# A Survey of Energy Efficient Network Protocols for Wireless Networks

Christine E. Jones[1], Krishna M. Sivalingam[2, †], Prathima Agrawal[3] and Jyh-Cheng Chen[3]
[1] *BBN Technologies, Cambridge, MA 02138*
[2] School of EECS, Washington State University, Pullman, WA 99164
[3] *Telcordia Technologies, Morristown, NJ 07960*

January 16, 2001

**Abstract.**
Wireless networking has witnessed an explosion of interest from consumers in recent years for its applications in mobile and personal communications. As wireless networks become an integral component of the modern communication infrastructure, energy efficiency will be an important design consideration due to the limited battery life of mobile terminals. Power conservation techniques are commonly used in the hardware design of such systems. Since the network interface is a significant consumer of power, considerable research has been devoted to low-power design of the entire network protocol stack of wireless networks in an effort to enhance energy efficiency. This paper presents a comprehensive summary of recent work addressing energy efficient and low-power design within all layers of the wireless network protocol stack.

**Keywords:** Wireless networks, Mobile computing, Energy efficient design, Network protocols, Power aware protocols, Low-power design

## 1. Introduction

The rapid expansion of wireless services such as cellular voice, PCS (Personal Communications Services), mobile data and wireless LANs in recent years is an indication that significant value is placed on accessibility and portability as key features of telecommunication (Salkintzis and Mathiopoulos (Guest Ed.), 2000). Wireless devices have maximum utility when they can be used "anywhere at anytime". One of the greatest limitations to that goal, however, is finite power supplies. Since batteries provide limited power, a general constraint of wireless communication is the short continuous operation time of mobile terminals. Therefore, power management is

† Corresponding Author: Dr. Krishna Sivalingam. Part of the research was supported by Air Force Office of Scientific Research grants F-49620-97-1-0471 and F-49620-99-1-0125; by Telcordia Technologies and by Intel. Part of the work was done while the first author was at Washington State University. The authors' can be reached at `cej@bbn.com`, `krishna@eecs.wsu.edu`, `pagrawal@research.telcordia.com`, `jcchen@research.telcordia.com`

one of the most challenging problems in wireless communication, and recent research has addressed this topic (Bambos, 1998). Examples include a collection of papers available in (Zorzi (Guest Ed.), 1998) and a recent conference tutorial (Srivastava, 2000), both devoted to energy efficient design of wireless networks.

Studies show that the significant consumers of power in a typical laptop are the microprocessor (CPU), liquid crystal display (LCD), hard disk, system memory (DRAM), keyboard/mouse, CDROM drive, floppy drive, I/O subsystem, and the wireless network interface card (Udani and Smith, 1996, Stemm and Katz, 1997). A typical example from a Toshiba 410 CDT mobile computer demonstrates that nearly 36% of power consumed is by the display, 21% by the CPU/memory, 18% by the wireless interface, and 18% by the hard drive. Consequently, energy conservation has been largely considered in the hardware design of the mobile terminal (Chandrakasan and Brodersen, 1995) and in components such as CPU, disks, displays, etc. Significant additional power savings may result by incorporating low-power strategies into the design of network protocols used for data communication. This paper addresses the incorporation of energy conservation at all layers of the protocol stack for wireless networks.

The remainder of this paper is organized as follows. Section 2 introduces the network architectures and wireless protocol stack considered in this paper. Low-power design within the physical layer is briefly discussed in Section 2.3. Sources of power consumption within mobile terminals and general guidelines for reducing the power consumed are presented in Section 3. Section 4 describes work dealing with energy efficient protocols within the MAC layer of wireless networks, and power conserving protocols within the LLC layer are addressed in Section 5. Section 6 discusses power aware protocols within the network layer. Opportunities for saving battery power within the transport layer are discussed in Section 7. Section 8 presents techniques at the OS/middleware and application layers for energy efficient operation. Finally, Section 9 summarizes and concludes the paper.

## 2. Background

This section describes the wireless network architectures considered in this paper. Also, a discussion of the wireless protocol stack is included along with a brief description of each individual protocol layer. The physical layer is further discussed.

### 2.1. Wireless Network Architectures

Two different wireless network architectures are considered in this paper: *infrastructure* and *ad hoc* networks. Below, a description of each system architecture is presented.

2.1.0.1. *Infrastructure:* Wireless networks often extend, rather than replace, wired networks, and are referred to as infrastructure networks. The infrastructure network architecture is depicted in Figure 1. A hierarchy of wide area and local area wired networks is used as the backbone network. The wired backbone connects to special switching nodes called *base stations*. Base stations are often conventional PCs and workstations equipped with custom wireless adapter cards. They are responsible for coordinating access to one or more transmission channel(s) for mobiles located within the coverage cell. Transmission channels may be individual frequencies in FDMA (Frequency Division Multiple Access), time slots in TDMA (Time Division Multiple Access), or orthogonal codes or hopping patterns in the case of CDMA (Code Division Multiple Access). Therefore, within infrastructure networks, wireless access to and from the wired host occurs in the last hop between base stations and mobile hosts that share the bandwidth of the wireless channel.

2.1.0.2. *Ad hoc:* Ad hoc networks, on the other hand, are multi-hop wireless networks in which a set of mobiles cooperatively maintain network connectivity (Macker and Corson, 1998). This on-demand network architecture is completely un-tethered from physical wires. An example of an ad hoc topology is pictured in Figure 2. Ad hoc networks are characterized by dynamic, unpredictable, random, multi-hop topologies with typically no infrastructure support. The mobiles must periodically exchange topology information which is used for routing updates. Ad hoc networks are helpful in situations in which temporary network connectivity is needed, and are often used for military environments, disaster relief, and so on. Mobile ad hoc networks have attracted considerable attention as evidenced by the IETF working group MANET (Mobile Ad hoc Networks). This has produced various Internet drafts, RFCs, and other publications (Macker and Corson, 1998, Macker and Corson, ). Also, a recent conference tutorial presents a good introduction to ad hoc networks (Vaidya, 2000).

## 2.2. Protocol Layers

This section provides an introduction to the software used in wireless data network systems. Application programs using the network do not interact directly with the network hardware. Instead, an application interacts with the protocol software. The notion of protocol layering provides a conceptual basis for understanding how a complex set of protocols work together with the hardware to provide a powerful communication system. Recently, communication protocol stacks such as the Infrared Data Association (IrDa) protocol stack for point-to-point wireless infrared communication and the Wireless Application Protocol (WAP) Forum protocol stack for enabling developers to build advanced services across differing wireless network technologies (Infrared Data Association, 2000, WAP Forum, 2000) have been developed specifically for wireless networks. This paper focuses on the traditional OSI protocol stack, depicted in Figure 3, for a generic wireless communication system. The application and services layer occupies the top of the stack followed by the operating system/middleware, transport, network, data link, and physical layers. The problems inherent to the wireless channel and issues related to mobility challenge the design of the protocol stack adopted for wireless networks. In addition, networking protocols need to be designed with energy efficiency in mind.

*Physical:* The physical layer consists of Radio Frequency (RF) circuits, modulation, and channel coding systems. From an energy efficient perspective, considerable attention has already been given to the design of this layer (Chandrakasan and Brodersen, 1995).

*Data Link:* The data link layer is responsible for establishing a reliable and secure logical link over the unreliable wireless link. The data link layer is thus responsible for wireless link error control, security (encryption/decryption), mapping network layer packets into frames, and packet retransmission.

A sublayer of the data link layer, the media access control (MAC) protocol layer is responsible for allocating the time-frequency or code space among mobiles sharing wireless channels in a region.

*Network:* The network layer is responsible for routing packets, establishing the network service type (connection-less versus connection-oriented), and transferring packets between the transport and link layers. In a mobile environment this layer has the added responsibility of rerouting packets and mobility management.

*Transport:* The transport layer is responsible for providing efficient and reliable data transport between network endpoints independent of the physical network(s) in use.

*OS/Middleware:* The operating system and middleware layer handles disconnection, adaptivity support, and power and quality of service (QoS) management within wireless devices. This is in addition to the conventional tasks such as process scheduling and file system management.

*Application:* The application and services layer deals with partitioning of tasks between fixed and mobile hosts, source coding, digital signal processing, and context adaptation in a mobile environment. Services provided at this layer are varied and application specific.

The next section further examines the low-power research completed at the physical layer.

### 2.3. Physical Layer

In the past, energy efficient and low-power design research has centered around the physical layer due to the fact that the consumption of power in a mobile computer is a direct result of the system hardware. Research addresses two different perspectives of the energy problem: (i) an increase in battery capacity and (ii) a decrease in the amount of energy consumed at the wireless terminal.

The primary problem concerning energy in wireless computing is that battery capacity is extremely limited. The focus of battery technology research has been to increase battery power capacity while restricting the weight of the battery. However, unlike other areas of computer technology such as micro-chip design, battery technology has not experienced significant advancement in the past 30 years. Therefore, unless a breakthrough occurs in battery technology, an attainable goal of research would be a decrease in the energy consumed in the wireless terminal (Lettieri and Srivastava, 1999).

Low-power design at the hardware layer uses different techniques including variable clock speed CPUs (Govil et al., 1995), flash memory (Marsh et al., 1994), and disk spindown (Douglis et al., 1994). Numerous energy efficient techniques for the physical layer are discussed in (Chandrakasan and Brodersen, 1995). Although the above techniques have resulted in considerable energy savings, other venues should also be explored to improve energy efficiency. One way to achieve this for future wireless networks is to design the higher layers of the protocol stack with energy efficiency as an important goal.

## 3. Power Consumption Sources and Conservation Mechanisms

This section first presents the chief sources of power consumption with respect to the protocol stack. Then, it presents an overview of the main mechanisms and principles that may be used to develop energy efficient network protocols.

### 3.1. Sources of Power Consumption

The sources of power consumption, with regard to network operations, can be classified into two types: communication related and computation related.

