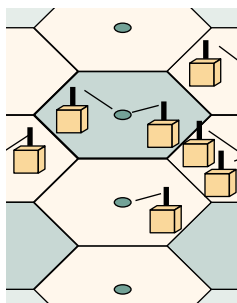# DESIGN CHALLENGES FOR ENERGY-CONSTRAINED AD HOC WIRELESS NETWORKS

ANDREA J. GOLDSMITH, STANFORD UNIVERSITY
STEPHEN B. WICKER, CORNELL UNIVERSITY

Ad hoc wireless networks enable new and exciting applications, but also pose significant technical challenges. We show that cross layer design of these protocols is imperative to meeting emerging application requirements, particularly when energy is a limited resource.

ABSTRACT

Ad hoc wireless networks enable new and exciting applications, but also pose significant technical challenges. In this article we give a brief overview of ad hoc wireless networks and their applications with a particular emphasis on energy constraints. We then discuss recent advances in the link, multiple access, network, and application protocols for these networks. We show that cross-layer design of these protocols is imperative to meeting emerging application requirements, particularly when energy is a limited resource.

ad hoc [Latin.]
a. **For this purpose, to this end; for the particular purpose in hand or in view.**
b. *attrib. or as* **adj. Devoted, appointed, and so on, to or for some particular purpose.**
Oxford English Dictionary, 2nd Edition

INTRODUCTION

An ad hoc wireless network is a collection of wireless mobile nodes that self-configure to form a network without the aid of any established infrastructure, as shown in Fig. 1. Without an inherent infrastructure, the mobiles handle the necessary control and networking tasks by themselves, generally through the use of distributed control algorithms. Multihop connections, whereby intermediate nodes send the packets toward their final destination, are supported to allow for efficient wireless communication between parties that are relatively far apart. Ad hoc wireless networks are highly appealing for many reasons. They can be rapidly deployed and reconfigured. They can be tailored to specific applications, as implied by Oxford's definition. They are also highly robust due to their distributed nature, node redundancy, and the lack of single points of failure. These characteristics

are especially important for military applications, and much of the groundbreaking research in ad hoc wireless networking was supported by the (Defense) Advanced Research Projects Agency (DARPA) [1–3]. Despite much research activity over the last several decades on wireless communications in general, and ad hoc wireless networks in particular, there remain many significant technical challenges in the design of these networks. In this tutorial we describe some of these technical challenges and possible approaches to solving them, with a special emphasis on how finite node energy impacts each layer of the network protocol stack.

The lack of infrastructure inherent to ad hoc wireless networks is best illustrated by contrast with the most prevalent wireless networks today: cellular systems and wireless local area networks (WLANs). In cellular telephone networks the geographic area of interest is divided into regions called cells. A mobile terminal located in a given cell communicates directly with a base station located at or near the center of each cell, as shown in Fig. 2. Thus, there is no peer-to-peer communication between mobiles. All communication is via the base station through single-hop routing, although recent work indicates that multihop routing to the base station can significantly improve performance [4–6]). Each base station is connected by a high-speed link (typically wired) to a mobile switching center (MSC) that in turn is connected to the public switched telephone network (PSTN). The base stations and MSC perform all control and networking functions, including authentication, call routing, and handoff. The mobile units that constitute the wireless portion of this network depend entirely on the base station/MSC/PSTN infrastructure for connectivity and centralized control. Most WLANs have a similar centralized single-hop architecture: mobile nodes communicate directly with a centralized access point that is connected to the backbone Internet, and the access point performs all networking and control functions for the mobile nodes. In contrast, an ad hoc wireless network has peer-to-peer communication, distributed networking and control functions among all nodes, and multihop routing.

This discussion should not be taken to mean that ad hoc wireless networks are completely flat; that is, cannot have any infrastructure or

pre-established node hierarchy. Indeed, many ad hoc wireless networks form a backbone infrastructure from a subset of nodes in the network to improve network reliability and capacity [7]. Similarly, some nodes may be chosen to perform as base stations for neighboring nodes [8]. The distinguishing emphasis in the ad hoc approach lies in the design requirements. Ad hoc wireless networks may exploit infrastructure to improve network performance. However, while the infrastructure provides the side benefit of enhanced performance, it is not a fundamental design principle of the network.

Ad hoc networks are quite common in the wired world. Indeed, most LANs, metropolitan area networks (MANs), and wide area networks (WANs), including the Internet, have an ad hoc structure. However, the broadcast nature of the radio channel introduces characteristics in ad hoc wireless networks that are not present in their wired counterparts. In particular, a radio channel allows a node to transmit a signal directly to any other node. The link signal-to-interference-plus-noise power ratio (SINR) between two communicating nodes will typically decrease as the distance between the nodes increases, and will also depend on the signal propagation and interference environment. Moreover, this link SINR varies randomly over time due to the mobility of the nodes that typically changes the transmission distance, propagation environment, and interference characteristics. Link SINR determines the communication performance of the link: the data rate and associated probability of packet error or bit error (bit error rate or BER) that can be supported on the link. Links with very low SINRs are not typically used due to their extremely poor performance, leading to partial connectivity among all nodes in the network, as shown in Fig. 1. However, link connectivity is not a binary decision, since nodes can back off on their transmission rate or increase their transmit power as link SINR degrades and still maintain connectivity [9, 10]. This is illustrated by the different line widths corresponding to different link qualities in Fig. 1. Link connectivity also changes as nodes enter and leave the network, and this connectivity can be controlled by adapting the transmit power of existing network nodes to the presence of a new node [11].

The flexibility in link connectivity that results from varying link parameters such as power and data rate has major implications for routing. Nodes can send packets directly to their final destination via single-hop routing as long as the link SINR is above some minimal threshold. However, single-hop routing can cause excessive interference to surrounding nodes. Routing over a single hop may also require a relatively low rate or have a high probability of bit or packet error if the link SINR is low, thereby introducing excessive delays. Alternatively, packets can be forwarded from source to destination by intermediate nodes at a link rate commensurate with the forwarding link SINR. Routing via forwarding by intermediate nodes is called multihop routing. Several recent research results indicate that ideal multihop routing significantly increases the capacity of ad hoc wireless networks [4, 12], but achieving these gains through a decen-
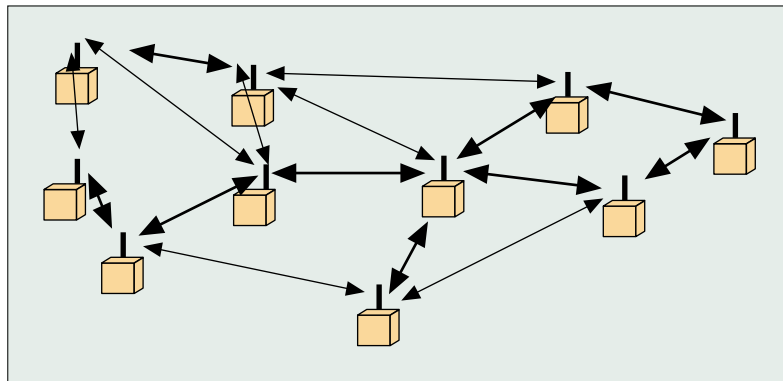
■ Figure 1. *Ad hoc network structure.*
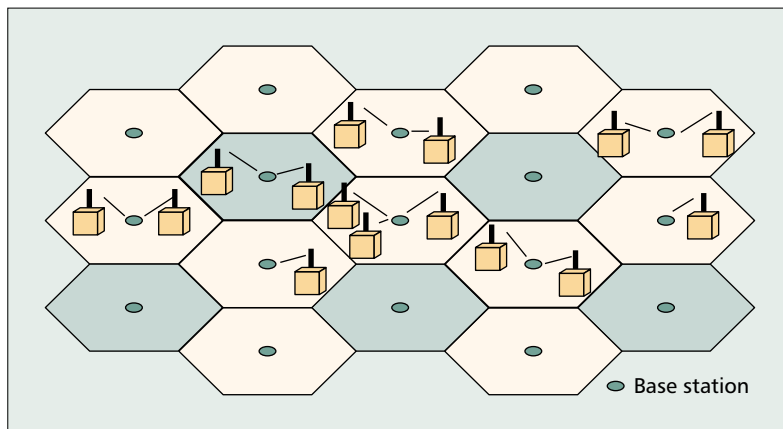
■ Figure 2. *Cellular system architecture.*

tralized routing strategy remains elusive. The channel and network dynamics of ad hoc wireless systems coupled with multihop routing make it difficult to support multimedia requirements of high speed and low delay. However, flexibility in the link, access, network, and application protocols can be exploited to compensate for and even take advantage of these dynamics.

Energy constraints are not inherent to all ad hoc wireless networks. Devices in an ad hoc wireless network may be stationary and attached to a large energy source. Mobile devices may be part of a large vehicle, such as a car or tank, that can generate significant amounts of power over the long term. However, many ad hoc wireless network nodes will be powered by batteries with limited lifetime. Some of the most exciting applications for ad hoc wireless networks follow this paradigm. This motivates our focus on energy constraints. Devices with rechargeable batteries must conserve energy to maximize time between recharging. Of particular interest are devices that cannot be recharged, that is, sensors that are imbedded in walls or dropped into a remote region. Energy constraints impact both the hardware operation and the signal transmission associated with node operation. It is often assumed that the transmit power associated with packet transmission dominates power consumption. However, signal processing associated with packet transmission and reception, and even hardware operation in a standby mode, consume nonnegligible power as well [13, 14]. This entails

> *the success of ad hoc wireless networks hinges on making them sufficiently flexible so that there can be accidental successes. Therein lies the design dilemma for ad hoc wireless networks.*

interesting energy trade-offs across protocol layers. At the link layer many communications techniques that reduce transmit power require a significant amount of signal processing. It is widely assumed that the energy required for this processing is small and continues to decrease with ongoing improvements in hardware technology [13, 15]. However, the results in [14] suggest that these energy costs are still significant. This would indicate that energy-constrained systems must develop energy-efficient processing techniques that minimize power requirements across all levels of the protocol stack and also minimize message passing for network control, since these entail significant transmitter and receiver energy costs. Sleep modes for nodes must be similarly optimized, since these modes conserve standby energy but may entail energy costs at other protocol layers due to associated complications in access and routing. The hardware and operating system design in the node can also be optimized to conserve energy: techniques for this optimization are described in [14, 16].

Another important characteristic of ad hoc wireless networks is mobility in the network nodes. Mobility impacts all layers of the network protocol stack. At the link layer it determines how fast the link characteristics change and whether or not the link connectivity is stable over time. At the medium access control (MAC) layer it affects how long measurements regarding channel and interference conditions remain in effect and how scheduling algorithms perform. At the network layer mobility has major implications for the performance of different routing protocols. The impact of mobility on network performance ultimately dictates which applications can be supported on a highly mobile network. The impact of mobility on ad hoc wireless network design will be discussed in more detail throughout the article.

The remainder of this article is organized as follows. In the next section we discuss applications for ad hoc wireless networks, including data networks, home networks, device networks, sensor networks, and distributed control. We address cross-layer design in ad hoc wireless networks: what it is, why it is needed, and how it can be done. Link layer design issues are discussed, where we describe the fundamental capacity limits at the link layer, and high-performance link layer designs that include coding, multiple antennas, power control, and adaptive resource allocation. We discuss the MAC layer design issues, including the trade-offs inherent to frequency/time/code channelization and the assignment of users to these channels via random access or scheduling. This section also describes the role power control can play in multiple access. Networking issues such as neighbor discovery, network connectivity, scalability, routing, and network capacity are outlined. We describe techniques for the network to adapt to the application requirements and the application to adapt to network capabilities. A summary of the design challenges inherent to ad hoc wireless networks and some parting thoughts make up the final section.

APPLICATIONS

Since the ad hoc wireless network paradigm tailors the network design to the intended application, it will be useful to consider potential applications in some detail. In what follows we consider both military and commercial applications. We see that several design requirements are common to both types of systems, especially the need for energy efficiency. Military applications often require the self-configuring nature and lack of infrastructure inherent to ad hoc wireless networks, even if it results in a significant cost or performance penalty. The lack of infrastructure is also highly appealing for commercial systems, since it precludes a large investment to get the network up and running, and deployment costs may then scale with network success. Other commercial advantages include ease of network reconfiguration and reduced maintenance costs. However, these advantages must be balanced against any performance penalty resulting from the need for distributed network control.

