF IFS(PIFS), DCF IFS (DIFS), and extended IFS (EIFS). The relationship among them is as follows:

$$SIFS < PIFS < DIFS < EIFS$$

DBASE

In an ad hoc WLAN, there is no fixed infrastructure (*i.e.,* AP) to coordinate the activity of individual stations. The stations are part of a single-hop wireless network and contend for the broadcast channel in a distributed manner. For real-time traffic (*rt-traffic*), a contention- based process is used in order to gain access to the channel. Once a station gains channel access, a reservation-based process is used to transmit the subsequent frames. The non-real-time stations (*nrt*-stations) regulate their accesses to the channel according to the standard CSMA/CA protocol used in 802.11 DCF. The DBASE protocol permits real-time stations (*rt*-stations) to acquire excess bandwidth on demand. It is still compliant with the IEEE 802.11 standard. Like the IEEE 802.11 standard, the DBASE protocol divides the frames into three priority classes. Frames belonging to different priority classes have to wait for different IFSs before they are transmitted. Stations have to wait for a minimum of PIFS before transmitting *rt* frames such as reservation frame (RF) and request-to-send (RTS). The *nrt*-frames have the lowest priority, and hence stations have to wait for DIFS before transmitting such frames.

NETWORK LAYER SOLUTIONS

The bandwidth reservation and real-time traffic support capability of MAC protocols can ensure reservation at the link level only, hence the network layer support for ensuring end-to-end resource negotiation, reservation, and reconfiguration is very essential. This section describes the existing network layer solutions that support QoS provisioning.

QoS Routing Protocols

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 module which assists the QoS routing protocol in its search for QoS feasible paths. The QoS routing protocol should find paths that consume minimum resources. The QoS metrics can be classified as additive metrics, concave metrics, and multiplicative metrics.

Ticket-Based QoS Routing Protocol

Ticket-based QoS routing is a distributed QoS routing protocol for ad hoc wireless networks. This protocol has the following features:

- It can tolerate imprecise state information during QoS route computation and exhibits good performance even when the degree of imprecision is high.
- It probes multiple paths in parallel for finding a QoS feasible path. This increases the chance of finding such a path. The number of multiple paths searched is limited by the number of tickets issued in the probe packet by the source node. State information maintained at intermediate nodes is used for more accurate route probing. An intelligent hop-by-hop selection mechanism is used for finding feasible paths efficiently.
- The optimality of a path among several feasible paths is explored. A low-cost path that uses minimum resources is preferred when multiple feasible paths are available.
- A primary-backup-based fault-tolerant technique is used to reduce service disruption during path breaks that occur quite frequently in ad hoc wireless networks.

Predictive Location-Based QoS Routing Protocol

The predictive location-based QoS routing protocol (PLBQR) is based on the prediction of the location of nodes in ad hoc wireless networks. The prediction scheme overcomes to some extent the problem arising due to the presence of stale routing information. No resources are reserved along the path from the source to the destination, but QoS-aware admission control is performed. The network does its best to support the QoS requirements of the connection as specified by the application. The QoS routing protocol takes the help of an update protocol and location and delay prediction schemes. The update protocol aids each node in broadcasting its geographic location and resource information to its neighbors. Using the update messages received from the neighbors, each node updates its own view of the network topology. The update protocol has two types of update messages, namely, *Type 1 update* and *Type 2 update.* Each node generates a Type 1 update message periodically. A Type 2 update message is generated when there is a considerable change in the node's velocity or direction of motion. From its recent update messages, each node can calculate an expected geographical location where it should be located at a particular instant and then periodically checks if it has deviated by a distance greater than δ from this expected location. If it has deviated, a Type 2 update message is generated.

Trigger-Based Distributed QoS Routing Protocol

The trigger-based (on-demand) distributed QoS routing (TDR) protocol was proposed by De *et al.* for supporting real-time applications in ad hoc wireless networks. It operates in a distributed fashion. Every node maintains only the local neighborhood information in order to reduce computation overhead and storage overhead. To reduce control overhead, nodes maintain only the active routes. When a link failure is imminent, TDR utilizes the global positioning system-based (GPS) location information of the destination to localize the reroute queries only to certain neighbors of the nodes along the source-to-destination active route. For a quick rerouting with reduced control overhead, rerouting is attempted from the location of an imminent link failure, called intermediate node-initiated rerouting (INIR). If INIR fails, then in order to keep the flow state disruption to a minimum, rerouting is attempted from the source, which is termed source-initiated rerouting (SIRR).