Communication involves usage of the transceiver at the source, intermediate (in the case of ad hoc networks), and destination nodes. The transmitter is used for sending control, route request and response, as well as data packets originating at or routed through the transmitting node. The receiver is used to receive data and control packets – some of which are destined for the receiving node and some of which are forwarded. Understanding the power characteristics of the mobile radio used in wireless devices is important for the efficient design of communication protocols. A typical mobile radio may exist in three modes: transmit, receive, and standby. Maximum power is consumed in the transmit mode, and the least in the standby mode. For example, the Proxim RangeLAN2 2.4 GHz 1.6 Mbps PCMCIA card requires 1.5W in transmit, 0.75W in receive, and 0.01W in standby mode. In addition, turnaround between transmit and receive modes (and vice-versa) typically takes between 6 to 30 microseconds. Power consumption for Lucent's 15 dBm 2.4 GHz 2 Mbps Wavelan PCMCIA card is 1.82W in transmit mode, 1.80W in receive mode, and 0.18W in standby mode. Thus, the goal of protocol development for environments with limited power resources is to optimize the transceiver usage for a given communication task.

The computation considered in this paper is chiefly concerned with protocol processing aspects. It mainly involves usage of the CPU and main memory and, to a very small extent, the disk or other components. Also, data compression techniques, which reduce packet length (and hence energy usage), may result in increased power consumption due to increased computation.

There exists a potential tradeoff between computation and communication costs. Techniques that strive to achieve lower communication costs may result in higher computation needs, and vice-versa.

Hence, protocols that are developed with energy efficiency goals should attempt to strike a balance between the two costs.

### 3.2. General Conservation Guidelines and Mechanisms

The following discussion presents some general guidelines that may be adopted for an energy efficient protocol design, and Figure 3 lists areas in which conservation mechanisms are beneficial. Examples are provided in which these guidelines have been adopted. Some mechanisms are better suited for infrastructure networks and others for ad hoc networks.

Collisions should be eliminated as much as possible within the MAC layer since they result in retransmissions. Retransmissions lead to unnecessary power consumption and to possibly unbounded delays. Retransmissions cannot be completely avoided in a wireless network due to the high error-rates. Similarly, it may not be possible to fully eliminate collisions in a wireless mobile network. This is partly due to user mobility and a constantly varying set of mobiles in a cell. For example, new users registering with the base station may have to use some form of random access protocol. In this case, using a small packet size for registration and bandwidth request may reduce energy consumption. The EC-MAC protocol (Sivalingam et al., 2000) is one example that avoids collisions during reservation and data packet transmission.

In a typical broadcast environment, the receiver remains on at all times which results in significant power consumption. The mobile radio receives all packets, and forwards only the packets destined for the receiving mobile. This is the default mechanism used in the IEEE 802.11 wireless protocol in which the receiver is expected to keep track of channel status through constant monitoring. One solution is to broadcast a schedule that contains data transmission starting times for each mobile as in (Sivalingam et al., 2000). This enables the mobiles to switch to standby mode until the receive start time. Another solution is to turn off the transceiver whenever the node determines that it will not be receiving data for a period of time. The PAMAS protocol (Singh and Raghavendra, 1998) uses such a method.

Furthermore, significant time and power is spent by the mobile radio in switching from transmit to receive modes, and vice versa. A protocol that allocates permission on a slot-by-slot basis suffers substantial overhead. Therefore, this turnaround is a crucial factor in the performance of a protocol. If possible, the mobile should be allocated contiguous slots for transmission or reception to reduce turnaround, resulting in lower power consumption. This is similar to buffering writes to the hard disk in order to minimize seek latency and head movement. Also, it is beneficial for mobiles to request multiple transmission slots with a single reservation packet when requesting bandwidth in order to reduce the reservation overhead. This leads to improved bandwidth usage and energy efficiency. The scheduling algorithms studied in (Chen et al., 1999b) consider contiguous allocation and aggregate packet requests.

Assuming that mobiles transmit data transmission requests to the base station, a centralized scheduling mechanism that computes the system transmission schedule at the base station is more energy efficient. A distributed algorithm in which each mobile computes the schedule independently may not be desirable because mobiles may not receive all reservation requests due to radio and error constraints, and schedule computation consumes energy. Thus, computation of the transmission schedule ought to be relegated to the base station, which in turn broadcasts the schedule to each mobile. Most reservation and scheduling based protocols require the base station to compute the schedule.

The scheduling algorithm at the base station may consider the node's battery power level in addition to the connection priority. This allows traffic from low-power mobiles that may be dropped due to depletion of power reserves to be transmitted sooner. Such a mechanism has been studied in (Price, 2000, Price et al., 2001). Also, under low-power conditions, it may be useful to allow a mobile to re-arrange allocated slots among its own flows. This may allow certain high-priority traffic to be transmitted sooner rather than waiting for the originally scheduled time. Such mobile-based adaptive algorithms have been considered in (Damodaran and Sivalingam, 1999) and (Chen, 2000) in the context of energy efficiency and channel error compensation.

At the link layer, transmissions may be avoided when channel conditions are poor, as studied in (Zorzi and Rao, 1997a). Also, error control schemes that combine Automatic Repeat Request (ARQ) and Forward Error Correction (FEC) mechanisms may be used to conserve power (i.e. tradeoff retransmissions with ARQ versus longer packets with FEC) as in (Lettieri et al., 1997).

Energy efficient routing protocols may be achieved by establishing routes that ensure that all nodes equally deplete their battery power, as studied in (Woo et al., 1998, Chang and Tassiulas, 2000). This helps balance the amount of traffic carried by each node. A related mechanism is to avoid routing through nodes with lower battery power, but this requires a mechanism for dissemination of node battery power. Also, the periodicity of routing updates can be reduced to conserve energy, but may result in inefficient routes when user mobility is high. Another method to improve energy performance is to take advantage of the broadcast nature of the network for broadcast and multicast traffic as in (Singh et al., 1999, Wieselthier et al., 2000). In (Ramanathan and Rosales-Hain, 2000), the topology of the network is controlled by varying the transmit power of the nodes, and the topology is generated to satisfy certain network properties.

At the OS level, the common factor to all the different techniques proposed is suspension of a specific sub-unit such as disk, memory, display, etc. based upon detection of prolonged inactivity. Several methods of extending battery lifetime within the operating system and middleware layer are discussed in (Tiwari et al., 1994, Chandrakasan and Brodersen, 1995, Mehta et al., 1997). Other techniques studied include power-aware CPU scheduling (Weiser et al., 1994, Lorch and Smith, 1997) and page allocation (Lebeck et al., 2000). Within the application layer, the power conserving mechanisms tend to be application specific – such as database access (Imielinski et al., 1994, Alonso and Ganguly, 1993) and video processing (Chandrakasan and Brodersen, 1995, Gordon et al., 1996, Agrawal et al., 1998). A summary of software strategies for energy efficiency is presented in (Lorch and Smith, 1998).

## 4. MAC Sublayer

The MAC (Media Access Control) layer is a sublayer of the data link layer which is responsible for providing reliability to upper layers for the point-to-point connections established by the physical layer. The MAC sublayer interfaces with the physical layer and is represented by protocols that define how the shared wireless channels are to be allocated among a number of mobiles. This section presents the details of three specific MAC protocols: IEEE 802.11 (IEEE, 1998), EC-MAC (Sivalingam et al., 2000), and PAMAS (Singh and Raghavendra, 1998).

### 4.1. IEEE 802.11 Standard

The IEEE 802.11 (IEEE, 1998) protocol for wireless LANs is a multiple access technique based on CSMA/CA (Collision Sense Multiple Access/Collision Avoidance), and is derived from the MACA protocol described in (Karn, 1990). The basic protocol is defined as follows. A mobile with a packet to transmit senses the transmission channel for activity. The mobile captures the channel and transmits all pending data packets if the channel is not busy. Otherwise, the mobile defers transmission and enters the backoff state. The time period that follows is called the contention window and consists of a pre-determined number of transmission slots. The mobile randomly selects a slot in the contention window, and continuously senses the medium until its selected contention slot. The mobile enters the backoff state again if it detects transmission from some other mobile during that period. However, if no transmission is detected, the mobile transmits the access packet and captures the channel. Extensions to the basic protocol include provisions for MAC-level acknowledgements and request-to-send (RTS)/clear-to-send (CTS) mechanisms.

The IEEE 802.11 (IEEE, 1998) standard recommends the following technique for power conservation. A mobile that wishes to conserve power may switch to sleep mode and inform the base station of this decision. The base station buffers packets received from the network that are destined for the sleeping mobile. The base station periodically transmits a beacon that contains information about such buffered packets. When the mobile wakes up, it listens for this beacon, and responds to the base station which then forwards the packets. This approach conserves power but results in additional delays at the mobile that may affect the quality of service (QoS). A comparison of power-saving mechanisms in the IEEE 802.11 and HIPERLAN standards is presented in (Woesner et al., 1998). Presented in (Dhaou, 1999) is a load-sharing method for saving energy in an IEEE 802.11 network. Simulation results indicate total power savings of 5 – 15 %.

Experimental measurements of per-packet energy consumption for an IEEE 802.11 wireless network interface is reported in (Feeney, 1999b). This work uses the Lucent WaveLAN card for its experiments. The cost of sending and receiving a packet is measured for a network using UDP point-to-point (or unicast) and broadcast traffic with varying packet sizes. The energy cost is studied in terms of fixed cost per packet which reflects MAC operation and incremental cost that depends on packet size. The results show that both point-to-point and broadcast traffic transmission incur the same incremental costs, but point-to-point transmission incurs higher fixed costs because of the MAC coordination. The reception of point-to-point traffic maintains higher fixed costs since the receiver must respond with CTS and ACK messages. However, incremental costs of packet reception were identical for both traffic types. The study also measures power consumption for non-destination mobiles that are in range of the sender and receiver. These experiments are a valuable source of information and represent an important step in expanding the knowledge of energy efficient protocol development.

### 4.2. EC-MAC Protocol

Although the IEEE 802.11 standard addresses energy efficiency, it was not one of the central design issues in developing the protocol. The EC-

MAC (Energy Conserving-Medium Access Control) protocol (Sivalingam et al., 2000, Chen et al., 1999a), on the other hand, was developed with the issue of energy efficiency as a primary design goal. The EC-MAC protocol is defined for an infrastructure network with a single base station serving mobiles in its coverage area. This definition can be extended to an ad hoc network by allowing the mobiles to elect a coordinator to perform the functions of the base station. The general guidelines outlined in the previous section and the need to support QoS led to a protocol that is based on reservation and scheduling strategies. Transmission in EC-MAC is organized by the base station into frames as shown in Figure 4, and each slot equals the basic unit of wireless data transmission.