In this section we consider the following applications: data networks, home networks, device networks, sensor networks, and distributed control systems. Note that this list is by no means comprehensive, and in fact the success of ad hoc wireless networks hinges on making them sufficiently flexible so that there can be accidental successes. Therein lies the design dilemma for ad hoc wireless networks. If the network is designed for maximum flexibility to support many applications (a one-size-fits-all network), it will be difficult to tailor the network to different application requirements. This will likely result in poor performance for some applications, especially those with high rate requirements or stringent delay constraints. On the other hand, if the network is tailored to a few specific applications, designers must predict in advance what these "killer applications" will be — a risky proposition. Ideally an ad hoc wireless network must be sufficiently flexible to support many different applications while adapting its performance to the given set of applications in operation at any given time. The cross layer design discussed below provides this flexibility along with the ability to tailor protocol design to the energy constraints in the nodes.

DATA NETWORKS

Ad hoc wireless networks can support data exchange between laptops, palmtops, personal digital assistants (PDAs), and other information devices. We focus on two types of wireless data networks: LANs with coverage over a relatively small area (a room, floor, or building) and MANs with coverage over several square miles (a metropolitan area or battlefield). The goal of WLANs is to provide peak data rates on the order of 10–100 Mb/s, similar to what is available on a wired LAN, for low-mobility and stationary users. Commercial WLAN standards such as 802.11a and 802.11b provide data rates on this order; however, the individual user rates are much less if there are many users accessing the system. Moreover, these commercial LANs are not really based on an ad hoc structure. The

normal 802.11 network configuration is a star topology with one wireless access point and single-hop routing from the mobile units to the access point. While the 802.11 standard does support a peer-to-peer architecture in the form of the Independent Base Service Set (IBSS) configuration option, it is not widely used and its performance is somewhat poor [17].

Wireless MANs typically require multihop routing since they cover a large area. The challenge in these networks is to support high data rates, in a cost-effective manner, over multiple hops, where the link quality of each hop is different and changes with time. The lack of centralized network control and potential for high-mobility users further complicates this objective. Military programs such as DARPA's Global Mobile Information Systems (GLOMO) have invested much time and money in building high-speed wireless MANs that support multimedia, with limited success [18, 19]. Wireless MANs have also permeated the commercial sector, with Metricom the best example [20]. While Metricom did deliver fairly high data rates throughout several major metropolitan areas, the deployment cost was quite large and significant demand never materialized. Metricom filed for protection under Chapter XI of the Federal Bankruptcy Code in fall 2000.

Note that energy efficiency is a major issue in the design of wireless data networks. The canonical example of an ad hoc wireless data network is a distributed collection of laptop computers. Laptops are highly limited in battery power, so power must be conserved as much as possible. In addition, a laptop acting as a router for other laptops could drain its battery forwarding packets for other users. This would leave no power for the laptop user and would initiate a change in network topology. Thus, these networks must conserve battery power in all communication functions, and devise routing strategies that use residual power at each node of the network in a fair and efficient manner.

HOME NETWORKS

Home networks are envisioned to support communication between PCs, laptops, PDAs, cordless phones, smart appliances, security and monitoring systems, consumer electronics, and entertainment systems anywhere in and around the home. The applications for such networks are limited only by the imagination. For example, using a PDA in the bedroom one could scan stored music titles on a PC and direct the bedroom stereo to play a favorite piece, check the temperature in the living room and increase it by a few degrees, check the daily TV programming from the Internet and direct the VCR to record a show that night, access voice messages and display them using a voice-to-text conversion software, check stocks on the Internet and send selling instructions to a broker, and start the coffee maker and toaster, all without getting up from the bed. Other applications include smart rooms that sense people and movement, and adjust light and heating accordingly, "aware homes" that network sensors and computers for assisted living of seniors and those with disabilities, video or sensor monitoring systems with the intelligence to coordinate and interpret data and alert the home owner and the appropriate police or fire department of unusual patterns, intelligent appliances that coordinate with each other and with the Internet for remote control, software upgrades, and to schedule maintenance, and entertainment systems that allow access to a VCR, Tivo box, or PC from any television or stereo system in the home [21–24].

There are several design challenges for such networks. One of the biggest is the need to support the varied quality of service (QoS) requirements for different home networking applications. QoS in this context refers to the requirements of a particular application, typically data rates and delay constraints, which can be quite stringent for home entertainment systems. Other big challenges include cost and the need for standardization, since all of the devices being supported on this type of home network must follow the same networking standard. Note that the different devices accessing a home network have very different power constraints: some will have a fixed power source and be effectively unconstrained, while others will have very limited battery power and may not be rechargeable. Thus, one of the biggest challenges in home network design is to leverage power in unconstrained devices to take on the heaviest communication and networking burden, such that the networking requirements for all nodes in the network, regardless of their power constraints, can be met.

One approach to home networking is to use an existing WLAN standard such as 802.11 [25]. But 802.11 has several limitations for this type of application. First, it most commonly supports a star architecture with a single access point and all devices talking directly to this access node. This star architecture eliminates the benefits of multihop routing, and while multihop routing is possible in 802.11, as noted above, its performance is poor. In addition, 802.11 uses a statistical multiple access protocol, which makes it difficult to support the quality required in home entertainment systems. 802.11b is also somewhat limited in data rate (1–10 Mb/s), and while the 802.11a standard supports much higher rates (10–70 Mb/s), it is mainly designed for packet data applications and not media streaming. While protocols to support media streaming on top of 802.11a are being developed (802.11e), this type of overlay will likely be insufficient to provide high-quality wireless home entertainment.

A natural choice for home networking is a peer-to-peer ad hoc wireless network. Much of the communication in home networks will take place between peer devices, so peer-to-peer communication eliminates the overhead of going through a centralized node. In addition, many of the devices in a home network will be low-power or battery-limited. In an ad hoc wireless network these devices need only communicate with their nearest neighbors (typically a short distance away) to maintain connectivity with (all) other devices in the home. Thus, multihop routing will be very beneficial to such devices in terms of energy savings. Most home networking applications involve stationary or low-mobility nodes, so the protocols need not support high mobility. Ad hoc wireless

Home networks are envisioned to support communication between PCs, laptops, PDAs, cordless phones, smart appliances, security and monitoring systems, consumer electronics, and entertainment systems anywhere in and around the home. The applications for such networks are limited only by the imagination.

Sensor networks have enormous potential for both consumer and military applications. For the military, it is now painfully clear that the wars of the 21st century will differ significantly from those of the 20th. Enemy targets will be small, mobile, and generally found in extremely hostile terrain.

networks will be challenged to provide high-quality media streaming for home entertainment, and this is an open area of active research.

Home networking is being pushed strongly by the HomeRF working group, which has developed an open industry standard for such networks that combines a centralized and peer-to-peer structure [23]. The working group for HomeRF was initiated by Intel, HP, Microsoft, Compaq, and IBM. The main component of the HomeRF protocol is its Shared Wireless Access Protocol (SWAP). SWAP is designed to carry both voice and data traffic and to interoperate with the PSTN and the Internet. SWAP is a combination of a managed network that provides isochronous services (e.g., real-time voice and video) via a centralized network controller (the main home PC) along with an ad hoc peer-to-peer network for data devices. The centralized network controller is not required, but greatly facilitates providing dedicated bandwidth to isochronous applications. Bandwidth sharing is enabled by frequency-hop spread spectrum at 50 hops/s. HomeRF also supports a time-division service for delivery of interactive voice and other time-critical services, and a random access protocol for high-speed packet data. The transmit power for HomeRF is specified at 100 mW, which provides a data rate of 1–2 Mb/s. However, in August 2000 the FCC authorized a fivefold increase in the HomeRF bandwidth, effectively increasing data rates to 10 Mb/s. The range of HomeRF covers a typical home and backyard. HomeRF products operating in the 2.4 GHz band are currently on the market in the $100–$200 price range. Details on these products can be found at http://www.homerf.org.

DEVICE NETWORKS

Device networks support short-range wireless connections between devices. Such networks are primarily intended to replace inconvenient cabled connections with wireless connections. Thus, the need for cables and the corresponding connectors between cell phones, modems, headsets, PDAs, computers, printers, projectors, network access points, and other such devices is eliminated. Clearly many of these devices have limited battery life, but are generally rechargeable. Thus, device networks require energy efficiency.

The main technology driver for such networks is Bluetooth [8, 26]. The Bluetooth standard is based on a tiny microchip incorporating a radio transceiver that is built into digital devices. The transceiver takes the place of a connecting cable for electronic devices. Up to eight Bluetooth devices can form a star topology network (a piconet) with one node acting as a master and the other nodes acting as slaves. The master node is responsible for synchronization and scheduling transmissions of the slave nodes. Piconets can also be interconnected, leading to a multihop topology. Bluetooth is mainly for short-range communications, such as from a laptop to a nearby printer or from a cell phone to a wireless headset. Its normal range of operation is 10 m (at 1 mW transmit power), and this range can be increased to 100 m by increasing the transmit power to 100 mW. The system operates in the unregulated 2.4 GHz frequency band; hence, it

can be used worldwide without any licensing issues. The Bluetooth standard provides 1 data channel at 721 kb/s and up to three voice channels at 56 kb/s for an aggregate bit rate on the order of 1 Mb/s. Networking is done via a packet switching protocol based on frequency hopping at 1600 hops/s. Energy constraints played a large role in the design of Bluetooth, with a goal of using as little energy from the host device as possible. Bluetooth uses a range of techniques in its hardware, communication, and networking protocols to preserve energy, including power-efficient modulation, a limited transmission range, smart packet detection, and intelligent sleep scheduling [26].

The Bluetooth standard was developed jointly by 3Com, Ericsson, Intel, IBM, Lucent, Microsoft, Motorola, Nokia, and Toshiba. Over 1300 manufacturers have now adopted the standard, and products compatible with Bluetooth are starting to appear on the market now. Some of the products currently available include a wireless headset for cell phones (Ericsson), a wireless USB or RS232 connector (RTX Telecom, Adayma), wireless PCMCIA cards (IBM), and wireless set-top boxes (Eagle Wireless). The prognosis for Bluetooth has been varied, progressing from the early euphoria of the late 1990s to pessimism and claims of premature death in 2000 to the current outlook of guarded optimism.

SENSOR NETWORKS

Sensor networks have enormous potential for both consumer and military applications. For the military, it is now painfully clear that the wars of the 21st century will differ significantly from those of the 20th. Enemy targets will be small, mobile, and generally found in extremely hostile terrain. If the war in Afghanistan is any indication, the targets in future combats will be small and difficult to detect from great distances. Future military missions will therefore require that sensors and other intelligence gathering mechanisms be placed close to their intended targets. The potential threat to these mechanisms is therefore quite high, so it follows that the technology used must be highly redundant and require as little support as possible from friendly forces. An apparent solution to these constraints lies in large arrays of passive electromagnetic, optical, chemical, and biological sensors. These can be used to identify and track targets, and can also serve as a first line of detection for various types of attacks. A third function lies in the support of the movement of unmanned robotic vehicles. For example, optical sensor networks can provide networked navigation, routing vehicles around obstacles while guiding them into position for defense or attack. The design considerations for some industrial applications are quite similar to those for military applications. In particular, sensor arrays can be deployed and used for remote sensing in nuclear power plants, mines, and other industrial venues.

Examples of sensor networks for the home environment include electricity, gas, and water meters that can be read remotely through wireless connections. The broad use of simple metering devices within the home can help consumers identify and regulate devices like air conditioners and hot water heaters that are significant

consumers of power and gas. Simple attachments to power plugs can serve as the metering and communication devices for individual appliances. One can imagine a user tracking various types of information on home energy consumption from a single terminal: the home computer. Remote control of television usage and content could be monitored in similar ways. Another important home application is smoke detectors that could not only monitor different parts of the house but also communicate to track the spread of fire. Such information could be conveyed to local fire-fighters before they arrived on the scene along with house blueprints. A similar type of array could be used to detect the presence and spread of gas leaks or other toxic fumes.