QoS-Enabled Ad Hoc On-Demand Distance Vector Routing Protocol

Perkins *et al.* have extended the basic ad hoc on-demand distance vector (AODV) routing protocol to provide QoS support in ad hoc wireless networks. To provide QoS, packet formats have been modified in order to specify the service requirements which must be met by the nodes forwarding a *RouteRequest* or a *RouteReply.*

Bandwidth Routing Protocol

The bandwidth routing (BR) protocol [18] consists of an end-to-end path bandwidth calculation algorithm to inform the source node of the available bandwidth to any destination in the ad hoc network, a bandwidth reservation algorithm to reserve a sufficient number of free slots for the QoS flow, and a standby routing algorithm to reestablish the QoS flow in case of path breaks. Here, only bandwidth is considered to be the QoS parameter. In TDMA-based networks, bandwidth is measured in terms of the number of free slots available at a node. The goal of the bandwidth routing algorithm is to find a shortest path satisfying the bandwidth requirement. The transmission time scale is organized into frames, each containing a fixed number of time-slots. The entire network is synchronized on a frame and slot basis. Each frame is divided into two phases, namely, the control phase and the data phase. The control phase is used to perform the control functions such as slot and frame synchronization, virtual circuit (VC) setup, and routing. The data phase is used for transmission/reception of data packets. For each node, a slot is assigned in the control phase for it to broadcast its routing information and slot requirements. At the end of the control phase, each node knows about the channel reservations made by its neighbors. This information helps nodes to schedule free slots, verify the failure of reserved slots, and drop expired real-time packets. The BR protocol assumes a half-duplex CDMA-over-T DMA system in which only one packet can be transmitted in a given slot.

On-Demand QoS Routing Protocol

Lin proposed an admission control scheme over an on-demand QoS routing (OQR) protocol to guarantee bandwidth for real-time applications. Since routing is on-demand in nature, there is no need to exchange control information periodically and maintain routing tables at each node. Similar to the bandwidth routing (BR) protocol, the network is time-slotted and bandwidth is the key QoS parameter. The path bandwidth calculation algorithm proposed in BR is used to measure the available end-to-end bandwidth. The on-demand QoS routing protocol is explained below.

On-Demand Link-State Multipath QoS Routing Protocol

Unlike the QoS routing protocols described above in this chapter which try to find a single path from the source to the destination satisfying the QoS requirements, the on-demand link-state multipath QoS routing (OLMQR) protocol searches for multiple paths which collectively satisfy the required QoS. The original bandwidth requirement is split into sub-bandwidth requirements. Notably, the paths found by the multipath routing protocol are allowed to share the same sub-paths. OLMQR has better call acceptance rate in ad hoc wireless networks where finding a single path satisfying all the QoS requirements is very difficult.

Asynchronous Slot Allocation Strategies

The QoS solutions discussed so far such as BR, OQR, and OLMQR assume a TDMA-based network or a CDMA-over-T DMA model for the network. This requires time synchronization across all nodes in the network. Time synchronization demands periodic exchange of control packets, which results in high bandwidth consumption. Ad hoc wireless networks experience rapid changes in topology leading to a situation where network partitions and merging of partitions can take place. Figure 5.4 shows the synchronization problems arising out of dynamic topological changes in an ad hoc wireless network. A completely connected and synchronized network A at time $t = t0$ (shown in Figure 5.4 (a)) may be partitioned into two disjoint networks A1 and A2 at time $t = t1$ (shown in Figure 5.4 (b)). These two networks may be synchronized to two different clock times as illustrated. Due to the dynamic topology experienced in an ad hoc wireless network, it is possible to have two separately synchronized networks A1 (synchronized to tA 1) and A2 (synchronized to tA 2) merge to form a combined network A (Figure 5.4 (c)). During the merging process, the real-time calls existing in the network may be affected while accommodating the changes in synchronization.