At the start of each frame, the base station transmits the frame synchronization message (FSM) which contains synchronization information and the uplink transmission order for the subsequent reservation phase. During the request/update phase, each registered mobile transmits new connection requests and status of established queues according to the transmission order received in the FSM. In this phase, collisions are avoided by having the BS send the explicit order of reservation transmission. New mobiles that have entered the cell coverage area register with the base station during the new-user phase. Here, collisions are not easily avoided and hence this may be operated using a variant of Aloha. This phase also provides time for the BS to compute the data phase transmission schedule. The base station broadcasts a schedule message that contains the slot permissions for the subsequent data phase. Downlink transmission from the base station to the mobile is scheduled considering the QoS requirements. Likewise, the uplink slots are allocated using a suitable scheduling algorithm.

Energy consumption is reduced in EC-MAC because of the use of a centralized scheduler. Therefore, collisions over the wireless channel are avoided and this reduces the number of retransmissions. Additionally, mobile receivers are not required to monitor the transmission channel as a result of communication schedules. The centralized scheduler may also optimize the transmission schedule so that individual mobiles transmit and receive within contiguous transmission slots. The priority round robin with dynamic reservation update and error compensation scheduling algorithm described in (Chen et al., 1999b) provides for contiguous slot allocation in order to reduce transceiver turnaround. Also, scheduling algorithms that consider mobile battery power level in addition to packet priority may improve performance for low-power mobiles. A family of algorithms based on this idea is presented in (Kishore et al., 1998, Price et al., 2001).

The frames may be designed to be fixed or variable length. Fixed length frames are desirable from the energy efficiency perspective, since a mobile that goes to sleep mode will know when to wake up to receive the FSM. However, variable length frames are better for meeting the demands of bursty traffic. The EC-MAC studies used fixed length frames.

The energy efficiency of EC-MAC is compared with that of IEEE 802.11 and other MAC protocols in (Chen et al., 1999a). This comparative study demonstrates how careful reservation and scheduling of transmissions avoids collisions that are expensive in energy consumption.

### 4.3. PAMAS Protocol

While the EC-MAC protocol described above was designed primarily for infrastructure networks, the PAMAS (Power Aware Multi-Access) protocol (Singh and Raghavendra, 1998) was designed for the ad hoc network, with energy efficiency as the primary design goal.

The PAMAS protocol modifies the MACA protocol described in (Karn, 1990) by providing separate channels for RTS/CTS control packets and data packets. In PAMAS, a mobile with a packet to transmit sends a RTS (request-to-send) message over the control channel, and awaits the CTS (clear-to-send) reply message from the receiving mobile. The mobile enters a backoff state if no CTS arrives. However, if a CTS is received, then the mobile transmits the packet over the data channel. The receiving mobile transmits a "busy tone" over the control channel enabling users tuned to the control channel to determine that the data channel is busy.

Power conservation is achieved by requiring mobiles that are not able to receive and send packets to turn off the wireless interface. The idea is that a data transmission between two mobiles need not be overheard by all the neighbors of the transmitter. The use of a separate control channel allows for mobiles to determine when and for how long to power off. A mobile should power itself off when: (i) it has no packets to transmit and a neighbor begins transmitting a packet not destined for it, and (ii) it does have packets to transmit but at least one neighbor-pair is communicating. Each mobile determines the length of time that it should be powered off through the use of a *probe* protocol, the details of which are available in (Singh and Raghavendra, 1998). Theoretical bounds on power savings for random, line, and fully connected topologies are also presented. The results from simulation and analysis show that between 10% to 70% power savings can be achieved for fully connected topologies.

## 5. LLC Sublayer

In this section, we focus on the error control functionality of the logical link control (LLC) sublayer. The two most common techniques used for error control are Automatic Repeat Request (ARQ) and Forward Error Correction (FEC). Both ARQ and FEC error control methods waste network bandwidth and consume power resources due to retransmission of data packets and greater overhead necessary in error correction. Care must be exercised while adopting these techniques over a wireless link where the error rates are high due to noise, fading signals, and disconnections caused from mobility. A balance needs to be maintained within this layer between competing measures for enhancing throughput, reliability, security, and energy efficiency. For example, channel encoding schemes for enhancing channel quality tend to reduce the throughput as more redundancy is added to the transmitted information. Also, increasing transmitted power to improve the channel to interference ratio depletes battery energy.

Recent research has addressed low-power error control and several energy efficient link layer protocols have been proposed. Three such protocols are described below.

### 5.1. Adaptive Error Control with ARQ

An ARQ strategy that includes an adaptive error control protocol is presented and studied in (Zorzi and Rao, 1997a, Zorzi and Rao, 1997b). First, though, the authors propose a new design metric for protocols developed specifically for the wireless environment and three guidelines in designing link layer protocols to be more power conserving. The new design metric introduced in (Zorzi and Rao, 1997a) is the energy efficiency of a protocol which is defined as the ratio between total amount of data delivered and total energy consumed. Therefore, as more data is successfully transmitted for a given amount of energy consumption, the energy efficiency of the protocol increases.

The following guidelines in developing a protocol should be considered in order to maximize the energy efficiency of the protocol.

1. Avoid persistence in retransmitting data.
2. Trade off number of retransmission attempts for probability of successful transmission.
3. Inhibit transmission when channel conditions are poor.

The energy efficient protocol proposed in (Zorzi and Rao, 1997a, Zorzi and Rao, 1997b) incorporates a probing protocol that slows down data transmission when degraded channel conditions are encountered. The ARQ protocol works as normal until the transmitter detects an error in either the data or control channel due to the lack of a received acknowledgement (ACK). At this time the protocol enters a *probing mode* in which a probing packet is transmitted every $t$ slots. The probe packet contains only a header with little or no payload and therefore consumes a smaller amount of energy. This mode is continued until a properly received ACK is encountered, indicating the recovered status of both channels. The protocol then returns to normal mode and continues data transmission from the point at which it was interrupted.

Using a Markov model based analysis and a recursive technique, the ARQ probing protocol is compared to traditional ARQ schemes, and the tradeoff between performance and energy efficiency is investigated. The results show that under slow fading channel conditions the proposed protocol is superior to that of standard ARQ in terms of energy efficiency, increasing the total number of packets that can be transmitted. The analysis also demonstrates that an optimal transmission power in respect to energy efficiency exists. Using a high transmission power to maximize the probability of a successful transmission may not be the best strategy. Although decreasing the transmission power results in an increased number of transmission attempts, it may be more efficient than attempting to maximize the throughput per slot. The conclusion reached is that although throughput is not necessarily maximized, the energy efficiency of a protocol may be maximized by decreasing the number of transmission attempts and/or transmission power in the wireless environment.

### 5.2. Adaptive Error Control with ARQ/FEC Combination

The above error control scheme included only ARQ strategies. However, the energy efficient error control scheme proposed in (Lettieri et al., 1997) combines ARQ and FEC strategies. The authors describe an error control architecture for the wireless link in which each packet stream maintains its own time-adaptive customized error control scheme based on certain set up parameters and a channel model estimated at runtime. The idea behind this protocol is that there exists no energy efficient "one-size-fits-all" error control scheme for all traffic types and channel conditions. Therefore, error control schemes should be customized to traffic requirements and channel conditions in order to obtain more optimal energy savings for each wireless connection.

The dynamic error control protocol described in (Lettieri et al., 1997) operates as follows. Service quality parameters, such as packet size and QoS requirements, used by the MAC sublayer and packet scheduler are associated with each data stream. These parameters are further used to select an appropriately customized combination of an ARQ scheme (Go-Back-N, Cumulative Acknowledgement (CACK), Selective Acknowledgement (SACK), etc.) and FEC scheme. In order to keep energy consumption at a minimum, the error control scheme associated with each stream may need to be modified as channel conditions change over time. Studies based on analysis and simulation under different scenarios were presented as a guideline in choosing an error control scheme to achieve low energy consumption while trading off QoS over various channel conditions, traffic types, and packet sizes. The authors extend their research in (Lettieri and Srivastava, 1998) to include a protocol for dynamically sizing the MAC layer frame, depending upon wireless channel conditions.

### 5.3. Adaptive Power Control and Coding Scheme

Finally, a dynamic power control and coding protocol for optimizing throughput, channel quality, and battery life is studied in (Agrawal et al., 1996, Narendran et al., 1997). This distributed algorithm, in which each mobile determines its own operating point with respect to power and error control parameters, maintains the goal of minimizing power utilization and maximizing capacity in terms of the number of simultaneous connections. Power control, as defined by the authors, is the technique of controlling the transmit power so as to affect receiver power, and ultimately the carrier-to-interference ratio (CIR).

The energy efficient power control and coding protocol operates in the following manner. Each transmitter operates at a power-code pair in which the power level lies between a specified minimum and maximum and the error code is chosen from a finite set. The algorithm is iterative in nature with the transmitter and receiver, at each iteration, cooperatively evaluating channel performance and determining if an adjustment in the power-code pair is necessary. The time between each iteration is referred to as a *timeframe*. After each timeframe the receiver involved in the data transmission evaluates the channel performance by checking the word error rate (WER). If the WER lies within an acceptable range, the power-code pair is retained; otherwise a new power-code pair is computed by the transmitter. The basic frame of the algorithm can be modified such that optimal levels of control overhead and channel quality are traded off. Also, variations of the base algorithm include the evaluation of the average WER, rather than the instantaneous WER in each timeframe, in determining channel quality and in the evaluation of anticipated channel performance. The latter adaptation of the algorithm attempts to predict changes in error rates due to mobility by sampling the received powers and extrapolating the values to the next timeframe. If the predicted WER is not within acceptable ranges, then the power-code pair is adapted to avoid unsatisfactory channel conditions.

A study of the dynamic power control and coding protocol was performed through simulation of a cellular system with roaming mobiles. Simulation results indicate that the proposed dynamic power control and coding protocol supports better quality channels as compared to schemes that use fixed codes; therefore power-control alone does not perform as well as an adaptive power-control/FEC protocol.

The next section discusses energy efficient routing protocols within the network layer.

## 6. Network Layer

The main functions of the network layer are routing packets and congestion control. In wireless mobile networks, the network layer has the added functionality of routing under mobility constraints and mobility management including user location, update, etc. In this section, we present energy efficient routing algorithms developed for wireless ad hoc networks. Energy efficient routing does not apply to infrastructure networks because all traffic is routed through the base station.