Sensor arrays also have great potential for use at the sites of large accidents. One may wish to consider, for example, the use of remote sensing in rescue operations following the collapse of a building. Sensor arrays could be rapidly deployed at the site of an accident and used to track heat, natural gas, and toxic substances. Acoustic sensors and triangulation techniques could be used to detect and locate trapped survivors. It may even be possible to avert such tragedies altogether through the use of sensor arrays. The collapse of walkways and balconies, for example, can be predicted and tragedy averted by building stress and motion sensors into the structures from the outset. One can imagine large numbers of low-cost low-power sensors being directly inserted into the concrete before it is poured. Material fatigue can be detected and tracked over time throughout the structure. Such sensors must be robust and self-configuring, and would require a very long lifetime, commensurate with the lifetime of the structure.

Most sensors will be deployed with non-rechargeable batteries. The problem of battery lifetime in such sensors may be averted through the use of ultra-small energy-harvesting radios. Research on such radios, coined the PicoRadio, promise radios smaller than 1 cm^3, weighing less than 100 g, with a power dissipation level below 100 μW [27]. This low level of power dissipation enables nodes to extract sufficient power from the environment — energy harvesting — to maintain operation indefinitely. Such picoradios open up new applications for sensor deployment in buildings, homes, and even the human body.

In short, important applications of the future are enabled by large numbers of very small, lightweight, battery-powered sensors. These sensors must be easily and rapidly deployed in large numbers and, once deployed, must form a suitable network with a minimum of human intervention. All of these requirements must be met with a minimum of power consumption due to battery limitations and, for many applications, the inability to recharge these batteries

DISTRIBUTED CONTROL SYSTEMS

Ad hoc wireless networks enable distributed control, with remote plants, sensors and actuators linked together via wireless communication channels. Such networks are imperative for coordinating unmanned mobile units, and greatly reduce maintenance and reconfiguration costs over distributed control systems with wired com-

munication links. Ad hoc wireless networks are currently under investigation for supporting coordinated control of multiple vehicles in an automated highway system (AHS), remote control of manufacturing and other industrial processes, and coordination of unmanned airborne vehicles (UAVs) for military applications.

Current distributed control designs provide excellent performance as well as robustness to uncertainty in model parameters. However, these designs are based on closed-loop performance that assumes a centralized architecture, synchronous clocked systems, and fixed topology. Consequently, these systems require that the sensor and actuator signals be delivered to the controller with a small fixed delay. Ad hoc wireless networks cannot provide any performance guarantee in terms of data rate, delay, or loss characteristics: delays are typically random and packets may be lost. Unfortunately, most distributed controllers are not robust to these types of communication errors, and effects of small random delays can be catastrophic [28, 29]. Thus, distributed controllers must be redesigned for robustness to the random delays and packet losses inherent to wireless networks. Ideally, the ad hoc wireless network can also be tailored to the requirements of the controller. This is a relatively new area of research: recent results in this area can be found in [30] and the references therein. Energy constraints in distributed control systems will be highly application-dependent: cars ON an automated highway will have a large renewable energy source, whereas sensors in most manufacturing applications will have nonrechargeable batteries.

CROSS-LAYER DESIGN

The different applications discussed in the previous section have a wide range of network requirements as well as different energy constraints for different network nodes. The network requirements must be met despite variations in the link characteristics on each hop, the network topology, and the node traffic. It is very difficult to ensure performance of the network or the support of real-time or mission critical data in the face of these random variations. There has been significant research directed toward energy constraints, application requirements, and network variability at different levels of the network protocol stack. Examples include multiple antennas, coding, power control, and adaptive techniques at the link layer, power control and scheduling at the MAC layer, energy-constrained and delay-constrained routing at the network layer, and application adaptation at the application layer. These techniques will be discussed in more detail in the following sections. However, these research efforts have mainly targeted isolated components of the overall network design, thereby ignoring important interdependencies. Specifically, current ad hoc wireless network protocol design is largely based on a layered approach, as shown in Fig. 3. In this model each layer in the protocol stack is designed and operated independently, with interfaces between layers that are static and independent of the individual network constraints and applications. This paradigm has

One may wish to consider, for example, the use of remote sensing in rescue operations following the collapse of a building. Sensor arrays could be rapidly deployed at the site of an accident and used to track heat, natural gas, and toxic substances.

It is this integrated approach to adaptive networking — how each layer of the protocol stack should respond to local variations given adaptation at higher layers — that forms the biggest challenge in adaptive protocol design.
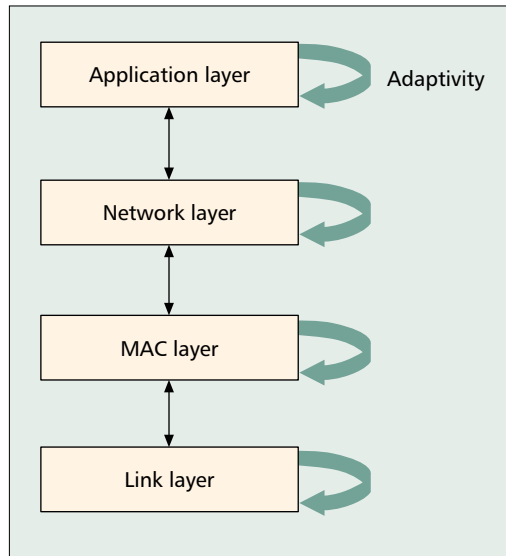
■ Figure 3. *The OSI model for protocol stack design and operation.*

greatly simplified network design and led to the robust scalable protocols in the Internet. However, the inflexibility and suboptimality of this paradigm result in poor performance for ad hoc wireless networks in general, especially when energy is a constraint or the application has high bandwidth needs and/or stringent delay constraints. To meet these requirements a cross-layer protocol design that supports adaptivity and optimization across multiple layers of the protocol stack is needed.

In an adaptive cross-layer protocol stack, the link layer can adapt rate, power, and coding to meet the requirements of the application given current channel and network conditions. The MAC layer can adapt based on underlying link and interference conditions as well as delay constraints and bit priorities. Adaptive routing protocols can be developed based on current link, network, and traffic conditions. Finally, the application layer can utilize a notion of soft QoS that adapts to the underlying network conditions to deliver the highest possible application quality. It is important that the protocols at each layer not be developed in isolation, but rather within an integrated and hierarchical framework to take advantage of the interdependencies between them. These interdependencies revolve around adaptivity at each layer of the protocol stack, general system constraints, such as energy and mobility, and the application(s) the network is supporting.

Adaptivity at each layer of the protocol stack should compensate for variations at that layer based on the timescale of these variations. Specifically, variations in link SINR are very fast, on the order of microseconds for vehicle-based users. Network topology changes more slowly, on the order of seconds, while variations of user traffic may change over tens to hundreds of seconds. The different timescales of the network variations suggest that each layer should attempt to compensate for variation at that layer first. If adapting locally is unsuccessful, information should be exchanged with other layers for a more general response. For example, suppose

the link connectivity (link SINR) in the wireless link of an end-to-end network connection is weak. By the time this connectivity information is relayed to a higher level of the protocol stack (i.e., the network layer for rerouting or the application layer for reduced-rate compression), the link SINR will most likely have changed. Therefore, it makes sense for each protocol layer to adapt to variations local to that layer. If this local adaptation is insufficient to compensate for the local performance degradation, the performance metrics at the next layer of the protocol stack will degrade as a result. Adaptation at this next layer may then correct or at least mitigate the problem that could not be fixed through local adaptation. For example, consider again the weak link scenario. Link connectivity can be measured quite accurately and quickly at the link level. The link protocol can therefore respond to weak connectivity by increasing its transmit power or its error correction coding. This will correct for variations in connectivity due to, for example, multipath flat fading. However, if the weak link is caused by something difficult to correct for at the link layer (e.g., the mobile unit is inside a tunnel), it is better for a higher layer of the network protocol stack to respond by, say, delaying packet transmissions until the mobile leaves the tunnel. Similarly, if nodes in the network are highly mobile, link characteristics and network topology will change rapidly. Informing the network layer of highly mobile nodes might change the routing strategy from unicast to broadcast in the general direction of the intended user. It is this integrated approach to adaptive networking — how each layer of the protocol stack should respond to local variations given adaptation at higher layers — that forms the biggest challenge in adaptive protocol design.

Energy conservation also requires a cross layer design. For example, Shannon theory indicates that the energy required to communicate one bit of information decreases as the bit time increases [31]. Thus, energy can be conserved by transmitting a bit over a longer period of time. However, this will clearly impact the MAC protocol and also the application. Routing is also an interesting example. The most energy efficient routing protocol in a sensor network may use a centrally-located sensor to forward packets from other sensors. However, the battery of this sensor will be quickly exhausted, which might be undesirable from an application standpoint. Thus, the need for energy efficiency must be balanced against the lifetime of each individual node and the overall life of the network.

The above discussion indicates that in order to support an adaptive cross-layer design, the design and operation of the protocol stack must evolve to that shown in Fig. 4. This figure indicates that information must be exchanged across all layers in the protocol stack. This information exchange allows the protocols to adapt in a global manner to the application requirements and underlying network conditions. In addition, all protocol layers must be jointly optimized with respect to global system constraints and characteristics such as energy and high-mobility nodes. In order to design a protocol stack based on Fig. 4, two fundamental questions must be answered:
• What information should be exchanged across

protocol layers and how should that information be adapted to?
- How should global system constraints and characteristics be factored into the protocol designs at each layer?

In the next several sections we will discuss the design of the different layers in the protocol stack, and then revisit cross-layer design relative to these two questions. Cross-layer design is an active theme in ad hoc wireless network design today. However, there remain many open questions in the understanding and implementation of this design philosophy.

LINK DESIGN ISSUES

The link layer design of energy-constrained ad hoc wireless networks introduces many challenges. Wireless channels are generally a difficult communications medium, with relatively low capacity per unit bandwidth, random amplitude and phase fluctuations due to multipath fading, intersymbol interference due to delay spread, and interference from other nodes due to the broadcast nature of the radio channel. The goal of link layer design in ad hoc wireless networks is to achieve rates close to the fundamental capacity limits of the channel while overcoming channel impairments using relatively little energy. In this section we explore capacity limits of energy-constrained nodes along with the techniques and design strategies to provide good link layer performance under an energy constraint.

FUNDAMENTAL CAPACITY LIMITS

The fundamental capacity of a channel dictates the maximum data rate that can be transmitted over the channel with arbitrarily small probability of error. Channel capacity for wireless channels under average or peak power constraints has been the subject of intense research for many years. This research was pioneered by Claude Shannon in 1948 when he determined the capacity of an additive white Gaussian noise (AWGN) channel to be $C = B\log_2(1 + \text{SNR})$ b/s, where B is the channel bandwidth and SNR is the received signal-to-noise power ratio within this bandwidth. There has been much recent work on obtaining capacity expressions for channel models that better reflect the channels underlying current wireless systems. Recent results in this area include the capacity of single- and multiuser fading channels under different assumptions about channel information [32–35], capacity of diversity channels [36], and capacity of channels with multiple antennas at both the transmitter and receiver [37, 38]. Capacity results for fading channels indicate that power and rate should be increased in good channel conditions and decreased in poor channel conditions. The multiple antenna results indicate that the capacity of wireless channels increases linearly with the number of antennas at the transmitter/receiver; however, this requires perfect channel estimates. Degradation in these estimates can significantly degrade the capacity gains resulting from multiple antennas. In general, the capacity-achieving codes for wireless channels have asymptotically large block lengths. The long codes and complex decoding in this optimal scheme drive the proba-
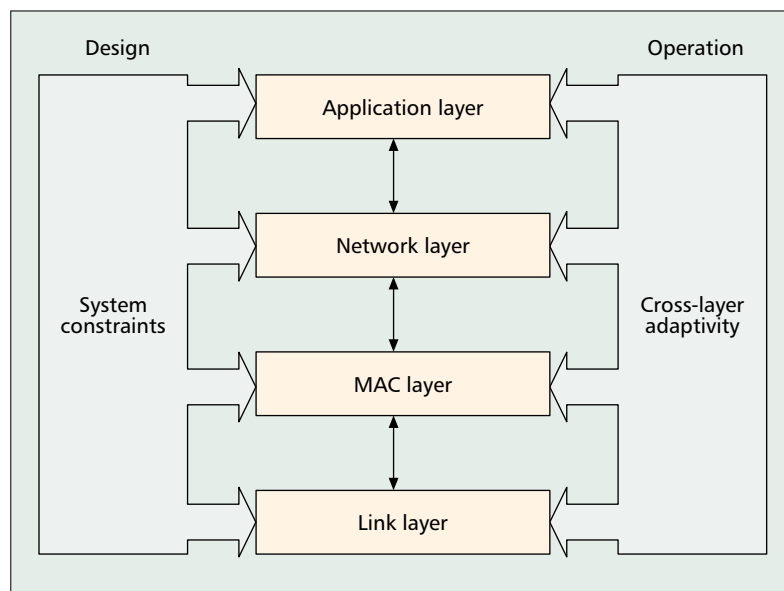

■ Figure 4. *Adaptive cross-layer design and operation.*

bility of error to zero for any data rate below capacity, but the complexity of these schemes makes them hard to approximate with practical implementations.