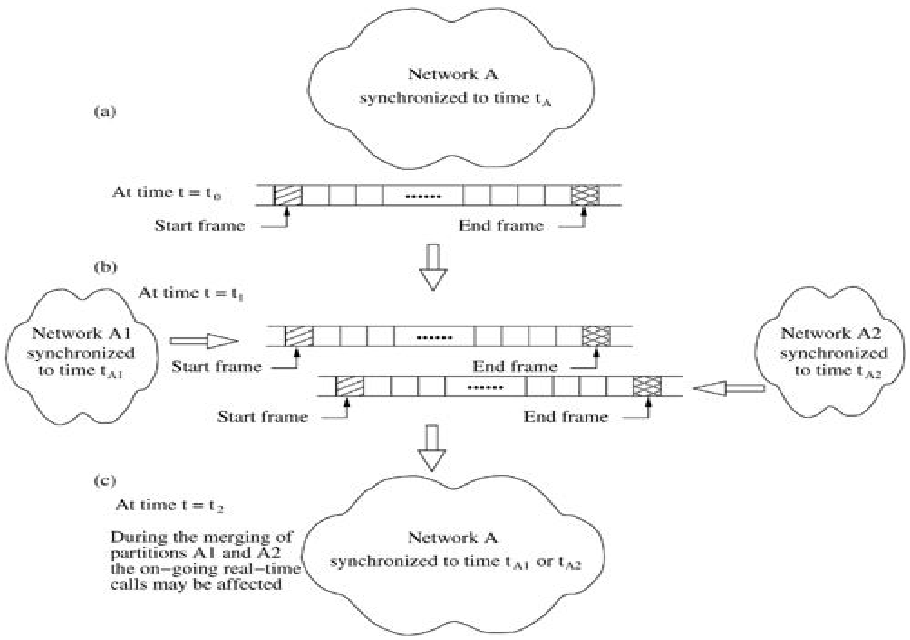

Figure 5.4. Illustration of synchronization problems in a dynamic network topology

QOS FRAMEWORKS FOR AD HOC WIRELESS NETWORKS

A framework for QoS is a complete system that attempts to provide required/promised services to each user or application. All components within this system cooperate in providing the required services. The key component of any QoS framework is the QoS service model which defines the way user requirements are met. The key design issue here is whether to serve users on a per session basis or on a per class basis. Each class represents an aggregation of users based on certain criteria. The other key components of the framework are QoS routing which is used to find all or some of the feasible paths in the network that can satisfy user requirements, QoS signaling for resource reservation, QoS medium access control, call admission control, and packet scheduling schemes. The QoS modules, namely, routing protocol, signaling protocol, and the resource management mechanism, should react promptly to changes in the network state (topology changes) and flow state (change in the end-to-end view of the service delivered). In what follows, each component's functionality and its role in providing QoS in ad hoc wireless networks will be described.

• ***Routing protocol***: Similar to the QoS routing protocols, discussed earlier in this chapter, the routing protocol module in any QoS framework is used to find a path from the source to the destination and to forward the data packet to the next intermediate relay node. QoS routing describes the process of finding suitable path(s) that satisfy the QoS service requirements of an application. If multiple paths are available, the information regarding such paths helps to restore the service quickly when the service becomes disturbed due to a path break. The performance of the routing protocol, in terms of control overhead, affects the performance of the QoS framework. The routing protocol should be able to track changes in the network topology with minimum control overhead. The routing protocol needs to work efficiently with other components of the QoS framework such as signaling, admission control, and resource management mechanisms in order to provide end-to-end QoS guarantees. These mechanisms should consume minimal resources in operation and react rapidly to changes in the network state and the flow state.

• **QoS resource reservation signaling**: Once a path with the required QoS is found, the next step is to reserve the required resources along that path. This is done by the resource reservation signaling protocol. For example, for applications that require certain minimum bandwidth guarantees, signaling protocol communicates with the medium access control subsystem to find and reserve the required bandwidth. On completion/termination of a session, the previously reserved resources are released.

• **Admission control**: Even though a QoS feasible path may be available, the system needs to decide whether to actually serve the connection or not. If the call is to be served, the signaling protocol reserves the resources; otherwise, the application is notified of the rejection. When a new call is accepted, it should not jeopardize the QoS guarantees given to the already admitted calls. A QoS framework is evaluated based on the number of QoS sessions it serves and it is represented by the average call acceptance ratio (ACAR) metric. Admission control ensures that there is no perceivable degradation in the QoS being offered to the QoS sessions admitted already.

• **Packet scheduling**: When multiple QoS connections are active at the same time through a link, the decision on which QoS flow is to be served next is made by the scheduling scheme. For example, when multiple delay-constrained sessions are passing through a node, the scheduling mechanism decides on when to schedule the transmission of packets when packets belonging to more than one session are pending in the transmission queue of the node. The performance of a scheduling scheme is reflected by the percentage of packets that meet their deadlines.