As mentioned earlier, in ad hoc networks the mobiles cooperate to maintain topology information and use multi-hop packet routing. The problem of routing is complicated due to user mobility resulting in frequently changed network topologies. The rate of topology change depends on many factors including user mobility speeds and terrain characteristics. Typical routing algorithms for ad hoc networks consider two different approaches:

1. Use frequent topology updates resulting in improved routing, but increased update messages consume precious bandwidth.

2. Use infrequent topology updates resulting in decreased update messages, but inefficient routing and occasionally missed packets results.

Typical metrics used to evaluate ad hoc routing protocols are shortest-hop, shortest-delay, and locality stability (Woo et al., 1998). However, these metrics may have a negative effect in wireless networks because they result in the overuse of energy resources of a small set of mobiles, decreasing mobile and network life. For example, consider the wireless network in Figure 5. Using shortest-hop routing, traffic from mobile A to mobile D will always be routed through mobile E, which will drain the energy reserves of E faster. If mobile E's battery becomes fully drained, then mobile F is disconnected from the network and communication to and from F is no longer possible. By using a routing algorithm that takes into account such issues, traffic from A to D may not always be routed through E, but through mobiles B and C, extending network life. Consequently, it is essential to consider routing algorithms from an energy efficient perspective, in addition to traditional metrics. Such research is described in the following paragraphs.

### 6.1. Unicast Traffic

Unicast traffic is defined as traffic in which packets are destined for a single receiver. In (Woo et al., 1998), routing of unicast traffic is addressed with respect to battery power consumption. The authors' research focuses on designing protocols to reduce energy consumption and to increase the life of each mobile, increasing network life as well. To achieve this, five different metrics are defined from which to study the performance of power-aware routing protocols.

*Energy consumed per packet:* It is easy to observe that if energy consumed per packet is minimized then the total energy consumed is also minimized. Under light loads, this metric will most likely result in the shortest-hop path. As network load increases, this is not necessarily the case because the metric will tend to route packets around areas of congestion in the network.

*Time to network partition:* Given a network topology, a minimal set of mobiles exist such that their removal will cause the network to partition. Routes between the two partitions must go through one of the "critical" mobiles; therefore a routing algorithm should divide the work among these mobiles in such a way that the mobiles drain their power at equal rates.

*Variance in power levels across mobiles:* The idea behind this metric is that all mobiles in a network operate at the same priority level. In this way, all mobiles are equal and no one mobile is penalized or privileged over any other. This metric ensures that all mobiles in the network remain powered-on *together* for as long as possible.

*Cost per packet:* In order to maximize the life of all mobiles in the network, metrics other than energy consumed per packet need to be used. When using these metrics, routes should be created such that mobiles with depleted energy reserves do not lie on many routes. Together, these metrics become the "cost" of a packet, which needs to be minimized.

*Maximum mobile cost:* This metric attempts to minimize the cost experienced by a mobile when routing a packet through it. By minimizing the cost per mobile, significant reductions in the maximum mobile cost result. Also, mobile failure is delayed and variance in mobile power levels is reduced due to this metric.

In order to conserve energy, the goal is to minimize all the metrics except for the second which should be maximized. As a result, a shortest-hop routing protocol may no longer be applicable; rather, a shortest-cost routing protocol with respect to the five energy efficiency metrics would be pertinent. For example, a cost function may be adapted to accurately reflect a battery's remaining lifetime. The premise behind this approach is that although packets may be routed through longer paths, the paths contain mobiles that have greater amounts of energy reserves. Also, energy can be conserved by routing traffic through lightly loaded mobiles because the energy expended in contention and retransmission is minimized.

The properties of the power-aware metrics and the effect of the metrics on end-to-end delay are studied in (Woo et al., 1998) using simulation. A comparison of shortest-hop routing and the power-aware shortest-cost routing schemes was conducted. The performance measures were delay, average cost per packet, and average maximum node cost. Results show that usage of power-aware metrics result in *no extra delay* over the traditional shortest-hop metric. This is true because congested paths are often avoided. However, there was significant improvement in average cost per packet and average maximum mobile cost in which the cost is in terms of the energy efficient metrics defined above. The improvements were substantial for large networks and heavily-loaded networks. Therefore, by adjusting routing parameters a more energy efficient routing scheme may be utilized for wireless networks.

The above approach to routing in wireless ad hoc networks requires, at the least, that every mobile have knowledge of the locations of every other mobile and the links between them. This creates significant communication overhead and increased delay. Research completed in (Stojmenovic and Lin, 2000) addresses this issue by proposing *localized* routing algorithms which depend only on information about the source location, the location of neighbors, and location of the destination. This information is collected through GPS receivers which are included within each mobile. Therefore, excessive network communication is not required which, the authors report, more than makes up for the extra energy consumed by the GPS units.

A new power-cost metric incorporating both a mobile's lifetime and distance based power metrics is proposed, and using the newly defined metric, three power-aware localized routing algorithms are developed: *power*, *cost*, and *power-cost*. The power algorithm attempts to minimize the total amount of power utilized when transmitting a packet, whereas the cost algorithm avoids mobiles that maintain low battery reserves in order to extend the network lifetime. Finally, the power-cost routing algorithm is a combination of the two algorithms. Experiments validated the performance of these routing algorithms.

### 6.2. Broadcast Traffic

In this section broadcast traffic, which is defined as traffic in which packets are destined for all mobiles in the system, is considered. With a single transmission, a mobile is able to broadcast a packet to all immediate neighbors. However, each mobile needs to receive a packet only once. Intermediate mobiles are required to retransmit the packet. The key idea in conserving energy is to allow each mobile's radio to turn off after receiving a packet if its neighbors have already received a copy of the packet. Addressed in (Singh et al., 1999) is the routing of broadcast traffic in terms of power consumption.

The broadcast technique used in traditional networks is a simple flooding algorithm. This algorithm gathers no global topology information, requiring little control overhead and completes the broadcast with minimum number of hops. However, the flooding algorithm is not suitable for wireless networks because many intermediate nodes must retransmit packets needlessly which leads to excessive power consumption. Therefore, the authors of (Singh et al., 1999) propose that it is more beneficial to spend some energy in gathering topology information in order to determine the most efficient broadcast tree.

In order to increase mobile and network life, any broadcast algorithm used in the wireless environment should focus on conserving energy and sharing the cost of routing among all mobiles in the system. One way to conserve power is by ensuring that a transmission reaches as many *new* nodes as possible. A broadcast tree approach is presented in (Singh et al., 1999), in which the tree is constructed starting from a source and expanding to the neighbor that has the lowest cost per outgoing degree, where the cost associated with each mobile increases as the mobile consumes more power. Therefore, priority for routing packets through the broadcast tree is given to nodes that have consumed lower amounts of power and nodes that have more neighbors which have not already received the data transmission. Since mobile costs continuously change, broadcast transmissions originating from the same source may traverse different trees, as they are determined based on current costs of nodes.

Simulations were conducted in order to study the performance of the proposed power-aware broadcast protocol as compared to traditional flooding in terms of energy savings as well as delay. Averaged over a period of time, the power-aware protocol demonstrates very little difference in broadcast delay. However, results indicate that savings in energy consumption of 20% or better are possible using the power-aware broadcast algorithm, with greater savings in larger networks and networks with increased traffic loads.

The construction of energy efficient broadcast and multicast trees for the wireless environment is also studied in (Wieselthier et al., 2000). The authors state that mobiles may experience greater energy conservation if routing decisions are combined with decisions concerning transmission power levels. An algorithm is presented for determining the minimum-energy source-based tree for each broadcast/multicast session request. The algorithm is based on the concept that there exists an optimal point in the trade-off between reaching greater number of mobiles in a single hop by using higher transmission power versus reaching fewer mobiles but using lower power levels. Performance results demonstrate that the combination of routing and transmission power decisions provide greater energy conservation. A similar idea concerning the incorporation of transmission power levels into routing algorithms is also presented for unicast traffic in (Chang and Tassiulas, 2000).

In (Feeney, 1999a), a simulation based comparison of energy consumption for two ad hoc routing protocols – Dynamic Source Routing (DSR) (Johnson et al., 2000) and Ad hoc On Demand Vector routing (AODV) (Perkins et al., 2000) protocols is presented. The analysis considers the cost for sending and receiving traffic, for dropped packets, and for routing overhead packets. User mobility is modeled in the analysis. The observations indicate that energy spent on receiving and discarding packets can be significant. Also, the costs of flooding-based broadcast traffic and MAC control were seen to be significant. For DSR, results show that the cost of source routing headers was not very high, but operating the receiver in promiscuous mode for caching and route response purposes resulted in high power consumption. Results also indicate that since AODV generates broadcast traffic more often, the energy cost is high given that broadcast traffic consumes more energy. Refer to (Feeney, 1999a) for more detailed results.

The next section presents work related to improving transport protocol performance in the wireless environment.

## 7. Transport Layer

The transport layer provides a reliable end-to-end data delivery service to applications running at the end points of a network. The most commonly used transport protocol for wired networks, where underlying physical links are fairly reliable and packet loss is random in nature, is the Transmission Control Protocol (TCP) (Postel, 1981). However, due to inherent wireless link properties, the performance of traditional transport protocols such as TCP degrades significantly over a wireless link. TCP and similar transport protocols resort to a larger number of retransmissions and frequently invoke congestion control measures, confusing wireless link errors and loss due to handoff as channel congestion. This can significantly reduce throughput and introduce unacceptable delays (Cáceres and Iftode, 1995). As stated earlier, increased retransmissions unnecessarily consume battery energy and limited bandwidth.

Recently, various schemes have been proposed to alleviate the effects of non congestion-related losses on TCP performance over networks with wireless links. These schemes, which attempt to reduce retransmissions, are classified into three basic groups: (i) split connection protocols, (ii) link-layer protocols, and (iii) end-to-end protocols.

Split connection protocols completely hide the wireless link from the wired network by terminating the TCP connections at the base station as shown in Figure 6. This is accomplished by splitting each TCP connection between the source and destination into two separate connections at the base station. The result is one TCP connection between the wired network and the base station and a second TCP connection between the base station and the mobile. The second connection over the wireless link may use modified versions of TCP that enhance performance over the wireless channel. Examples of split connection protocols include Indirect-TCP (Bakre and Badrinath, 1995), Berkeley Snoop Module (Balakrishnan et al., 1995), and M-TCP (Brown and Singh, 1997).