Channel capacity under a hard transmit energy constraint, as opposed to a peak or average power constraint, is a relatively new area of research. With finite energy it is not possible to transmit any number of bits with asymptotically small error probability. This is easy to see intuitively by considering the transmission of a single bit. The only way to ensure that two different values in signal space, representing the two possible bit values, can be decoded with arbitrarily small error is to make their separation arbitrarily large, which requires arbitrarily large energy. Since arbitrarily small error probability is not possible under a hard energy constraint, a different notion of reliable communication is needed. The pioneering work by Gallager [39] in this area defines reliable communication under a finite energy constraint in terms of the capacity per unit energy. This capacity per unit energy is defined as the maximum number of bits per unit energy that can be transmitted such that the maximum likelihood random coding error exponent is positive. This definition ensures that for all rates below the capacity per unit energy, error probability decreases exponentially with the total energy, although it will not be asymptotically small for finite-energy channels. Gallager also shows that the capacity per unit energy is achieved using an unlimited number of degrees of freedom per transmitted bit. This translates to either very wideband communication [40] or using many symbols per bit, the opposite of high-rate transmission schemes under a power constraint (e.g., M-QAM, with M b/symbol for M large).

Capacity per unit energy is also explored in [31], and these results can be used to obtain the capacity of finite-energy channels in terms of bits [41]. Capacity in bits dictates the maximum number of bits that can be transmitted over a channel using finite energy given some nonzero
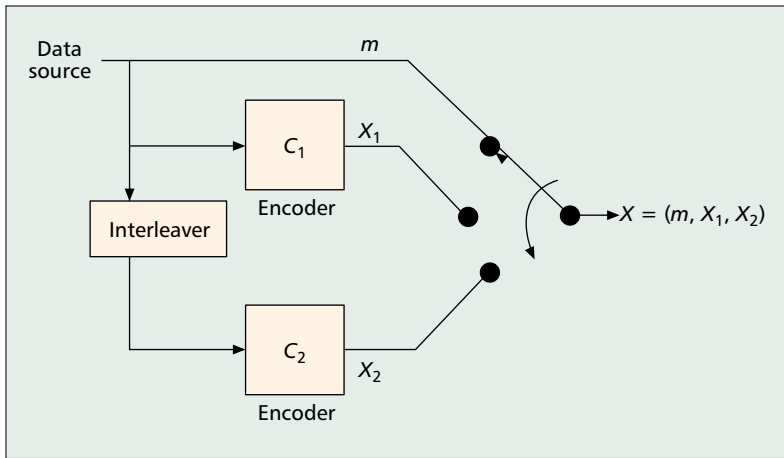
■ Figure 5. *A turbo encoder.*

probability of bit error (recall that this error probability cannot be driven to zero with finite energy). The capacity of a finite-energy channel in bits is an important concept, since it indicates that ad hoc wireless networks with finite energy nodes only have a finite number of bits that a given node can transmit before exhausting its energy. Allocating those bits to the different requirements of the network — information transmission, exchange of routing information, forwarding bits for other nodes, channel estimation, and so on — becomes an interesting and challenging optimization probl em that clearly requires cross-layer design.

CODING

Channel coding can significantly reduce the power required to achieve a given BER and is therefore a common feature in energy-constrained link layer design. Deep space communications engineers have aggressively pursued the use of very complex codes to minimize power. Most wireless systems also use some form of error control coding to reduce power consumption. Conventional error control codes use block or convolutional code designs: the error correction cabability of these
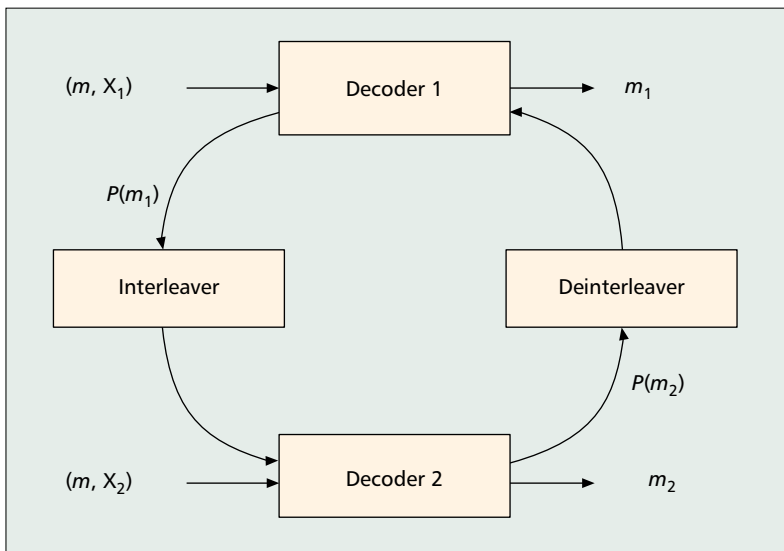
codes is obtained at the expense of an increased signal bandwidth or a lower data rate [42]. Trellis codes, invented by Ungerboeck in the '70s [43], use a joint design of the channel code and modulation to provide good error correction without any bandwidth or rate penalty.

For AWGN channels the challenge of achieving very low BERs with minimum power has for the most part been met by turbo codes [44] and the more general family of codes on graphs with iterative decoding algorithms [45, 46]. While the main ideas behind current research in codes on graphs was introduced by Gallager in 1962, at the time these coding techniques were thought impractical and were generally not pursued by researchers in the field. The landmark 1993 paper by Berrou, Glavieux, and Thitimajshima on turbo codes [44] provided more than enough motivation to revisit Gallager's and other's work on iterative, graph-based decoding techniques.

As first described by Berrou *et al.*, turbo error control consists of two key components: parallel concatenated encoding and iterative "turbo" decoding [44, 47]. A typical parallel concatenated encoder is shown in Fig. 5. It consists of two parallel convolutional encoders separated by an interleaver, with the input to the channel being the data bits $\mathbf{m}$ along with the parity bits $\mathbf{X}_1$ and $\mathbf{X}_2$ output from each of the encoders in response to input $\mathbf{m}$. The key to parallel concatenated encoding lies in the recursive nature of the encoders and the impact of the interleaver on the information stream. Interleavers also play a significant role in the elimination of error floors [47].

Iterative or turbo decoding exploits the component-code substructure of the turbo encoder by associating a component decoder with each of the component encoders. More specifically, each decoder performs soft input/soft output decoding, as shown in Fig. 6. In this figure decoder 1 generates a soft decision in the form of a probability measure $p(\mathbf{m}_1)$ on the transmitted data bits based on the received codeword $(\mathbf{m}',\mathbf{X}_1')$. This reliability information is passed to decoder 2, which generates its own probability measure $p(\mathbf{m}_2)$ from its received codeword $(\mathbf{m}',\mathbf{X}_2')$ and the probability measure $p(\mathbf{m}1)$. This reliability information is input to decoder 1, which revises its measure based on this information and the original received codeword. Decoder 1 sends the new reliability information to decoder 2, which revises its measure using this new information. Turbo decoding proceeds in an iterative manner, with the two component decoders alternately updating their probability measures. Ideally the decoders eventually agree on probability measures that reduce to hard decisions $\mathbf{m} = \mathbf{m}_1 = \mathbf{m}_2$. However, the stopping condition for turbo decoding is not well defined, in part because there are many cases in which the turbo decoding algorithm does not converge; that is, the decoders cannot agree on the value of $\mathbf{m}$. Several methods have been proposed for detecting convergence (if it occurs), including bit estimate variance [48] and neural net-based techniques [49].

Figure 7 indicates the performance of turbo codes. It can be seen that at a BER of 10^{-5} a recursive convolutional systematic (RCS) turbo code provides an 8.5 dB coding gain over uncoded systems and lies within 1.5 dB of the Shannon

■ Figure 6. *A turbo decoder.*

capacity limit. The intuitive explanation for this amazing performance is that the code complexity introduced by the encoding structure is similar to the codes that achieve Shannon capacity. The iterative procedure of the turbo decoder allows these codes to be decoded without excessive complexity [46]. Similar performance gains have been shown for bandwidth-constrained systems through turbo trellis coded modulation [50–52].

The soft decisions of the turbo decoding algorithm are highly dependent on the channel model. The performance shown in Fig. 7 is for a simple AWGN channel model. Turbo codes designed for AWGN channels do not necessarily perform well on wireless channels. However, recent work has shown that turbo codes work well on many wireless channels as long as an accurate channel model is embedded into the turbo decoding algorithm [53]. Recent results also indicate that for wireless systems with poor channel estimates, turbo codes may have inferior performance to simpler codes like trellis codes [54]. Note also that the turbo decoding algorithm typically requires more signal processing power than other decoders due to its iterative nature. This additional energy requirement may reduce the net energy gain of turbo codes.

MULTIPLE ANTENNAS

Multiple antennas at the transmitter and/or receiver play a powerful role in improving the performance and reducing the required transmit power for wireless link layer designs. Multiple antenna systems typically use either diversity, beamsteering, or multiple input multiple output (MIMO) techniques. Diversity combining is a common technique to mitigate flat fading by coherently combining multiple independently fading copies of the signal [55]. By significantly reducing the impact of flat fading, diversity combining can lead to significant power savings [56]. This is indicated in Fig. 8, where we show the performance of maximal-ratio combining (MRC) diversity in Rayleigh fading with one (no diversity), two, and four branches. We see that compared to no diversity, at a BER of 10^{-3} two-branch diversity saves 5.5 dB of power and four-branch diversity saves 8.5 dB.

Beamsteering creates an effective antenna pattern at the receiver with high gain in the direction of the desired signal and low gain in all other directions. Beamsteering is accomplished by combining arrays of antennas with signal processing in both space and time. The signal processing typically adjusts the phase shifts at each antenna to "steer" the beam in the desired direction. A simpler technique uses sectorized antennas with switching between the sectors. Beamsteering significantly improves energy efficiency since transmitter power is focused in the direction of its intended receiver. Beamsteering also reduces interference power along with fading and intersymbol interference due to multipath, since the interference and multipath signals are highly attenuated when they arrive from directions other than that of the line-of-sight (or dominant) signal. Results in [57] indicate that beamsteering can significantly improve the transmission range, data rates, and BER of wireless links. Highly mobile nodes can diminish these
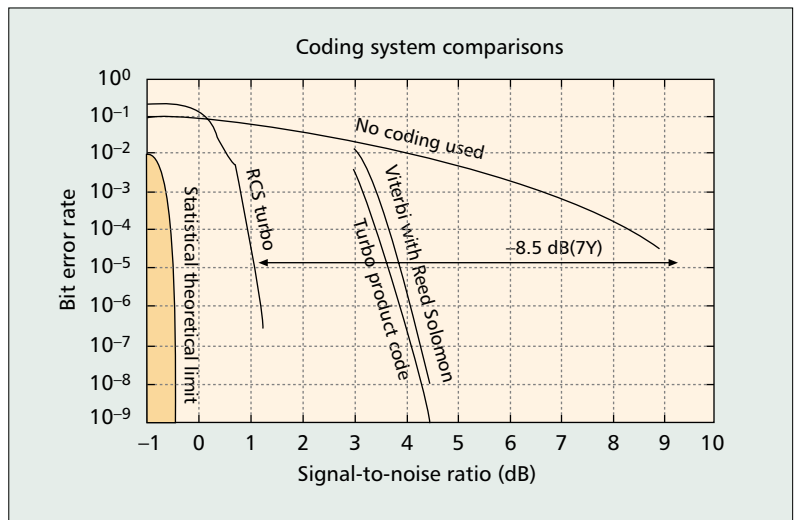
■ Figure 7. *Turbo code performance.*

gains, since the beamsteering direction will be shifting and difficult to determine accurately.