ENERGY MANAGEMENT IN AD HOC WIRELESS NETWORKS

The nodes in an ad hoc wireless network are constrained by limited battery power for their operation. Hence, energy management is an important issue in such networks. The use of multi-hop radio relaying requires a sufficient number of relaying nodes to maintain the network connectivity. Hence, battery power is a precious resource that must be used efficiently in order to avoid early termination of any node.

Energy management deals with the process of managing energy resources by means of controlling the battery discharge, adjusting the transmission power, and scheduling of power sources so as to increase the lifetime of the nodes of an ad hoc wireless network. Efficient battery management, transmission power management, and system power management are the three major means of increasing the life of a node. Battery management is concerned with problems that lie in the selection of battery technologies, finding the optimal capacity of the battery, and scheduling of batteries, that increase the battery capacity. Transmission power management techniques attempt to find an optimum power level for the nodes in the ad hoc wireless network. On the other hand, system power management deals mainly with minimizing the power required by hardware peripherals of a node (such as CPU, DRAM, and LCD display) and incorporating low-power strategies into the protocols used in various layers of the protocol stack. This chapter concentrates on the issues involved and the solutions for energy management in ad hoc wireless networks.

NEED FOR ENERGY MANAGEMENT IN AD HOC WIRELESS NETWORKS

The energy efficiency of a node is defined as the ratio of the amount of data delivered by the node to the total energy expended. Higher energy efficiency implies that a greater number of packets can be transmitted by the node with a given amount of energy reserve. The main reasons for energy management in ad hoc wireless networks are listed below:

• **Limited energy reserve:** The main reason for the development of ad hoc wireless networks is to provide a communication infrastructure in environments where the setting up of a fixed infrastructure is impossible. Ad hoc wireless networks have very limited energy resources. Advances in battery technologies have been negligible as compared to the recent advances that have taken place in the field of mobile computing and communication. The increasing gap between the power consumption requirements and power availability adds to the importance of energy management.

• **Difficulties in replacing the batteries:** Sometimes it becomes very difficult to replace or recharge the batteries. In situations such as battlefields, this is almost impossible. Hence, energy conservation is essential in such scenarios.

• **Lack of central coordination:** The lack of a central coordinator, such as the base station in cellular networks, introduces multi-hop routing and necessitates that some of the intermediate nodes act as relay nodes. If the proportion of relay traffic is large, then it may lead to a faster depletion of the power source for that node. On the other hand, if no relay traffic is allowed through a node, it may lead to partitioning of the network. Hence, unlike other networks, relay traffic plays an important role in ad hoc wireless networks.

• **Constraints on the battery source:** Batteries tend to increase the size and weight of a mobile node. Reducing the size of the battery results in less capacity which, in turn, decreases the active lifespan of the node. Hence, in addition to reducing the size of the battery, energy management techniques are necessary to utilize the battery capacity in the best possible way.

• **Selection of optimal transmission power:** The transmission power selected determines the reachability of the nodes. The consumption of battery charge increases with an increase in the transmission power. An optimal value for the transmission power decreases the interference among nodes, which, in turn, increases the number of simultaneous transmissions.

• **Channel utilization:** A reduction in the transmission power increases frequency reuse, which leads to better channel reuse. Power control becomes very important for CDMA based systems in which the available bandwidth is shared among all the users. Hence, power control is essential to maintain the required signal to interference ratio (SIR) at the receiver and to increase the channel reusability.

CLASSIFICATION OF ENERGY MANAGEMENT SCHEMES

The need for energy management in ad hoc wireless networks, discussed in the previous section, points to the fact that energy awareness needs to be adopted by the protocols at all the layers in the protocol stack, and has to be considered as one of the important design objectives for such protocols. Energy conservation can be implemented using the following techniques:

- Battery management schemes
- Transmission power management schemes
- System power management schemes

Maximizing the life of an ad hoc wireless network requires an understanding of the capabilities and the limitations of energy sources of the nodes. A greater battery capacity leads to a longer lifetime of the nodes. Increasing the capacity of the batteries can be achieved by taking into consideration either the internal characteristics of the battery (battery management) or by minimizing the activities that utilize the battery capacity (power management). The system power management approach can be further divided into the following categories:

- Device management schemes
- Processor power management schemes

Figure 5.5 provides an overview of some of the techniques at different layers of the protocol stack that fall into three categories: battery management, transmission power management, and system power management schemes. Though these schemes cannot be strictly classified under the different layers of the OSI protocol stack as they reside in more than one layer, the classification provided in this section is based on the highest layer in the protocol stack used by each of these protocols.