Figure 7 depicts the link layer approach which attempts to hide link related losses from the TCP source by using a combination of local retransmissions and forward error correction over the wireless link. Local retransmissions use techniques that are tuned to the characteristics of the wireless channel to provide significant increase in performance. One example of a link layer protocol is the AIRMAIL protocol (Ayanoglu et al., 1995), which employs a combination of both FEC and ARQ techniques for loss recovery.

Finally, end-to-end protocols include modified versions of TCP that are more sensitive to the wireless environment. End-to-end protocols require that a TCP source handle losses through the use of such mechanisms as selective acknowledgements and explicit loss notification (ELN). Selective acknowledgements allow the TCP source to recover from multiple packet losses, and ELN mechanisms aid the TCP source in distinguishing between congestion and other forms of loss.

### 7.1. Energy Consumption Analysis of TCP

The procotols described previously generally achieve higher throughput rates over the wireless channel than standard TCP because the protocols are better able to adapt to the dynamic mobile environment. However, the performance of a particular protocol is largely dependent upon various factors such as mobility handling, amount of overhead costs incurred, frequency and handling of disconnections, etc. Therefore, performance and energy conservation may range widely for these protocols depending upon both internal algorithm and external environmental factors. Although the above protocols, along with many others proposed in research, have addressed the unique needs of designing transport protocols in the wireless environment which may or may not lead to greater energy efficiency, they have not directly addressed the idea of a low-power transport protocol.

The energy consumption of Tahoe, Reno, and New Reno versions of TCP is analyzed in (Zorzi and Rao, 2000). Energy consumption is the main parameter studied with the objective of measuring the effect of TCP transmission policies on energy performance. The energy efficiency of a protocol is defined as the average number of successful transmissions per energy unit, which can be computed as the average number of successes per transmission attempt. A two-state Markov packet error process is used in the performance evaluation of a single transceiver running the various versions of TCP on a dedicated wireless link. Results of the study demonstrate that error correlation significantly affects the energy performance of TCP and that congestion control algorithms of TCP actually allow for greater energy savings by backing off and waiting during error bursts. It is also seen that energy efficiency may be quite sensitive to the version of TCP implemented and the choice of protocol parameters.

The same versions of TCP were studied in (Tsaoussidis et al., 2000a) in terms of energy/throughput tradeoffs. Simulation results show that no single TCP version is most appropriate within wired/wireless heterogenous networks, and that the key to balancing energy and throughput performance is through the error control mechanism. Using these results, the authors propose a modified version of TCP, referred to as TCP-Probing, in (Tsaoussidis and Badr, 2000).

In TCP-Probing, data transmission is suspended and a probe cycle is initiated when a data segment is delayed or lost, rather than immediately invoking congestion control. A probe cycle consists of an exchange of probe segments between sender and receiver. Probe segments are implemented as extensions to the TCP header and carry no payload. The TCP sender monitors the network through the probe cycle which terminates when two consecutive round-trip-times (RTT) are successfully measured. The sender invokes standard TCP congestion control if persistent error conditions are detected. However, if monitored conditions indicate transient random error, then the sender resumes transmission according to available network bandwidth. Simulation results provided in (Tsaoussidis and Badr, 2000) indicate that TCP-Probing achieves higher throughput rates while consuming less energy. Therefore, the authors believe that TCP-Probing provides a universal error control mechanism for heterogenous wired/wireless networks. The authors also present in (Tsaoussidis et al., 2000b) an experimental transport protocol, called Wave and Wait Protocol (WWP), developed specifically for a wireless environment with limited power.

## 8. OS/Middleware and Application Layers

This section addresses research completed at the OS/middleware and application layers with respect to energy efficiency.

### 8.1. OS/Middleware

One important advantage of integrating wireless communication with computing is that it facilitates user mobility and connectivity to the network. Mobility, directly or indirectly, impacts the design of operating systems, middleware, file systems, and databases. It also presents a new set of challenges that result from power constraints and voluntary disconnections. To be consistent with fixed counterparts like PCs and workstations, mobile computers need to process multimedia information. However, such processing is expensive in terms of both bandwidth and battery power. In general, the majority of the techniques used in the design of today's applications to conserve bandwidth also conserve battery life.

The main function of an operating system is to manage access to physical resources like CPU, memory, and disk space from the applications running on the host. To reduce power dissipation, CPUs used in the design of portable devices can be operated at lower speeds by scaling down the supply voltage (Chandrakasan and Brodersen, 1995).

Due to the quadratic relationship between power and supply voltage, halving the supply voltage results in one fourth of the power being consumed. To maintain the same throughput, the reduction in circuit speed can be compensated by architectural techniques like pipelining and parallelism. These techniques increase throughput resulting in an energy efficient system operating at a lower voltage but with the same throughput. The operating system is active in relating scheduling and delay to speed changes.

Another technique of power management at this layer is predictive shutdown (Chandrakasan and Brodersen, 1995). This method exploits the *event driven* nature of computing in that sporadic computation activity is triggered by external events and separated by periods of inactivity. A straightforward means of reducing average energy consumption is to shut down the system during periods of inactivity. However, preserving the latency and throughput of applications requires intelligent activity-based predictive shutdown strategies.

In (Lebeck et al., 2000), a study of different page placement algorithms that exploit the new power management features of memory technology is presented. The study considers DRAM chips that support different power modes: active, standby, nap and powerdown. Trace-driven and execution-driven simulations show that improvement of 6% to 55% in the Energy $\times$ Delay metric are obtained using power-aware page allocation mechanisms that operate in conjunction with hardware policies.

CPU scheduling techniques that attempt to minimize power consumption are presented in (Weiser et al., 1994, Lorch and Smith, 1997). The impact of software architecture on power consumption is studied in (Tiwari et al., 1994, Mehta et al., 1997).

### 8.2. Application Layer

The application layer in a wireless system is responsible for such things as partitioning of tasks between the fixed and mobile hosts, audio and video source encoding/decoding, and context adaptation in a mobile environment. Energy efficiency at the application layer is becoming an important area of research as is indicated by industry. APIs such as Advanced Configuration and Power Interface (Intel Corporation, Microsoft and Toshiba Corporation, 2000) and power management analysis tools such as Power Monitor (Intel Corporation, 2000) are being developed to assist software developers in creating programs that are more power conserving. Another power management tool developed at Carnegie Mellon University is PowerScope (Flinn and Satyanarayanan, 1999). PowerScope maps energy consumption to program structure, producing a profile of energy usage by process and procedure. The authors report a 46% reduction in energy consumption of an adaptive video playing application by taking advantage of the information provided by PowerScope. This section summarizes some of the research being conducted at the application layer with respect to power conservation.

*Load Partitioning:* Challenged by power and bandwidth constraints, applications may be selectively partitioned between the mobile and base station (Weiser et al., 1994, Narayanaswamy et al., 1996). Thus, most of the power intensive computations of an application are executed at the base station, and the mobile host plays the role of an intelligent terminal for displaying and acquiring multimedia data (Narayanaswamy et al., 1996).

*Proxies:* Another means of managing energy and bandwidth for applications on mobile clients is to use proxies. Proxies are middleware that automatically adapt the applications to changes in battery power and bandwidth. A simple example of proxy usage during multimedia transmissions in a low-power or low bandwidth environment is to suppress video and permit only audio streams. Another example is to direct a file to be printed at the nearest printer when the host is mobile. Proxies are either on the mobile or base station side of the wireless link.

*Databases:* Impact of power efficiency on database systems is considered by some researchers. For example, energy efficiency in database design by minimizing power consumed per transaction through embedded indexing has been addressed in (Imielinski et al., 1994). By embedding the directory in the form of an index, the mobile only needs to become active when data of interest is being broadcast (the system architecture consists of a single broadcast channel). When a mobile needs a piece of information an initial probe is made into the broadcast channel. The mobile is then able to determine the next occurrence of the required index and enters *probe wait* mode while it waits for the index to be broadcast. After receiving the index information relevant to the required data, the mobile enters *bcast wait* mode while it waits for the information to be broadcast. *Access time* is defined as the sum of the two waiting periods, *probe wait* and *bcast wait.* The goal of the authors is to provide methods to combine index information together with data on the single broadcast channel in order to minimize access time. The authors propose two such strategies which are further described in (Imielinski et al., 1994). Also, energy efficient query optimization for database systems is described in (Alonso and Ganguly, 1993).

*Video Processing:* Multimedia processing and transmission require considerable battery power as well as network bandwidth. This is especially true for video processing and transmission. However, reducing the effective bit rate of video transmissions allows lightweight video encoding and decoding techniques to be utilized thereby reducing power consumption. Under severe bandwidth constraints or low-power situations, video frames can even be carefully discarded before transmission while maintaining tolerable video quality.

In (Agrawal et al., 1998), research on processing encoded video for transmission under low battery power conditions is presented. The basic idea of this work is to decrease the number of bits transmitted over the wireless link in response to low-power situations. The challenge is to accomplish this goal while preserving or minimally degrading the video quality. Decreasing the number of transmitted bits reduces the energy consumption due to reduced transmitter usage. In fact, several studies have shown that transmission accounts for more than a third of the energy consumption in video processing and exchange in a portable device. The reduction in the number of bits can be achieved in one of two ways: (i) reducing the number of bits in the compressed video stream generated by the video encoder, and (ii) discarding selected packets at the wireless network interface card (WNIC).

The first approach is possible only if two conditions are satisfied. The portable device must be encoding a video stream as opposed to transmitting a stored video, and the application must be able to modify parameters in the video encoder. The second approach is possible if the WNIC is flexible and sensitive to battery power conditions. Further details on how the different encoding schemes affect the choice of discarding may be found in (Agrawal et al., 1998). Also, a testbed implementing this research was developed, and preliminary results reported in (Mahadevan, 1999).

Power-aware video processing is an important and exciting topic. There are several approaches for developing efficient encoding schemes that will impact performance and energy consumption as in (Liu and Zarki, 1998, Swann and Kingsbury, 1998). However, a complete discussion is not presented due to space constraints.