MIMO systems, where both transmitter and receiver use multiple antennas, can significantly increase the data rates possible on a given channel. In MIMO systems, if both the transmitter and receiver have channel estimates, with *N* antennas at the transmitter and receiver the MIMO system can be transformed into *N* separate channels that do not interfere with each other, providing a roughly *N*-fold capacity increase over a system with a single antenna at both the transmitter and receiver. When the transmitter does not know the channel, the optimal transmission strategy is a space-time code, where bits are encoded over both space and time [58]. These codes are highly complex, so in practice suboptimal schemes like layered space-time codes are used and tend to perform very well [59].
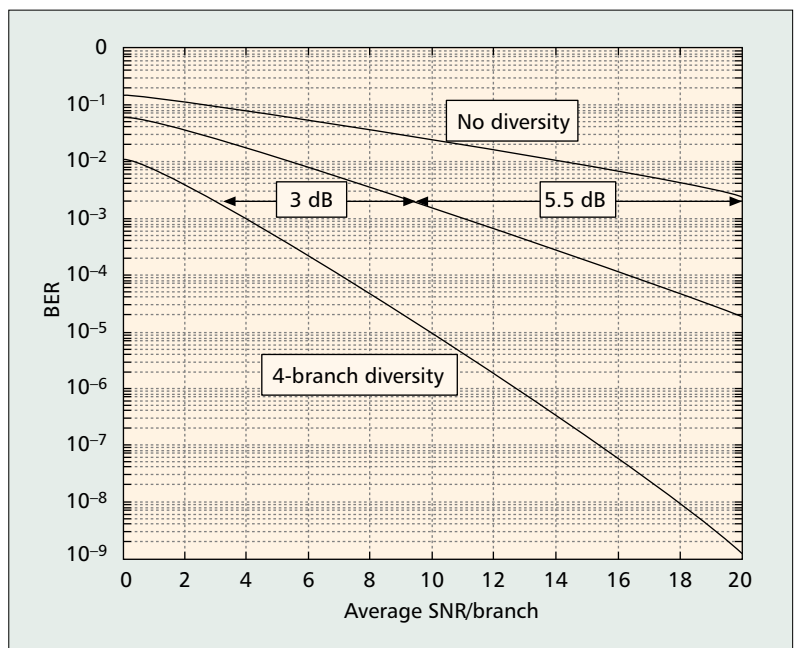
While multiple antenna techniques save trans-

■ Figure 8. *Performance of diversity systems.*

> At the link layer power control can be used to compensate for random channel variations due to multipath fading, reduce the transmit power required to obtain a given data rate and error probability, minimize the probability of link outage, and reduce interference to neighboring nodes.

mission power, they are often highly complex and therefore require significant power for signal processing. Given a total energy constraint this trade-off must be examined relative to each system to determine if multiple antenna techniques result in a net savings in energy.

POWER CONTROL

Power control is a potent mechanism for improving wireless network performance. At the link layer power control can be used to compensate for random channel variations due to multipath fading, reduce the transmit power required to obtain a given data rate and error probability, minimize the probability of link outage, and reduce interference to neighboring nodes. It can also be used to meet hard delay constraints and prevent buffer overflow.

Power control strategies at the link layer typically either maintain SINR on the link above a required threshold by increasing power relative to fading and interference or use a "water-filling" approach where power and rate are increased for good channel conditions, decreased for poor channel conditions, and set to zero when the channel quality falls below a given cutoff threshold [36]. The constant SINR strategy works well for continuous stream traffic with a delay constraint, where data is typically sent at a fixed rate regardless of channel conditions. However, this power control strategy is not power-efficient, since much power must be used to maintain the constant SINR in deep fading conditions. Optimal variation of transmission rate and power maximizes average throughput [60] and channel capacity [61], but the associated variable-rate transmission and channel-dependent delay may not be acceptable for some applications. Power control has also been used to meet delay constraints for wireless data links. In this approach power for transmission of a packet increases as the packet approaches its delay constraint, thereby increasing the probability of successful transmission [62]. A more complex approach uses dynamic programming to minimize the transmit power required to meet a hard delay constraint [63], and the resulting power consumption is much improved over power control that maintains a constant SINR.

Before closing this section, we want to emphasize that power control has a significant impact on protocols above the link layer. The level of transmitter power defines the "local neighborhood" — the collection of nodes that can be reached in a single hop — and thus in turn defines the context in which access, routing, and other higher-layer protocols operate. Power control will therefore play a key role in the development of efficient cross-layer networking protocols. We will discuss integration of power control with multiple access and routing protocols in later sections.

ADAPTIVE RESOURCE ALLOCATION

Adaptive resource allocation in link layer design is a relatively new technique that provides robust link performance with high throughput while meeting application-specific constraints. The basic premise is to adapt the link transmission scheme to the underlying channel, interference, and data characteristics through variation of the transmitted power level, symbol transmission rate, constellation size, coding rate/scheme, or any combination of these parameters. Recent work shows that adaptation can increase the link spectral efficiency by a factor of five or more over nonadaptive techniques while maintaining the required link performance [10]. Moreover, adaptive modulation can compensate for SINR variations due to interference as well as multipath fading and can be used to meet different QoS requirements of multimedia [64] by prioritizing delay-constrained bits and adjusting transmit power to meet BER requirements.

Recent work in adaptive resource allocation has investigated combinations of power, rate, code, and BER adaptation ([60] and the references therein). These schemes typically assume some finite number of power levels, modulation schemes, and codes, and the optimal combination is chosen based on system conditions and constraints. Only a small number of power levels, rates, and/or codes are needed to achieve near-optimal performance, since there is a critical number of degrees of freedom needed for good performance of adaptive resource allocation, and beyond this critical number additional degrees of freedom provide minimal performance gain [60]. In particular, power control in addition to variable-rate transmission provides negligible capacity increase in fading channels [61], cellular systems [65, 66], and ad hoc wireless networks [4]. Code-division multiple access (CDMA) systems, in addition to varying power, data rate, and channel coding, can also adjust their spreading gain or the number of spreading codes assigned to a given user [67, 68]. The benefits of assigning multiple spreading codes per user are greatest when some form of multiuser detection is used, since otherwise self-interference is introduced [69]. Note also that in adaptive CDMA systems all transmitters sending to a given receiver must coordinate since they interfere with each other.

Other adaptive techniques include variation of the link layer retransmission strategy as well as its frame size. The frame is the basic information block transmitted over the link and includes overhead in the form of header and error control bits. Shorter frames entail higher overhead, but are less likely to be corrupted by sporadic interference and require less time for retransmission. Recent results have shown that optimizing frame length can significantly improve throughput as well as energy efficiency [70].

Data communications require corrupted packets to be retransmitted so that all bits are correctly received. Current protocols typically discard the corrupted packet and start over again on the retransmission. However, recent work has shown that diversity combining of retransmitted packets or retransmitting additional redundant code bits instead of the entire packet can substantially increase throughput [71, references therein]. A performance comparison of incremental redundancy against that of adaptive modulation is given in [72].

MEDIUM ACCESS CONTROL DESIGN ISSUES

The medium access control protocol dictates

how different users share the available spectrum. There are two components to this spectrum allocation: how to divide the spectrum into different channels, and then how to assign these different channels to different users. In the first subsection below we discuss channelization: the different methods that can be used to divide the spectrum into different channels. The next two sections deal with channel assignment. When users have very bursty traffic the most efficient mechanism to assign channels is random access, where users contend for a channel whenever they have data to transmit. This contention is inefficient when users have continuous stream data or long packet bursts. In this case some form of scheduling helps to prevent collisions and ensure continuous connections. We conclude this section with a discussion of the role power control plays in multiple access strategies.

CHANNELIZATION

Methods to divide the spectrum into different channels include frequency division, time division, code division, and hybrid combinations of these methods. In code division either orthogonal or semi-orthogonal spreading code techniques may be used, although orthogonal codes often do not maintain orthogonality after transmission through a wireless channel. Time division, frequency division, and orthogonal code division are equivalent in that they all divide up the spectrum orthogonally, and this orthogonal division results in the same number of channels [73]. Without frequency reuse, code division with semi-orthogonal codes has inferior capacity to that of orthogonal techniques due to its introduction of interference between users. However, multiuser detection can increase the capacity of semi-orthogonal code division above that of the orthogonal channelization methods. These capacity conclusions are based on idealized implementations. In practice, due to implementation and signal propagation issues, these capacity conclusions are not necessarily accurate. Frequency reuse and bursty traffic make the capacity trade-offs between the different channelization methods even more difficult to determine. We now describe each of these division techniques in more detail.

In frequency division the system bandwidth is divided into nonoverlapping channels, and each active user is assigned a different channel. If the channel bandwidth does not exceed the inverse of the multipath delay spread in the channel, frequency-division systems do not require intersymbol interference compensation (e.g., equalization or multicarrier modulation) [55, 74]. Frequency division is generally the simplest channelization technique to implement, but it is rather inflexible. In particular, it is difficult to allocate multiple channels on demand to a single user, since this requires simultaneous demodulation of multiple channels in different frequency bands.

Time division is an alternate technique, where time is divided into orthogonal time slots. Active users with continuous stream data are typically assigned a cyclically repeating timeslot. These active users do not transmit continuously, so functions such as channel probing can be done during idle times. One difficulty of using time division is the need for synchronization among all nodes transmitting to the same receiver. Since these different nodes will have different propagation delays to the receiver, the timeslots must be synchronized so that they remain orthogonal after these respective delays. The channels associated with a time-division system are typically wideband relative to multipath delay spread, and thus require intersymbol interference mitigation, often in the form of an equalizer. Time division has more flexibility for dynamic channel allocation than frequency division, since multiple time slots can be assigned to a user on demand with no complexity increase in transmission or reception.

In code division, time and bandwidth are used simultaneously by different users, modulated by orthogonal or semi-orthogonal spreading codes. The receiver then uses the spreading code properties to separate out the different users. One of the big advantages of channelization using spread spectrum is that little dynamic coordination of users in time or frequency is required, since the users can be separated by the code properties alone. In addition, since time and frequency division carve up time and bandwidth into orthogonal pieces, there is a hard limit on how many users can simultaneously occupy the system. This is also true for code division using orthogonal codes. In contrast, if semi-orthogonal codes are used, the number of users is *interference-limited*. Specifically, there is no hard limit on how many users can simultaneously share the channel. However, since semi-orthogonal codes interfere with each other, the more users that are packed into the same channel, the higher the level of interference, which degrades the performance of all the users. Moreover, semiorthogonal code-division systems typically require power control to compensate for the near-far problem. The near-far problem arises because users modulating their signals with different spreading codes interfere with each other. Suppose that one user is very close to the receiver, and another user is very far away. If both users transmit at the same power level, the interference from the close user will swamp the signal from the far user. Thus, power control is used such that the received signal powers from all users are roughly the same. This form of power control, which essentially inverts any attenuation and/or fading introduced by the channel, causes each interferer to contribute an equal amount of power, thereby eliminating the near-far problem.