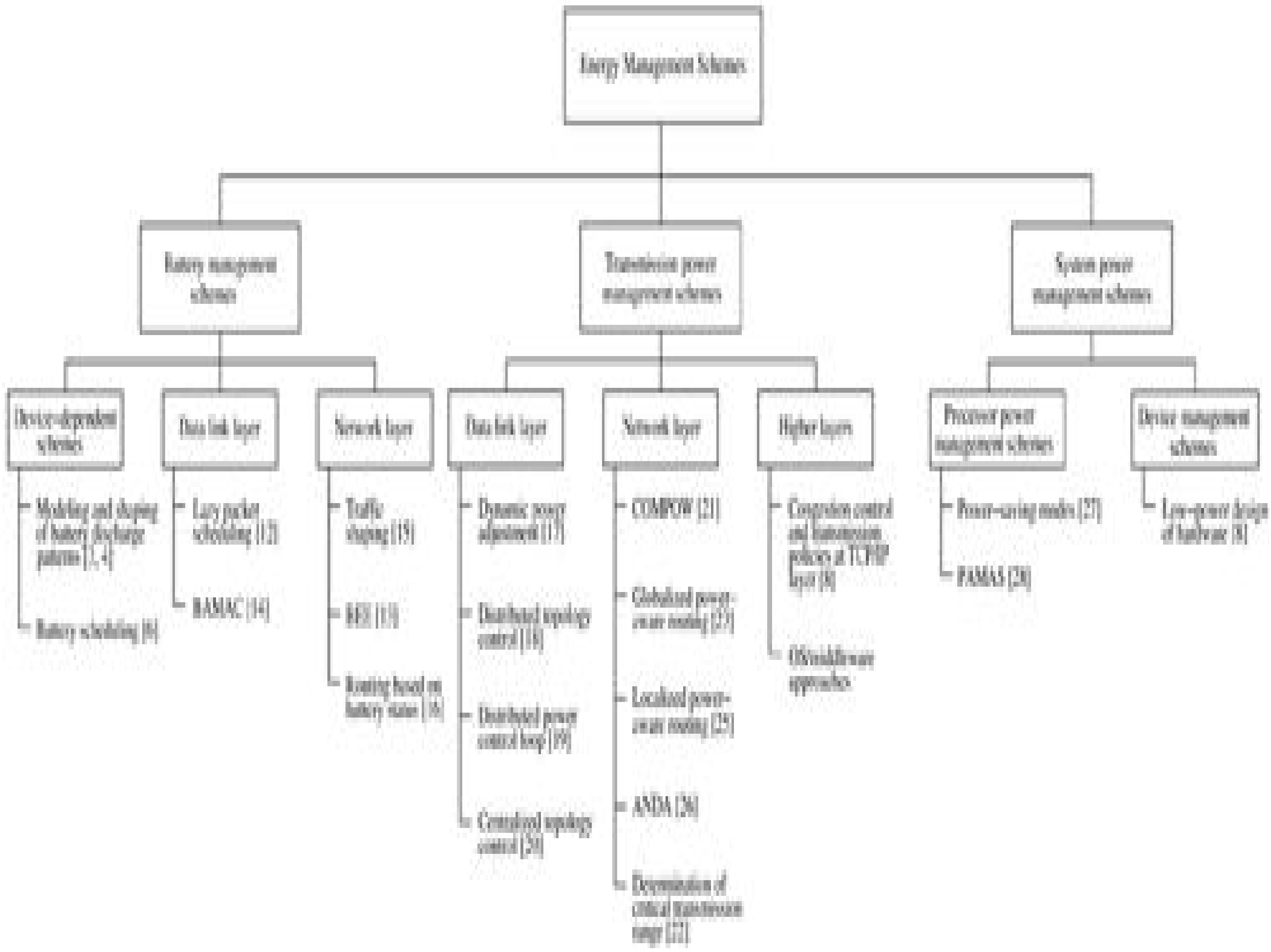

Figure 5.5. Classification of energy management schemes.