## 9. Summary

As wireless services continue to add more capabilities such as multimedia and QoS, low-power design remains one of the most important research areas within wireless communication. Research must focus on decreasing the amount of energy consumed by the wireless terminal. Power conservation has typically been considered at the physical layer. However, most of the energy savings at the physical layer have already been achieved. Therefore, the key to energy conservation in wireless communications lies within the higher levels of the wireless protocol stack. This paper describes research completed at the data link, network, transport, OS/middleware, and application protocol layers that have addressed energy efficiency for wireless networks. However, power conservation within the wireless protocol stack remains a very crucial research area for the viability of wireless services in the future.

## Acknowledgements

The authors wish to thank the reviewers and the Editor for their valuable suggestions and comments that helped improve the paper; and Stephanie Lindsey and Harini Krishnamurthy who assisted with editing the paper.

Correspondence Information

Prof. Krishna Sivalingam
Boeing Associate Professor of Computer Science
School of Elect. Engg. & Computer Science
102 EME Building
Washington State University
Pullman, WA 99164-2752
Phone: 509 335 3220
Fax: 253-295-9458 (Please email to addr below after sending Fax)
Email: krishna@eecs.wsu.edu

*Author Biographies:*

**Christine Jones** received her M.S. in Computer Science from Washington State University, Pullman in 2000 and her B.S. in Computer Science from Whitworth College in 1998. She is presently with BBN Technologies in Cambridge, MA, USA.

Photo for Jones is included as a JPEG file.

**Krishna M. Sivalingam** (ACM '93, IEEE SM '00 M '95) received his Ph.D. and M.S. degrees in Computer Science from State University of New York at Buffalo in 1994 and 1990 respectively. While at SUNY Buffalo, he was a Presidential Fellow from 1988 to 1991. Prior to that, he received the B.E. degree in Computer Science and Engineering in 1988 from Anna University, Madras, India. He is Boeing Associate Professor of Computer Science in the School of Electrical Engineering and Computer Science, at Washington State University, Pullman, where he was an Assistant Professor from 1997 to 2000. Earlier, he was an Assistant Professor at University of North Carolina Greensboro from 1994 until 1997. He has conducted research at Lucent Technologies' Bell Labs in Murray Hill, NJ, and at AT&T Labs in Whippany, NJ.

His research interests include wireless networks, optical wavelength division multiplexed networks, and performance evaluation. He has served as a Guest Co-Editor for a special issue of the IEEE Journal on Selected Areas in Communications on optical WDM networks. He is co-recipient of the Best Paper Award at the IEEE International Conference on Networks 2000 held in Singapore. He has published an edited book on optical WDM networks in 2000. His work is supported by AFOSR, Laboratory for Telecommunication Sciences, NSF, Cisco, Bellcore, Alcatel, Intel, and Washington Technology Center. He holds three patents in wireless networks and has published several papers including 18 journal publications. He has served on several conference committees including ACM Mobicom 2001, Opticom 2000, ACM Mobicom 99, MASCOTS 99, and IEEE INFOCOM 1997. He is a Senior Member of IEEE and a member of ACM. Email: krishna@eecs.wsu.edu

Photo of Krishna Sivalingam is enclosed in hard-copy.

**Prathima Agrawal** is Assistant Vice President of the Internet Architecture Research Laboratory and Executive Director of the Computer Networking Research Department at Telcordia Technologies (formerly Bellcore), Morristown, NJ. She is also an adjunct Professor of Electrical and Computer Engineering, Rutgers University, NJ. She worked for 20 years in AT&T/Lucent Bell Laboratories in Murray Hill, NJ, where she was Head of the Networked Computing Research Department. Presently, she leads the ITSUMO joint research project between Telcordia and Toshiba Corp. ITSUMO is a third generation wireless access system for multimedia communication over end-to-end packet networks. The ITSUMO project received the Telcordia CEO Award for 2000. Dr. Agrawal received her Ph.D. degree in Electrical Engineering from the University of Southern California. Her research interests are computer networks, mobile and wireless computing and communication systems and parallel processing. She has published over 150 papers and has received or applied for more than 50 U.S. patents. Dr. Agrawal is a Fellow of the IEEE and a member of the ACM. She is the recipient of an IEEE Third Millennium Medal in 2000. She chaired the IEEE Fellow Selection Committee during 1998-2000.

Photo of Agrawal is enclosed as a JPEG file.

**Jyh-Cheng Chen** received the B.S. degree in information science from Tunghai University, Taichung, Taiwan, in 1990, the M.S. degree in computer engineering from Syracuse University, Syracuse, NY, in 1992, and the Ph.D. degree in electrical engineering from the State University of New York, Buffalo, in 1998.

Since August 1998, he has been a Research Scientist in Applied Research at Telcordia Technologies, Morristown, NJ. At Telcordia, he is one of the key architects and implementers of the ITSUMO (Internet Technologies Supporting Universal Mobile Operation) project, in which he has been working on QoS for mobile and wireless IP networks, IP-based base station design, SIP-based mobility management, and multimedia applications, etc. He received the 2000 Telcordia CEO Award that was intended to recognize and honor the Company's most exceptional teams and individuals. While working toward his Ph.D., he also worked on energy efficient MAC protocols for wireless ATM networks at AT&T Labs, Whippany, NJ, and non-destructive coating thickness measurement at ASOMA-TCI Inc., N. Tonawanda, NY.

Dr. Chen is a member of IEEE and ACM.

Email: jcchen@research.telcordia.com

Photo of Chen is enclosed as a JPEG file.

## References


Agrawal, P., J.-C. Chen, S. Kishore, P. Ramanathan, and K. M. Sivalingam: 1998, 'Battery power sensitive video processing in wireless networking'. In: *Proc. IEEE PIMRC '98*.

Agrawal, P., B. Narendran, J. Sienicki, and S. Yajnik: 1996, 'An adaptive power control and coding scheme for mobile radio systems'. In: *Proc. of IEEE Personal Communications Conf.* New Delhi, India, pp. 283–288.

Alonso, R. and S. Ganguly: 1993, 'Energy efficient query optimization'. Technical Report MITL-TR-33-92, Matsushita Info Tech Lab, Princeton, NJ.

Ayanoglu, E., S. Paul, T. F. LaPorta, K. K. Sabnani, and R. D. Gitlin: 1995, 'AIRMAIL: A link-layer protocol for wireless networks'. *ACM/Baltzer Wireless Networks* **1**(2), 161–174.

Bakre, A. and B. R. Badrinath: 1995, 'I-TCP: indirect TCP for mobile hosts'. In: *Proc. of 15th Int. Conf. on Distributed Computing Systems.* pp. 136–143.

Balakrishnan, H., S. Seshan, and R. Katz: 1995, 'Improving reliable transport and handoff performance in cellular wireless networks'. *Wireless Networks* **1**(4).

Bambos, N.: 1998, 'Toward power-sensitive network architectures in wireless communications: concepts, issues and design aspects'. *IEEE Personal Communications* **5**(3), 50–59.

Brown, K. and S. Singh: 1997, 'M-TCP: TCP for mobile cellular networks'. *Computer Communication Review* **27**(5), 19–43.

Cáceres, R. and L. Iftode: 1995, 'Improving the performance of reliable transport protocols in mobile computing environments'. *IEEE Journal on Selected Areas in Communications* **13**(5), 850–857.

Chandrakasan, A. and R. W. Brodersen: 1995, *Low Power Digital CMOS Design.* Norwell, MA: Kluwer Academic Publishers.

Chang, J.-H. and L. Tassiulas: 2000, 'Energy conserving routing in wireless ad-hoc networks'. In: *Proc. IEEE INFOCOM.* Tel-Aviv, Israel.

Chen, J.-C., K. M. Sivalingam, and P. Agrawal: 1999a, 'Performance comparison of battery power consumption in wireless multiple access protocols'. *ACM/Baltzer Wireless Networks* **5**(6), 445–460.

Chen, J.-C., K. M. Sivalingam, P. Agrawal, and R. Acharya: 1999b, 'Scheduling Multimedia Services for a Low-Power MAC in Wireless and Mobile ATM Networks'. *IEEE Transactions on Multimedia* **1**(2), 187–201.

Chen, L.: 2000, 'Mobile-based error compensation algorithms for wireless networks'. Master's thesis, Washington State University, Pullman.

Damodaran, S. and K. M. Sivalingam: 1999, 'Adaptive Scheduling at Mobiles for Wireless Networks with Multiple Priority Traffic and Multiple Transmission Channels'. In: *Proc. of Intl. Conf. on High Performance Computing.* Calcutta, India.

Dhaou, I. B.: 1999, 'A novel load-sharing algorithm for energy efficient MAC protocol compliant with 802.11 WLAN'. In: *Proc. IEEE Vehicular Technology Conference.* Amsterdam, Holland.

Douglis, F., P. Krishnan, and B. Marsh: 1994, 'Thwarting the Power-Hungry Disk'. In: *Proceedings of the 1994 Winter USENIX Conference.*

Feeney, L. M.: 1999a, 'An Energy Consumption Model for Performance Analysis of Routing Protocols for Mobile Ad Hoc Networks'. Proc. of the 45th IETF Meeting: MANET Working Group.

Feeney, L. M.: 1999b, 'Investigating the Energy Consumption of an IEEE 802.11 Network Interface'. Technical Report T1999-11, Swedish Institute of Computer Science, Kista, Sweden.

Flinn, J. and M. Satyanarayanan: 1999, 'PowerScope: A tool for profiling the energy usage of mobile applications'. In: *Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications.* http://www.cs.cmu.edu/user/jflinn/www/pscope.html.

Gordon, B. M., E. Tsern, and T. H. Meng: 1996, 'Design of a low power video decompression chip set for portable applications'. *Journal of VLSI Signal Processing Systems* **13**, 125–142.

Govil, K., E. Chan, and H. Wasserman: 1995, 'Comparing algorithms for dynamic speed-setting of a low-power CPU'. In: *Proc. Mobicom.* pp. 13–25.

IEEE: 1998, 'Wireless LAN medium access control (MAC) and physical layer (PHY) Spec'. IEEE 802.11 standard.

Imielinski, T., S. Vishwanathan, and B. R. Badrinath: 1994, 'Energy efficient indexing on air'. In: *Proceedings of the International conference on Management of Data (ACM-SIGMOD).*

Infrared Data Association: 2000, 'IrDA SIR data specification'. http://www.irda.org/standards/specifications.asp.

Intel Corporation: 2000, 'Intel Power Measurement Tools'. http://developer.intel.com/design/mobile/intelpower/tools/.