Performance of code-division systems can be significantly improved using both RAKE receivers and multiuser detection. A RAKE receiver takes advantage of the structure of the spreading codes to coherently combine the multipath components of the received signal [55]. Multiuser detection schemes, pioneered by Verdu, take advantage of the fact that the interference from other users arises from a known semi-orthogonal spreading code sequence. Thus, this interference can be decoded and subtracted out, thereby eliminating most interference between users and the need for power control [75]. While multiuser detection promises significant performance enhancement, the detection scheme must have a low probability of bit error, since bits that are incorrectly detected are sub-

The benefits of assigning multiple spreading codes per user are greatest when some form of multiuser detection is used, since otherwise self-interference is introduced. Note also that in adaptive CDMA systems all transmitters sending to a given receiver must coordinate since they interfere with each other.
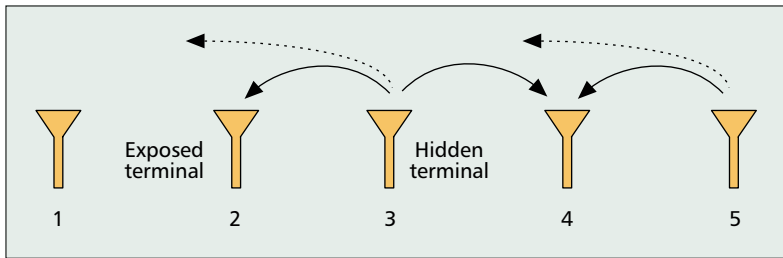
■ Figure 9. *Hidden and exposed terminals in random access.*

tracted from the signals of other users, which may cause them to be decoded in error as well (error propagation). These detection strategies can also be highly sensitive to channel estimation errors. Finally, multiuser detection significantly increases the complexity (and associated energy requirements) of the receivers. For these reasons multiuser detection has not yet permeated existing code-division system designs. Even without multiuser detection, code-division systems have the most complex receivers of all the channelization techniques due to spreading code acquisition and synchronization requirements along with power control. Note also that multiuser detection requires knowledge at the receiver of the spreading codes associated with each signal it receives, which entails some control overhead in an ad hoc wireless network.

RANDOM ACCESS

Given a channelization scheme, each user can be assigned a different channel for some period of time. However, most data users do not require continuous transmission, so dedicated channel assignment can be extremely inefficient. Moreover, most systems have many more total users (active plus idle users) than channels, so at any given time channels can only be allocated to users that need them. Random access strategies are used in such systems to assign channels to active users. Random access techniques were pioneered by Abramson with the Aloha protocol [76]. In Aloha users are assumed to share a set of common channels. In unslotted or "pure" Aloha, the users transmit on one of these channels whenever they have data to send. Should two users "collide," they both wait a random amount of time before retransmitting. The goal, of course, is to prevent the users from colliding once again when they retransmit. In slotted Aloha, the users are further constrained by a requirement that they only begin transmitting at the start of a time slot. The use of such time slots increases the maximum possible throughput of the channel [77], but also introduces the need for a synchronization mechanism of some sort. Even in a slotted system, collisions occur whenever two or more users attempt transmission in the same slot. Error control coding can result in correct detection of a packet even after a collision, but if the error correction is insufficient, the packet must be retransmitted, resulting in a complete waste of the energy consumed in the original transmission. A study on design optimization between error correction and retransmission is described in [78].

Collisions can be reduced by carrier sense

multiple access (CSMA), where users sense the channel and delay transmission if they detect that another user is currently transmitting [77]. CSMA only works when all users can hear each other's transmissions, which is typically not the case in wireless systems due to the nature of wireless propagation. This gives rise to the hidden terminal problem, illustrated in Fig. 9, where each node can hear its immediate neighbor but no other nodes in the network. In this figure both nodes 3 and 5 wish to transmit to node 4. Suppose node 5 starts its transmission. Since node 3 is too far away to detect this transmission, it assumes that the channel is idle and begins its transmission, thereby causing a collision with node 5's transmission. Node 3 is said to be hidden from node 5 since it cannot detect node 5's transmission. Aloha with CSMA also creates inefficiencies in channel utilization from the exposed terminal problem, also illustrated in Fig. 9. Suppose the exposed terminal in this figure, node 2, wishes to send a packet to node 1 at the same time node 3 is sending to node 4. When node 2 senses the channel it will detect node 3's transmission and assume the channel is busy, even though node 3 does not interfere with the reception of node 2's transmission by node 1. Thus, node 2 will not transmit to node 1 even though no collision would have occurred.

The collisions introduced by hidden terminals and inefficiencies introduced by exposed terminals are often addressed by a four-way handshake prior to transmission, as in the 802.11 WLAN protocol [79, 80]. However, this handshake protocol is based on single-hop routing, and recent work indicates that its performance in multihop networks may be suboptimal [81, 82]. Another technique to avoid hidden and exposed terminals is busy tone transmission. In this strategy users first check to see whether the transmit channel is busy by listening for a busy tone on a separate control channel [77]. This is typically not an actual busy tone; instead, a bit is set in a predetermined field on the control channel. This scheme works well in preventing collisions when a centralized controller can be "heard" by users throughout the network. In a flat network without centralized control, more complicated measures are used to ensure that any potential interferer on the first channel can hear the busy tone on the second [83]. Hybrid techniques using handshakes, busy tone transmission, and power control are investigated in [84]. Note that while the four-way handshake and busy tone transmission both reduce collisions due to the hidden terminal problem, they tend to aggravate the exposed terminal problem, leading to less efficient utilization of the available channels in the network. A solution to this problem is to have both transmitter and receiver send busy tones [83].

Random access protocols can be more energy-efficient by limiting the amount of time a given node spends transmitting and receiving. The paging industry developed a solution to this problem several decades ago by scheduling "sleep" periods for pagers. The basic idea is that each pager need only listen for transmissions during certain short periods of time. This is a simple solution to implement when a central

controller is available. It is less obvious how to implement such strategies within the framework of a distributed control algorithm. Access protocols that utilize node sleep times to minimize energy consumption are investigated in [13].

Random access schemes can be made more flexible in general, and more energy aware in particular, by adopting a dynamic programming approach to decisions about transmissions. Under dynamic programming, decision making is based on utility (cost) functions: an agent will act or not, depending on utility of the action as indicated by a utility function computed over some time period. A given protocol can be made energy aware by introducing the cost of a transmission into the utility function.

Consider the case of Aloha. In work conducted by MacKenzie at Cornell, a game-theoretic version of Aloha was developed that initially focused on a simple "collision game" [85]. It is assumed in this game that players know the number of backlogged users, n. $G(n)$ is then the game in which there are currently n users backlogged. In each stage of $G(n)$ each of the n backlogged players must decide whether to transmit or wait. If one player decides to transmit and the rest decide to wait, the player who transmits will receive a payoff of $1 - c$, where c is the cost of a transmission, and each of the other $(n - 1)$ players will then proceed to play $G(n - 1)$ in the next period. If either no users transmit or more than one user transmits, all players will play $G(n)$ again in the next period. Players place a lower value on payoffs in later stages than on current payoffs via a discount factor. The cost of delay and the energy cost of transmission are parameters of the cost function that can be tailored to the particular application. The resulting system is both stable (in the language of game theory, there is a Nash equilibrium) and distributed. It allows for individual nodes to make autonomous decisions on retransmission strategies. This simple version of the game assumes that the users know the number of backlogged users within the local neighborhood, but it is possible to develop utility functions that reflect less ideal situations. In general, the decision-theoretic approach provides a convenient way to embed the cost of transmission decisions into random access protocols.

SCHEDULING

Random access protocols work well with bursty traffic where there are many more users than available channels, yet these users rarely transmit. If users have long strings of packets or continuous stream data, then random access works poorly as most transmissions result in collisions. Thus channels must be assigned to users in a more systematic fashion by transmission scheduling. In scheduled access the available bandwidth is channelized into multiple time, frequency, or code division channels. Each node schedules its transmission on different channels in such a way as to avoid conflicts with neighboring nodes while making the most efficient use of the available time and frequency resources. While there has been much work on transmission scheduling, or channel assignment, in cellular systems [86], the centralized control in these systems greatly simplifies the problem. Distributed scheduled

access in ad hoc wireless networks in general is an NP-hard problem [87]. Selman *et al.* have recently discovered that NP-hard problems exhibit a rapid change in complexity as the size of the problem grows [88, 89]. The identification of this "phase transition" provides an opportunity for bounding the complexity of problems like scheduled access by staying on the good side of the phase transition.

Even with a scheduling access protocol, some form of Aloha will still be needed since a predefined mechanism for scheduling will be, by definition, unavailable at startup. Aloha provides a means for initial contact and the establishment of some form of scheduled access for the transmission of relatively large amounts of data. A systematic approach to this initialization that also combines the benefits of random access for bursty data with scheduling for continuous data is packet reservation multiple access (PRMA) [90]. PRMA assumes a slotted system with both continuous and bursty users (e.g., voice and data users). Multiple users vie for a given time slot under a random access strategy. A successful transmission by one user in a given time slot reserves that time slot for all subsequent transmissions by the same user. If the user has a continuous or long transmission, after successfully capturing the channel he/she has a dedicated channel for the remainder of his/her transmission (assuming subsequent transmissions are not corrupted by the channel: this corruption causes users to lose their slots, and they must then recontend for an unreserved slot, which can entail significant delay). When this user has no more packets to transmit, the slot is returned to the pool of available slots that users attempt to capture via random access. Thus, data users with short transmissions benefit from the random access protocol assigned to unused slots, and users with continuous transmissions get scheduled periodic transmissions after successfully capturing an initial slot. A similar technique using a combined reservation and Aloha policy is described in [14].

Scheduling under an energy constraint further complicates the problem. As discussed in an earlier section, channel capacity under a finite energy constraint is maximized by transmitting each bit over a very long period of time. However, when multiple users wish to access the channel, the transmission time allocated to each user must be limited. Recent work has investigated optimal scheduling algorithms to minimize transmit energy for multiple users sharing a channel [91]. In this work scheduling was optimized to minimize the transmission energy required by each user subject to a deadline or delay constraint. The energy minimization was based on judiciously varying packet transmission time (and corresponding energy consumption) to meet the delay constraints of the data. This scheme was shown to be significantly more energy efficient than a deterministic schedule with the same deadline constraint.

POWER CONTROL

Access protocols can be made more efficient and distributed by taking advantage of power control. Work in this area has mainly focused on maintaining the SINR of each user sharing the channel above a given threshold, which may be

Even with a scheduling access protocol, some form of Aloha will still be needed since a predefined mechanism for scheduling will be, by definition, unavailable at startup. Aloha provides a means for initial contact and the establishment of some form of scheduled access for the transmission of relatively large amounts of data.
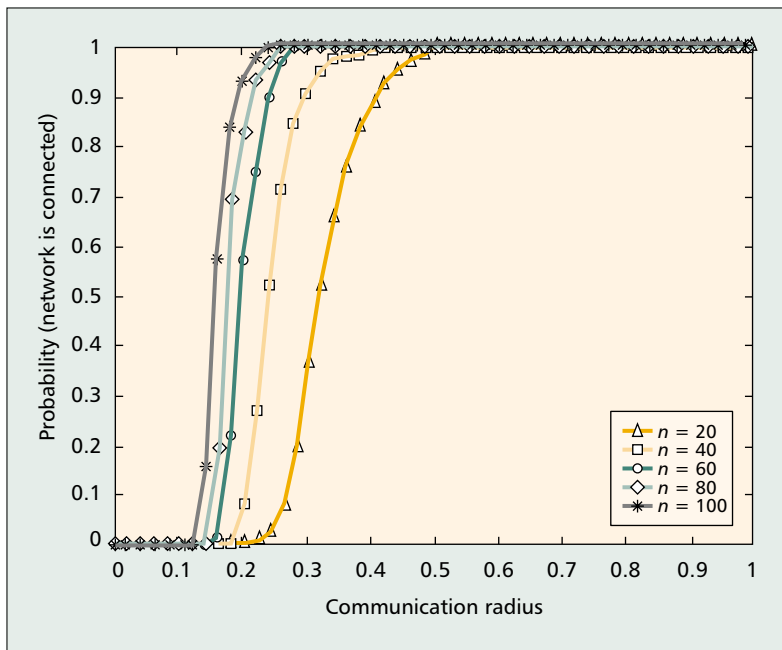
■ **Figure 10.** *Network connectivity vs. node density.*

different for different users [11]. Necessary and sufficient conditions to ensure that a feasible set of transmit powers for all users exists under which these users can meet their threshold SINR levels given the link gains between them are determined in [11]. Battery power for each user is minimized by finding the minimum power vector within the feasible set. This algorithm can also be performed in a distributed manner, which eliminates the need for centralized power control. Access to the system can be based on whether the new user causes other users to fall below their SINR targets. Specifically, when a new user requests access to the system, a centralized controller can determine if a set of transmit powers exists such that he/she can be admitted without degrading existing users below their desired SINR threshold. This admission can also be done using the distributed algorithm, where the new user gradually ramps up his/her power, which causes interference to other existing users in the system. If the new user can be accommodated in the system without violating the SINR requirements of existing users, then the power control algorithms of the new and existing users eventually converge to the feasible power vector under which all users (new and existing) meet their SINR targets. If the new user cannot be accommodated, as he/she ramps up his/her power the other users will increase their powers to maintain their SINRs such that the new user remains far from his/her SINR target. After some number of iterations without reaching his/her target, the new user will either back off from the channel and try again later or adjust his/her SINR target to a lower value and try again.