BATTERY MANAGEMENT SCHEMES

Battery-driven systems are those systems which are designed taking into consideration mainly the battery and its internal characteristics. They try to maximize the amount of energy provided by the power source by exploiting the inherent property of batteries to recover their charge when kept idle. Recent research results in this area have proved that, by varying the manner in which energy is drawn from the batteries, significant improvement can be obtained in the total amount of energy supplied by them. In the section that follows, we also discuss some of the battery characteristics which are used throughout our discussions on battery management.

Overview of Battery Characteristics

The major components of batteries are illustrated in Figure 5.6 A battery mainly consists of an anode, a cathode, an electrolyte medium, and a case. The anode is often a metal and the cathode a metallic oxide. The electrolyte is a salt solution that promotes the ion flow. The porous separator is used to prevent a short circuit between anode and cathode by keeping them from touching one another. The battery is contained in a structural support (case) that provides dimensional stability and a positive and a negative electrode for discharging (or recharging) the cell. The positive ions move from the anode toward the cathode through the electrolyte medium and the electrons flow through the external circuit. A number of separate electrochemical cells can also be combined within the same case to create a battery.

• **Battery technologies:** The most popular rechargeable battery technologies developed over the last two decades are comprised of nickel-cadmium, lithium ion, nickel metal hydride, reusable alkaline, and lithium polymer. The main factors considered while designing a battery technology are the energy density (the amount of energy stored per unit weight of the battery), cycle life [the number of (re)charge cycles prior to battery disposal], environmental impact, safety, cost, available supply voltage, and charge/discharge characteristics.

• **Principles of battery discharge:** A battery typically consists of an array of one or more cells. Hence, in the subsequent sections, the terms "battery" and "cell" are used interchangeably. The three main voltages that characterize a cell are: (1) the open circuit voltage (V_{oc}), that is, the initial voltage under a no-load condition of a fully charged cell, (2) the operating voltage (V_i), that is, the voltage under loaded conditions, and (3) the cut-off voltage (V_{cut}) at which the cell is said to be discharged. All the cells are defined by three main capacities:

– **Theoretical capacity**: The amount of active materials (the materials that react chemically to produce electrical energy when the cell is discharged and restored when the cell is charged) contained in the cell refers to its theoretical capacity. A cell cannot exceed its theoretical capacity.

– **Nominal (standard) capacity**: This corresponds to the capacity actually available when discharged at a specific constant current. It is expressed in ampere-hours.

– **Actual capacity**: The energy delivered under a given load is said to be the actual capacity of the cell. A cell may exceed the actual capacity but not the theoretical capacity.

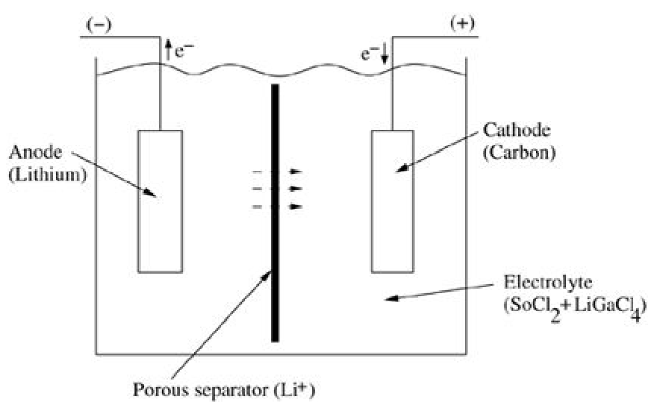

Figure 5.6 Basic structure of a lithium/thionyl chloride battery

The constant current discharge behavior of lithium-manganese dioxide ($LiMnO_2$) cells with $V_{oc} = 3V$ and $V_{cut} = 1V$ is shown in Figure.5.7 The discharge curve is flat most of the time and a gradual slope is developed as the voltage reaches the cut-off voltage. The performance of a cell's discharge is measured using the following parameters:

– **Discharge time**: The time elapsed when a fully charged cell reaches its cut-off voltage and has to be replaced or recharged is called the discharge time of the cell.

– **Specific power (energy):** This is the power (energy) delivered by a fully charged cell under a specified discharge current. It is expressed in watt-per- kilogram (watt-hour-per-kilo gram).