Intel Corporation, Microsoft and Toshiba Corporation: 2000, 'Advanced Configuration & Power Interface'. http://www.teleport.com/~acpi/.

Johnson, D. B., D. A. Maltz, Y.-C. Hu, and J. G. Jetcheva: 2000, 'The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks'. IETF Draft MANET Working Group.

Karn, P.: 1990, 'MACA - A new channel access method for packet radio'. In: *ARRL/CRRL Amateur Radio* $9^{th}$ *Computer Networking Conference.* pp. 134–140.

Kishore, S., J.-C. Chen, K. M. Sivalingam, and P. Agrawal: 1998, 'Adaptive Power Control and Scheduling Algorithms Based on Battery Power Level for CDMA Wireless Networks'. In: *Proc. of IEEE ICUPC.* Florence, Italy, pp. 967–971.

Lebeck, A. R., X. Fan, H. Zeng, and C. Ellis: 2000, 'Power-aware Page Allocation'. In: *Proc. of Ninth ACM ASPLOS.* pp. 105–116.

Lettieri, P., C. Fragouli, and M. B. Srivastava: 1997, 'Low power error control for wireless links'. In: *Proc. ACM MobiCom.* Budapest, Hungary.

Lettieri, P. and M. Srivastava: 1999, 'Advances in wireless terminals'. *IEEE Personal Communications* **6**(1), 6–18.

Lettieri, P. and M. B. Srivastava: 1998, 'Adaptive frame length control for improving wireless link throughput, range, and energy efficiency'. In: *INFOCOM*, Vol. 2. pp. 564–571.

Liu, H. and M. E. Zarki: 1998, 'Adaptive source rate control for real-time wireless video transmission'. *ACM/Baltzer Mobile Networks and Applications* **3**, 49–60.

Lorch, J. and A. J. Smith: 1997, 'Scheduling Techniques for Reducing Processor Energy Use in MacOS'. *ACM/Baltzer Wireless Networks* **3**(5), 311–324.

Lorch, J. R. and A. J. Smith: 1998, 'Software strategies for portable computer energy management'. *IEEE Personal Communications* **5**(3), 60–73.

Macker, J. and M. Corson, 'IETF working group: mobile ad-hoc networks (MANET)'. http://www.ietf.org/html.charters/manet-charter.html.

Macker, J. and M. Corson: 1998, 'Mobile ad-hoc networking and the IETF'. *ACM Mobile Computing and Communications Review* **2**(1).

Mahadevan, I.: 1999, 'Architecture and algorithms for Quality of Service support and energy efficient protocols for wireless/mobile networks'. Ph.D. thesis, Washington State University.

Marsh, B., F. Douglis, and P. Krishnan: 1994, 'Flash Memory File Caching for Mobile Computers'. In: *Proceedings of the 27th Hawaii Conference on Systems Science*.

Mehta, H., R. M. Owens, M. J. Irwin, R. Chen, and D. Ghosh: 1997, 'Techniques for Low Energy Software'. In: *Proc. of the International Symposium on Low Power Electronics and Design*. pp. 72–75.

Narayanaswamy, S., S. Seshan, E. Amir, E. Brewer, R. W. Brodersen, F. Burghart, A. Burstein, Y. Chang, A. Fox, J. M. Gilbert, R. Han, R. H. Katz, A. C. Long, D. G. Messerschmitt, and J. M. Rabaey: 1996, 'A low-power, lightweight unit to provide ubiquitous information access applications and network support for Infopad'. *IEEE Personal Communications* pp. 4–17.

Narendran, B., J. Sienicki, S. Yajnik, and P. Agrawal: 1997, 'Evaluation of an adaptive power and error control algorithm for wireless systems'. In: *Proc. of IEEE Communications Conf.* Montreal, Canada, pp. 349–355.

Perkins, C. E., E. M. Royer, and S. R. Das: 2000, 'Ad hoc On-Demand Distance Vector (AODV) Routing'. IETF Draft MANET Working Group.

Postel, J.: 1981, 'Transmission Control Protocol'. RFC 793.

Price, C. E.: 2000, 'Power-aware scheduling algorithms for wireless networks'. Master's thesis, Washington State University, Pullman.

Price, C. E., K. M. Sivalingam, J.-C. Chen, and P. Agrawal: 2001, 'Power-Aware Scheduling Algorithms for Wireless Networks'. In: *Proc. Intl. Conference on Intelligence Computing and VLSI*. Kalyani, India.

Ramanathan, R. and R. Rosales-Hain: 2000, 'Topology Control of Multihop Wireless Networks using Transmit Power Adjustment'. In: *Proc. IEEE INFOCOM*. Tel-Aviv, Israel.

Salkintzis, A. and P. T. Mathiopoulos (Guest Ed.): 2000, 'The Evolution of Mobile Data Networking'. *IEEE Personal Communications* **3**(2).

Singh, S. and C. S. Raghavendra: 1998, 'PAMAS: Power aware multi-access protocol with signalling for ad hoc networks'. *Computer Communication Review* **28**(3), 5–26.

Singh, S., C. S. Raghavendra, and J. Stephanek: 1999, 'Power-aware broadcasting in mobile ad hoc networks'. Technical report, Oregon State University, Department of Electrical and Computer Engineering.

Sivalingam, K. M., J.-C. Chen, P. Agrawal, and M. Srivastava: 2000, 'Design and analysis of low-power access protocols for wireless and mobile ATM networks'. *ACM/Baltzer Wireless Networks* **6**(1), 73–87.

Srivastava, M.: 2000, 'TUTORIAL: Energy Efficiency in Mobile Computing and Networking'. In: *ACM MobiCom Tutorials*. Boston, MA.

Stemm, M. and R. H. Katz: 1997, 'Measuring and reducing energy consumption of network interfaces in hand-held devices'. *IEICE Transactions on Fundamentals of Electronics, Communications, and Computer Science*.

Stojmenovic, I. and X. Lin: 2000, 'Power-aware localized routing in wireless networks'. In: *Proc. of the IEEE International Parallel and Distributed Processing Symposium*. Cancum, Mexico.

Swann, R. and N. Kingsbury: 1998, 'Bandwidth efficient transmission of MPEG-II video over noisy mobile links'. *Signal Processing* **12**(2), 105–115.

Tiwari, V., S. Malik, and A. Wolfe: 1994, 'Power analysis of embedded software: A first step towards software power minimization'. *IEEE Transactions on Very Large Scale Integration* **2**(4), 437–445.

Tsaoussidis, V. and H. Badr: 2000, 'TCP-Probing: Towards an Error Control Schema with Energy and Throughput Performance Gain'. In: *International Conference on Network Protocols*. Osaka, Japan.

Tsaoussidis, V., H. Badr, X. Ge, and K. Pentikousis: 2000a, 'Energy/Throughput Tradeoffs of TCP Error Control Strategies'. In: *Proc. of 5th IEEE Symposium on Computers and Communications (ISCC) 2000*. France.

Tsaoussidis, V., A. Lahanas, and H. Badr: 2000b, 'Wave & Wait Protocol (WWP): Low Energy, High Throughput for Mobile IP-Devices'. In: *Proc. of 8th IEEE Conference on Networks (ICON) 2000*. Singapore.

Udani, S. and J. Smith: 1996, 'Power Management in Mobile Computing (A Survey)'. http://www.cis.upenn.edu/~udani/papers.html. University of Pennsylvania.

Vaidya, N.: 2000, 'TUTORIAL: Mobile Ad hoc Networks: Routing, MAC and Transport Issues'. In: *ACM MobiCom Tutorials*. Boston, MA.

WAP Forum: 2000, 'Wireless application protocol (WAP) architecture specification'. http://www.wapforum.org/what/technical.htm.

Weiser, M., B. Welch, A. Demers, and S. Shenker: 1994, 'Scheduling for Reduced CPU Energy'. In: *USENIX Association, Proc. of First Symposium on Operating Systems Design and Implementation (OSDI)*. Monterey, CA.

Wieselthier, J. E., G. D. Nguyen, and A. Ephremides: 2000, 'On the Construction and Energy-Efficient Broadcast and Multicast Trees in Wireless Networks'. In: *Proc. IEEE INFOCOM*. Tel-Aviv, Israel.

Woesner, H., J.-P. Ebert, M. Schlager, and A. Wolisz: 1998, 'Power-saving mechanisms in emerging standards for wireless LANs: the MAC level perspective'. *IEEE Personal Communications* **5**(3), 40–48.

Woo, M., S. Singh, and C. S. Raghavendra: 1998, 'Power aware routing in mobile ad hoc networks'. In: *Proc. ACM MobiCom*. Dallas, TX, pp. 181–190.

Zorzi, M. and R. Rao: 1997a, 'Energy constrained error control for wireless channels'. *IEEE Personal Communications*.

Zorzi, M. and R. R. Rao: 1997b, 'Error control and energy consumption in communications for nomadic computing'. *IEEE Transactions on Computers* **46**(3), 279–289.

Zorzi, M. and R. R. Rao: 2000, 'Energy efficiency of TCP in a local wireless environment'. *ACM/Baltzer Mobile Networks and Applications*. (Accepted for Publication).

Zorzi (Guest Ed.), M.: 1998, 'Energy management in personal communications and mobile computing'. *IEEE Personal Communications* **5**(3).