A power control strategy for multiple access that takes into account delay constraints is proposed and analyzed in [62]. This strategy optimizes the transmit power relative to both channel conditions and the delay constraint via dynamic programming. The optimal strategy

exhibits three modes: very low-power transmission when the channel is poor and the tolerable delay large, higher power when the channel and delay are average, and very high-power transmission when the delay constraint is tight. This strategy exhibits significant power savings over constant power transmission while meeting the delay constraints of the traffic.

Power control has also been extensively studied for cellular systems [92, references therein]. In cellular systems the power adaptation of each mobile is typically controlled by a centralized base station in each cell, and the base stations coordinate to optimize power allocation across all mobiles in the network. Distributed power control for cellular systems has also been investigated [93]. In general, power control in cellular systems, often in combination with channel or rate allocation, improves spatial channel reuse by managing interference, thereby significantly increasing cellular network capacity. The benefits of power control for both voice and data in cellular systems has been well established, and most next-generation cellular system designs include some form of power control. However, there are few results outside of [11] on the design and performance of power control schemes in ad hoc wireless networks, and this remains an active area of research.

NETWORK DESIGN ISSUES

NEIGHBOR DISCOVERY AND NETWORK CONNECTIVITY

"Neighbor discovery" is one of the first steps in the initialization of a network of randomly distributed nodes. From the perspective of the individual node, this is the process of determining the number and identity of network nodes with which direct communication can be established given some maximum power level and minimum link performance requirements (typically in terms of data rate and associated BER). Clearly the higher the allowed transmit power, the greater the number of nodes in a given neighborhood.

Neighbor discovery begins with a probe of neighboring nodes using an initial power constraint. If the number of nodes thus contacted is insufficient to ensure some minimal connectivity requirements then the power constraint is relaxed and probing repeated. The minimal connectivity requirements will depend on the application, but most ad hoc wireless network applications assume a fully-connected network whereby each node can reach every other node, often through multiple hops. The exact number of neighbors that each node requires to obtain a fully-connected network depends on the exact network configuration but is generally on the order of six to eight for randomly distributed immobile nodes [77, 94]. An analysis of the minimum transmit power required at each node to maintain full connectivity is done in [95, 96]. Clearly the ability of the network to stay connected will decrease with node mobility, so maintaining full connectivity under high mobility will require larger neighborhoods and an associated increase in transmit power at each node.

It is interesting to note that, given a random
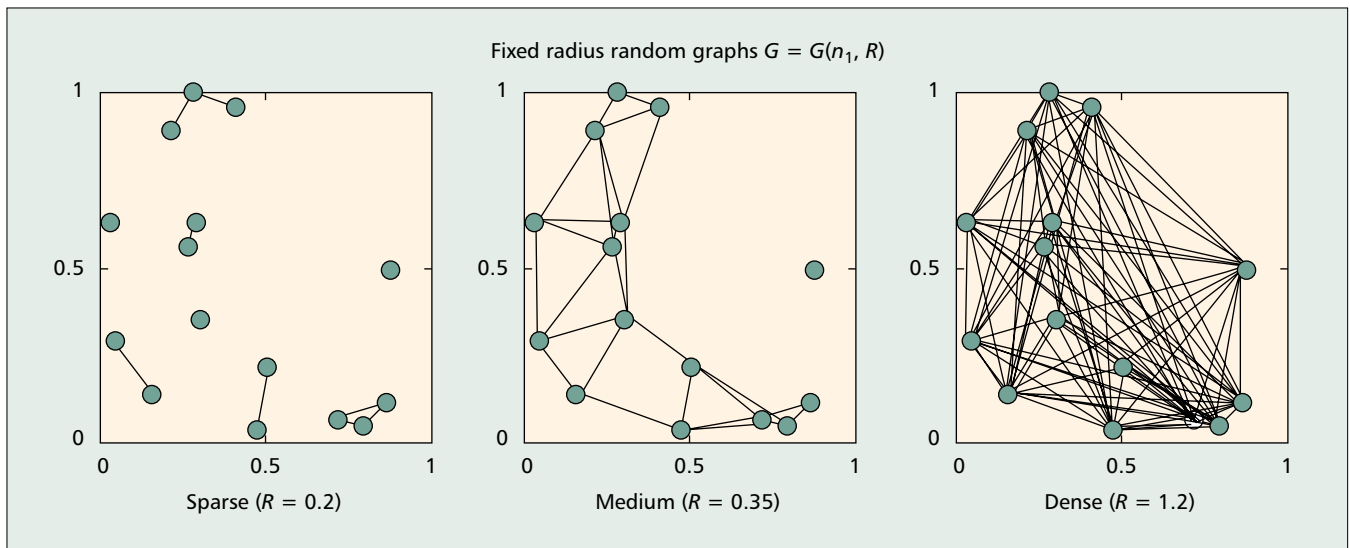
Figure 11. *Connectivity increases with transmission radius* R *(scales with power).*

distribution of nodes, the likelihood of complete connectivity changes abruptly from zero to one as the transmission range of each node is increased. The sensitivity of the location of this "zero-one" transition to the density of nodes is shown in Fig. 10 from [97] for a static network. This figure shows several curves, each corresponding to a different number (n) of randomly distributed nodes in a region of fixed size. The curves thus reflect, from left to right, a decreasing level of node density. Note that the transmission range required for the network to be fully connected increases as the node density decreases, reflecting the increased probability of deep holes, to borrow a term from the theory of lattices.

Connectivity is heavily influenced by the ability to adapt various parameters at the link layer such as rate, power, and coding, since communication is possible even on links with low SINR if these parameters are adapted [19]. The sensitivity of connectivity to transmit power level can be represented abstractly through the theory of transmission graphs [97]. Figure 11 shows a series of transmission graphs: graphs in which edges exist between pairs of nodes if and only if the distance between the nodes is less than or equal to a given bound. The series of graphs shows how connectivity can increase dramatically with transmit power given a random distribution of nodes.

From the standpoint of power efficiency and operational lifetime, it is also very important that nodes be able to decide whether or not to take a nap. These sleep decisions must take into account network connectivity, so it follows that these decisions are local but not autonomous. Mechanisms that support such decisions can be based on neighbor discovery coupled with some means of ordering decisions within the neighborhood. In a given area, the opportunity to sleep should be circulated among the nodes, ensuring that connectivity is not lost through the coincidence of several identical decisions to go to sleep.

ROUTING

The multihop routing protocol in an ad hoc wireless network is a significant design challenge, especially under energy constraints where the exchange of routing data consumes precious energy resources. Most work in multihop routing protocols falls into three main categories: flooding, proactive routing (centralized or distributed), and reactive routing [98–100, references therein].

In flooding a packet is broadcast to all nodes within receiving range. These nodes also broadcast the packet, and the forwarding continues until the packet reaches its ultimate destination. Flooding has the advantage that it is highly robust to changing network topologies and requires little routing overhead. In fact, in highly mobile networks flooding may be the only feasible routing strategy. The obvious disadvantage is that multiple copies of the same packet traverse through the network, wasting bandwidth and battery power of the transmitting nodes. This disadvantage makes flooding impractical for all but the smallest of networks.

The opposite philosophy to flooding is centralized route computation. In this approach information about channel conditions and network topology are determined by each node and forwarded to a centralized location that computes the routing tables for all nodes in the network. The criterion used to compute the "optimal" route depends on the optimization criterion: common criteria include minimum average delay, minimum number of hops, and recently, minimum network congestion [101]. While centralized route computation provides the most efficient routing according to the optimality condition, it cannot adapt to fast changes in the channel conditions or network topology, and also requires much overhead for collecting local node information and then disseminating the routing information. Centralized route computation, like flooding, is typically only used in very small networks.

Distributed route computation is the most common routing procedure used in ad hoc wireless networks. In this protocol nodes send their connectivity information to neighboring nodes, and then routes are computed from this local information. In particular, nodes determine the

Both centralized and distributed routing require fixed routing tables that must be updated at regular intervals. An alternate approach is reactive (on-demand) routing, where routes are created only at the initiation of a source node that has traffic to send to a given destination. In this strategy a source node initiates a route-discovery process when it has data to send.

next hop in the route of a packet based on this local information. There are several advantages of distributed route computation. First, the overhead of exchanging routing information with local nodes is minimal. In addition, this strategy adapts quickly to link and connectivity changes. The disadvantages of this strategy are that global routes based on local information are typically suboptimal, and routing loops are often common in the distributed route computation.

Both centralized and distributed routing require fixed routing tables that must be updated at regular intervals. An alternate approach is reactive (on-demand) routing, where routes are created only at the initiation of a source node that has traffic to send to a given destination. This eliminates the overhead of maintaining routing tables for routes not currently in use. In this strategy a source node initiates a route discovery process when it has data to send. This process will determine if one or more routes are available to the destination. The route or routes are maintained until the source has no more data for that particular destination. The advantage of reactive routing is that globally efficient routes can be obtained with relatively little overhead, since these routes need not be maintained at all times. The disadvantage is that reactive routing can entail significant delay, since the route discovery process is initiated when there is data to send, but this data cannot be transmitted until the route discovery process has concluded. Recently a combination of reactive and proactive routing has been proposed to reduce the delay associated with reactive routing as well as the overhead associated with proactive routing [99].

Mobility has a huge impact on routing protocols as it can cause established routes to no longer exist. High mobility especially degrades the performance of proactive routing, since routing tables quickly become outdated, requiring an enormous amount of overhead to keep them up to date. Flooding is effective in maintaining routes under high mobility, but has a huge price in terms of network efficiency. A modification of flooding called multipath routing has been recently proposed, whereby a packet is duplicated on only a few paths with a high likelihood of reaching its final destination [100]. This technique has been shown to perform well under dynamically changing topologies.

Energy constraints in the routing protocol significantly change the problem. First of all, the exchange of routing information between nodes entails an energy cost: this cost must be traded against the energy savings that result from using this information to make routes more efficient. In addition, even with perfect information about the links and network topology, the route computation must change to take energy constraints into account. Specifically, a route utilizing a small number of hops (low delay) may use significantly more energy (per node and/or total energy) than a route consisting of a larger number of hops. Moreover, if one node is often used for forwarding packets the battery of that node will die out quickly, making that node unavailable for transmitting its own data or forwarding packets for others. Thus, the routing protocol under energy constraints must somehow balance delay constraints, battery lifetime, and routing efficiency.

There has been much recent work on evaluating routing protocols under energy constraints. In [102] simulations were used to compare the energy consumption of different well-known routing protocols. Their results indicate that reactive routing is more energy efficient. This is not surprising since proactive routing must maintain routing tables via continuous exchange of routing information, which entails a significant energy cost. This work was extended in [103] to more accurately model the energy consumption of radios in a "listening" mode. The energy consumption for this mode, ignored in [101], was significant; based on this more accurate model, it was concluded that the proactive and reactive routing schemes analyzed in [102] have roughly the same energy consumption. The article goes on to propose a sleep mode for nodes that reduces energy consumption by up to 40 percent. Other work in this area combines routing, power control, and adaptive coding to minimize the energy cost of routes [104]. Power control to optimize energy efficiency in routing is also studied in [105].

SCALABILITY AND DISTRIBUTED PROTOCOLS

Scalability arises naturally in the design of self-configuring ad hoc wireless networks. The key to self-configuration lies in the use of distributed network control algorithms — algorithms that adjust local performance to account for local conditions. To the extent that these algorithms forgo the use of centralized information and control resources, the resulting network will be scalable. Work on scalability in ad hoc wireless networks has mainly focused on self-organization [13, 106], distributed routing [107], mobility management [7], QoS support, and security [108]. Note that distributed protocols often consume a fair amount of energy in local processing and message exchange: this is analyzed in detail for security protocols in [109]. Thus, interesting trade-offs arise as to how much local processing should be done vs. transmitting information to a centralized location for processing. Most work on scalability in ad hoc wireless networks has focused on relatively small networks, less than 100 nodes. Many of the applications described earlier, especially sensor networks, could have hundreds to thousands of nodes or even more. The ability of existing network protocols to scale to such large network sizes remains an open question.