– Discharge current: There are mainly two models of battery discharge: constant current discharge and pulsed current discharge. In pulsed current discharge, the battery switches between short discharge periods and idle periods (rest periods). Chiasserini and Rao in illustrate the performance of the bipolar lead-acid battery subjected to a pulsed discharge current of six current pulses. After each discharge, which lasts for 3 ms, the cell was idled for 22 ms during which no recharging was allowed to take place. Figure 5.7 shows the current density and the corresponding cell voltage. The cell is able to recover and revert to its initial open circuit voltage during the first four rest periods. After the fifth current pulse, the rest period of 22 ms turns out to be inadequate for the cell recovery.

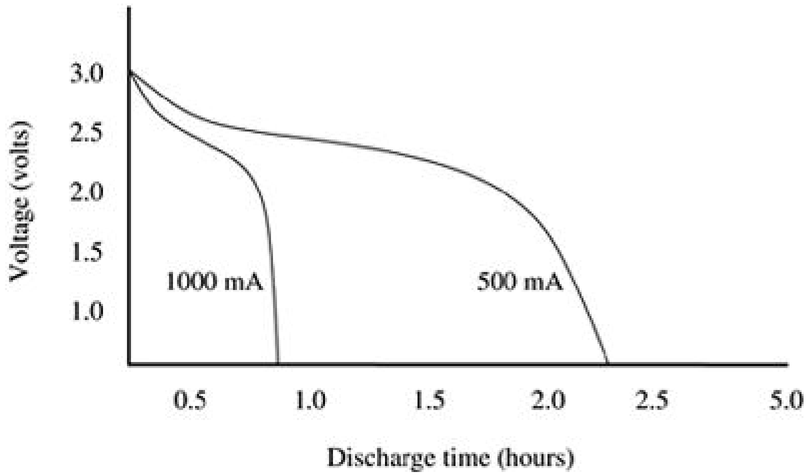

Figure 5.7. Discharge pattern of a cell when *Voc =* 3 *V* and *Vcut* = 1 *V*.

• **Impact of discharge characteristics on battery capacity:** The important chemical processes that affect the battery characteristics are given below.
– Diffusion process: When the battery is actively involved in discharging, that is, at a nonzero current, the active materials move from the electrolyte solution to the electrodes and are consumed at the electrode. If this current is above a threshold value called the *limiting current,* the active materials get depleted very quickly. But as the current decreases, the concentration of the active materials around the electrode drops. By increasing the rest time periods of the battery, longer lifetimes can be achieved due to the recovery capacity effect, which is explained later in this section. In the following discussion, we will concentrate on some of the battery management techniques which increase idle periods for batteries.
– Passivation process: The cell discharge is limited not only by the diffusion process but also by a process called *passivation,* which induces in the cell the precipitation of crystals which are produced by the discharge due to the chemical reactions on the electrode. This phenomenon increases during higher current densities.
Two important effects to be considered for understanding the battery's discharge properties are stated below.
– **Rate capacity effect**: As the intensity of the discharge current increases, an insoluble component develops between the inner and outer surfaces of the cathode. The inner layer becomes inaccessible as a result of this phenomenon, rendering the cell unusable even while a sizable amount of active materials still exists. This effect depends on the actual capacity of the cell and the discharge current. – Recovery capacity effect: This effect is concerned with the recovery of charges under idle conditions. By increasing the idle time, one may be able to completely utilize the theoretical capacity of the cell. • Battery models: Battery models depict the characteristics of the batteries used in real life. The pros and cons of following battery models are summarized in[8]: analytical models, stochastic models, electric circuit models, and electrochemical models. Finally, battery efficient system architecture is proposed and the following approaches are suggested to enable longer life of the nodes of an ad hoc wireless network

PART- A

1. What are the limitations of IEEE 802.11 MAC protocol that prevent from supporting QOS traffic?
2. Compare IEEE 802.1e with DBASE protocol.
3. Which battery is being commonly used for portable mobile nodes such as laptops? Give few reasons to support your answer.
4. Suggest a few metrics that can be associated with battery –aware routing techniques.
5. What are disadvantages of clustering in adhoc wireless networks?
6. What are the advantages of distributed power control algorithms in adhoc wireless networks over centralized power control algorithms?

PART- B

1. Explain the need for energy management in MANET.
2. Discuss in detail about battery management schemes.
3. Explain the issues and challenges in provisioning QOS.
4. Explain the classification of QOS solutions.
5. Explain the QOs frameworks for adhoc wireless networks.