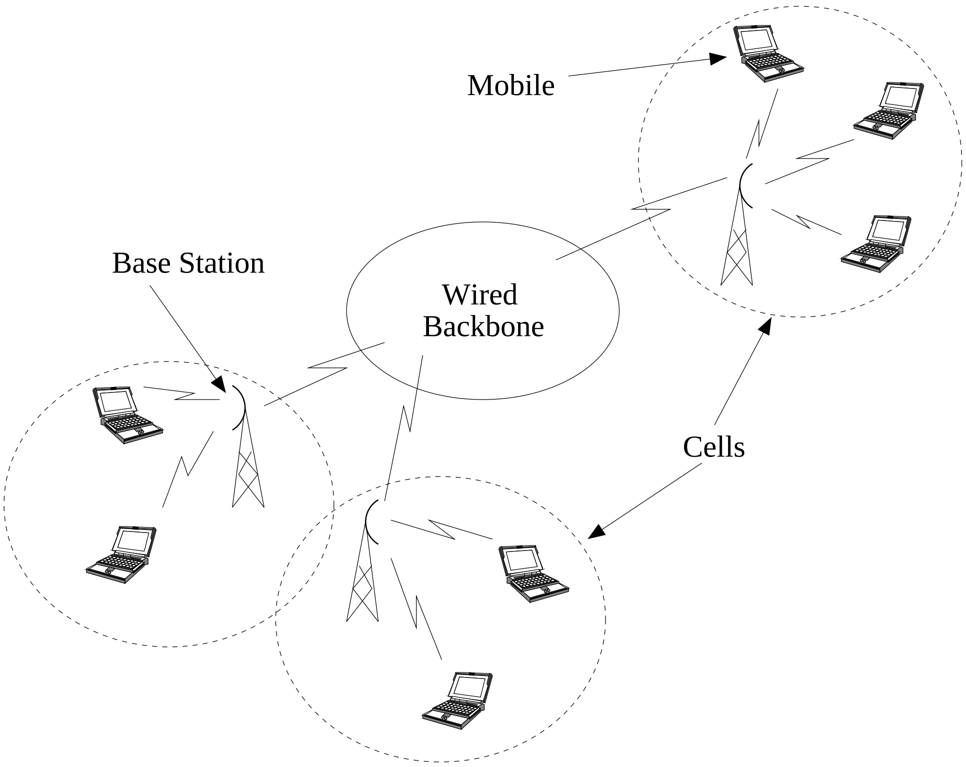


*Figure 1.* Infrastructure wireless network architecture.

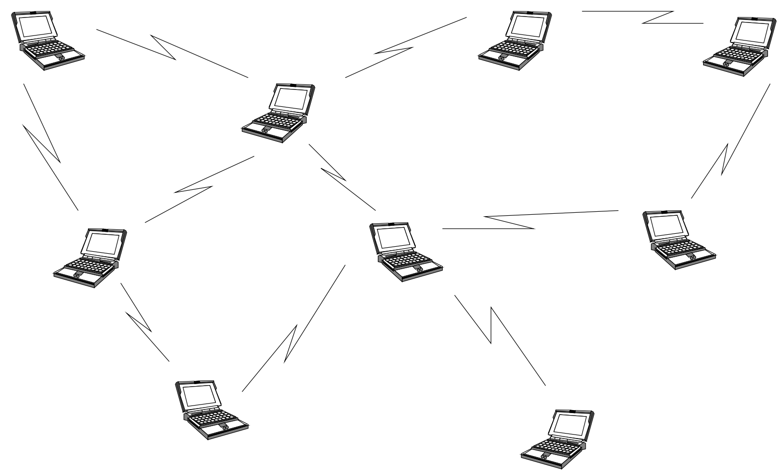

*Figure 2.* Ad Hoc wireless network architecture.

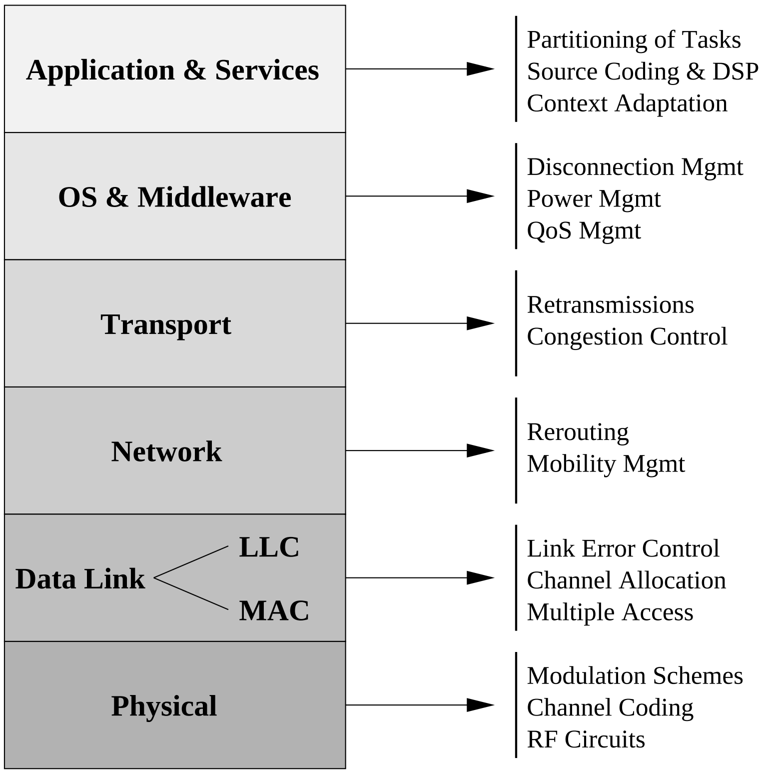


*Figure 3.* Protocol stack of a generic wireless network, and corresponding areas of energy efficient research.

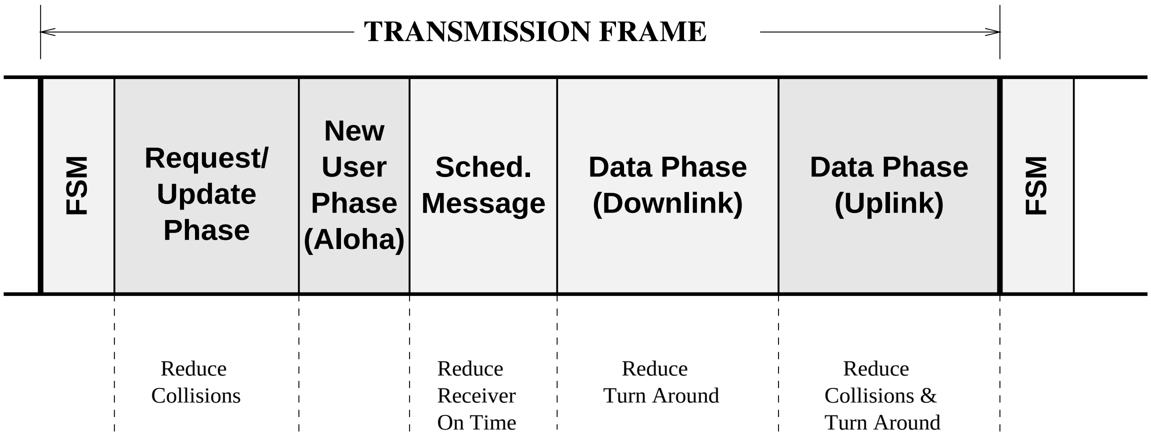


*Figure 4.* EC-MAC protocol frame structure.

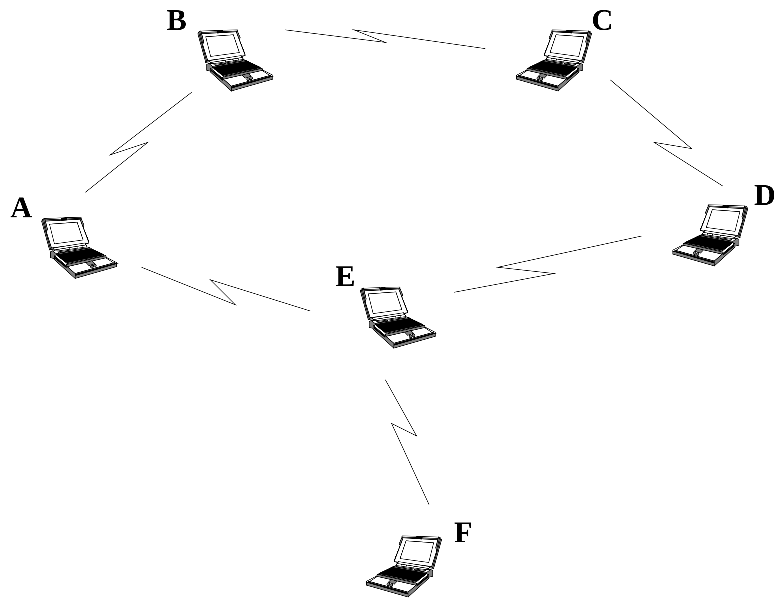


*Figure 5.* Example ad hoc topology.

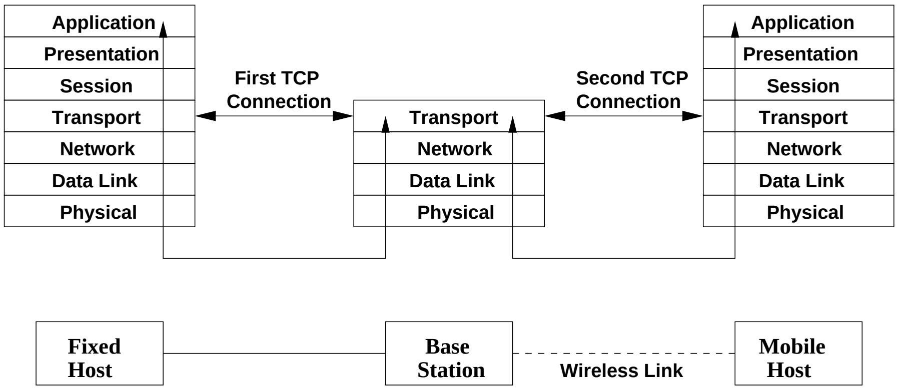


*Figure 6.* Split connection TCP approach.

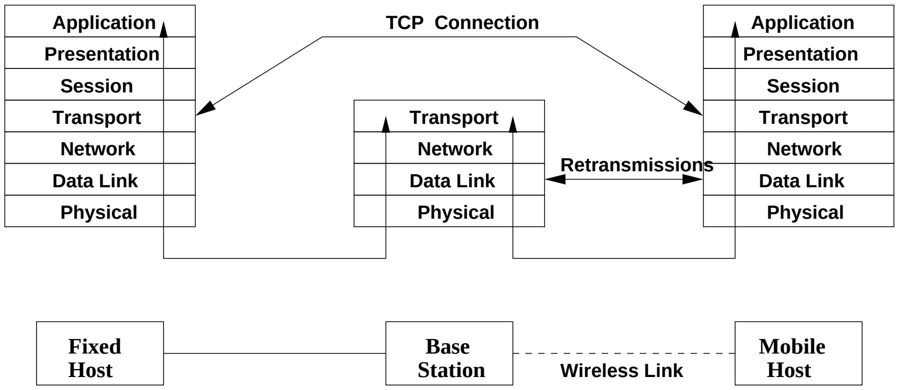


*Figure 7.* Link layer TCP approach.