NETWORK CAPACITY

The fundamental capacity limit of an ad hoc wireless network — the set of maximum data rates possible between all nodes — is a highly challenging problem in information theory. In fact, the capacity for simple channel configurations within an ad hoc wireless network, such as the general relay and interference channel, remain unsolved [110]. In a recent landmark paper an upper bound on the performance of an asymptotically large ad hoc wireless network in terms of the uniformly achievable maximum data rate was determined [12]. Surprisingly, this result indicates that even with optimal routing and scheduling, the per-node rate in a large ad hoc wireless network goes to zero. To a large extent this pessimistic result indicates that in a large network all nodes should not communicate with

all other nodes: there should be distributed processing of information within local neighborhoods. This work was extended in [111] to show that node mobility actually increases the per-node rate to a constant (i.e., mobility increases network capacity). This result follows from the fact that mobility introduces variation in the network that can be exploited to improve per-user rates. Other recent work in this area has determined achievable rate regions for ad hoc wireless networks using adaptive transmission strategies [4] and an information theoretic analysis on achievable rates between nodes [112].

APPLICATION DESIGN ISSUES

In true cross-layer protocol design, the highest layer — the application — can play a significant role in network efficiency. In this section we consider network adaptation to the application requirements and application adaptation to the underlying network capabilities.

ADAPTIVE QoS

The Internet today, even with high-speed high-quality fixed communication links, is unable to deliver guaranteed QoS to applications in terms of guaranteed end-to-end rates or delays. For ad hoc wireless networks, with low-capacity error-prone time-varying links, mobile users, and a dynamic topology, the notion of being able to guarantee these forms of QoS is simply unrealistic. Therefore, ad hoc wireless network applications must adapt to time-varying QoS parameters offered by the network. While adaptivity at the link and network levels as described earlier will provide the best possible QoS to the application, this QoS will vary with time as channel conditions, network topology, and user demands change. Applications must therefore adapt to the QoS offered. There can also be a negotiation for QoS such that users with higher priority can obtain better QoS by lowering the QoS of less important users.

As a simple example, the network may offer the application a rate-delay trade-off curve derived from the capabilities of the lower-layer protocols [30]. The application layer must then decide at which point on this curve to operate. Some applications may be able to tolerate a higher delay but not a lower overall rate. Examples include data applications in which the overall data rate must be high but latency might be tolerable. Other applications might be extremely sensitive to delay (e.g., a distributed control application) but might be able to tolerate a lower rate (e.g., via a coarser quantization of sensor data). Energy constraints introduce another set of trade-offs related to network performance versus longevity. Thus, these trade-off curves will typically be multidimensional to incorporate rate, delay, BER, longevity, and so on. These trade-off curves will also change with time as the number of users on the network and the network environment change.

APPLICATION ADAPTATION AND CROSS LAYER DESIGN REVISITED

In addition to adaptive QoS, the application itself can adapt to the QoS offered. For example, for applications like video with a hard delay constraint, the video compression algorithm might change its compression rate such that the source rate adjusts to the rate the network can deliver under the delay constraint. Thus, under poor network conditions compression would be higher (lower transmission rate) and the end quality poorer. There has been much recent work on application adaptation for wireless networks [113–115, references therein]. This work indicates that even demanding applications like video can deliver good overall performance under poor network conditions if the application is given the flexibility to adapt.

The concept of application adaptation returns us to the cross-layer design issue discussed earlier. While the application can adapt to a rate-delay-performance trade-off curve offered by the network and underlying links, by making the lower-layer protocols aware of the trade-offs inherent to the application adaptation, that trade-off curve might be adjusted to improve end-to-end performance without using up more resources in the network. In other words, if the application is aware of the lower-layer protocol trade-offs and these protocols are aware of the application trade-offs, these trade-off curves can be merged to operate at the best point relative to end-to-end performance. While implementing this philosophy remains a wide open research problem, it holds significant promise for the performance of ad hoc wireless networks.

SUMMARY AND CONCLUSIONS

We have described recent advances in ad hoc wireless network protocols and the strong interaction that occurs across different protocol layers. While there is still work to be done on improving link, medium access, network, and application protocols, the interactions across these different protocol layers invite novel cross-layer designs that exploit these interdependencies. Cross-layer design is particularly important under energy constraints, since energy across the entire protocol stack must be minimized.

The biggest design challenges in ad hoc wireless networks are the lack of centralized control, limited node capability, and variability of the links and network topology. Methods to address these challenges include intelligent distributed control protocols, node redundancy and coordination, and adaptivity at each layer of the protocol stack to compensate for and exploit variability. Recent breakthroughs in ad hoc wireless network design have come from thinking outside the box: the box of layered protocol designs, the box of wireline protocols, the box of guaranteed QoS for demanding applications, and the box of application-agnostic networks. Much more out-of-box thinking is required to deliver the great promise offered by ad hoc wireless networks.

ACKNOWLEDGMENTS

> The Internet today, even with high-speed high-quality fixed communication links, is unable to deliver guaranteed QoS to an application in terms of guaranteed end-to-end rates or delays. For ad hoc wireless networks, with low-capacity error-prone time-varying links, mobile users, and a dynamic topology, the notion of being able to guarantee these forms of QoS is simply unrealistic.

Energy constraints introduce another set of trade-offs related to network performance versus longevity. Thus, these trade-off curves will typically be multidimensional to incorporate rate, delay, BER, longevity, and so on. These trade-off curves will also change with time as the number of users on the network and the network environment change.

REFERENCES

[1] F. A. Tobagi, "Modeling and Performance Analysis of Multihop Packet Radio Networks," *Proc. EEE*, Jan. 1987, pp. 135–55.

[2] M. B. Pursley, "The Role of Spread Spectrum in Packet Radio Networks," *IEEE Proc.*, Jan. 1987.

[3] L. Kleinrock and J. Silvester, "Spatial Reuse in Multihop Packet Radio Networks," *IEEE Proc.*, Jan. 1987.

[4] S. Toumpis and A. Goldsmith, "Capacity Regions for Ad Hoc Networks," to appear, *ICC 2002*, Apr. 2002, also submitted to *IEEE Trans. WL Commun.*

[5] Y.-D. Lar and Y.-C. Hse, "Multihop Cellular: A New Architecture for Wireless Communications," *IEEE INFO-COM 2000*, June 2000, pp. 1273–82.

[6] A.N. Zadeh and B. Jabbari, "Performance Analysis of Multihop Packet CDMA Cellular Networks," *IEEE GLOBE-COM 2001*, Nov. 2001, pp. 2875–79.

[7] S. Basagni, D. Turgut, and S.K. Das, "Mobility-adaptive Protocols for Managing Large Ad Hoc Networks," *Proc. IEEE ICC*, June 2001, pp. 1539–43.

[8] J. Haartsen, "The Bluetooth Radio System," *IEEE Pers. Commun.*, Feb. 2000, pp. 28–36.

[9] J. Zander, "Performance of Optimum Transmitter Power Control in Cellular Radio Systems," *IEEE Trans. Vehic. Tech.*, Feb. 1992, pp. 57–62.

[10] A. J. Goldsmith and S. G. Chua, "Variable-rate Variable-power MQAM for Fading Channels," *IEEE Trans. Commun.*, Oct. 1997, pp. 1218–30.

[11] N. Bambos, "Toward Power-sensitive Network Architectures in Wireless Communications: Concepts, Issues, and Design Aspects," *IEEE Pers. Commun.*, June 1998, pp. 50–59.

[12] P. Gupta and P. R. Kumar, "The Capacity of Wireless Networks," *IEEE Trans. Inform. Theory*, Mar. 2000, pp. 388–404.

[14] P. Agrawal, "Energy Efficient Protocols for Wireless Systems," *Proc. IEEE Intl. Symp. Personal, Indoor, Mobile Radio Commun.*, Sept. 1998, pp. 564–69.

[15] J. M. Kahn, R. H. Katz, and K. S. Pister, "Emerging Challenges: Mobile Networking for "Smart Dust," *J. Commun. Net.*, Aug. 2000, pp. 188–96.

[16] A. Chandrakasan and R. W. Brodersen, *Low Power Digital CMOS Design*, Norwell, MA: Kluwer, 1995.

[17] S. Xu and T. Saadawi, "Does the IEEE802.11 MAC Protocol Work Well in Multihop Wireless Ad Hoc Networks?" *IEEE Commun. Mag.*, vol 39, no. 6, June 2001, pp. 130–37.

[18] B. Leiner, R. Ruther, and A. Sastry, "Goals and Challenges of the DARPA Glomo Program (Global Mobile Information Systems)," *IEEE Pers. Commun.*, Dec. 1996, pp. 34–43.

[19] R. Ramanathan and R. Rosales-Hain, "Topology Control of Multihop Wireless Networks Using Transmit Power Adjustment," *Proc. IEEE INFOCOM*, Mar. 2000, pp. 404–13.

[20] M. Ritter, R. Friday, and M. Cunningham, "The Architecture of Metricom's Microcellular Data Network and Details of Its Implementation as the 2nd and 3rd Generation Ricochettm Wide-area Mobile Data Service," *IEEE Emerging Tech. Symp. Broadband Commun. Internet Era*, 2001, pp. 143–52.

[21] M.N. Huhns, "Networking Embedded Agents," *IEEE Int'l. Comp.*, Jan./Feb. 1999, pp. 91–93.

[22] W. W. Gibbs, "As We May Live," *Sci. Amer.*, Nov. 2000.

[23] K. Negus, R. Stephens, and Lansford, "HomeRF: Wireless Networking for the Connected Home," *IEEE Pers. Commun.*, Feb. 2000, pp. 20–7.

[24] A. Schmidt, "How to Build Smart Appliances," *IEEE Pers. Commun.*, Aug. 2001, pp. 66–71.

[25] B. P. Crow *et al.*, "IEEE 802.11 Wireless Local Area Networks," *IEEE Commun. Mag.*, Sept. 1997, pp. 116–26.

[26] J. Haartsen and S. Mattisson, "Bluetooth — A New Low-power Radio Interface Providing Short-range Connectivity," *IEEE Proc.*, Oct. 2000, pp. 1651–61.

[27] J. Rabaey *et al.*, "PicoRadio Supports Ad Hoc Ultra-low Power Wireless Networking," *IEEE Comp.*, July 2000, pp. 42–48.

[28] J. Nilsson, B. Bernhardsson, and B. Wittenmark, "Stochastic Analysis and Control of Real-time Systems with Random Time Delays," *Automatica*, 1998, pp. 57–64.

[29] X. Liu *et al.*, "Effects of Communication Delay on String Stability in Vehicle Platoons," *IEEE Int'l. Conf. Intell. Transp. Sys.*, Aug. 2001.

[30] L. Xiao *et al.*, "Joint Optimization of Communication Rates and Linear Systems," *Proc. IEEE Conf. Dec. Cntl.*, Dec. 2001, also submitted to *IEEE Trans. Auto. Cntl.*

[31] Sergio Verdú, "On Channel Capacity Per Unit Cost," *IEEE Trans. Info. Theory*, Sept. 1990, pp. 1019–30.

[32] E. Biglieri, J. Proakis, and S. Shamai, "Fading Channels: Information Theoretic and Communications Aspects," *IEEE Trans Info. Theory*, Oct. 1998, pp. 2619–92.

[33] G. Caire and S. Shamai (Shitz), "On the Capacity of Some Channels with Channel State Information," *IEEE Trans. Info. Theory*, Sept. 1999, pp. 2007–19.

[34] D. N. Tse and S. Hanly, "Multiaccess Fading Channels- Part I: Polymatroidal Structure, Optimal Resource Allocation and Throughput Capacities and Part II: Delay Limited Capacities, *IEEE Trans. Info. Theory*, Nov. 1998, pp. 2796–831.

[35] L. Li and A.J. Goldsmith, "Capacity and Optimal Resource Allocation for the Fading Broadcast Channel, Part I: Ergodic capacity and Part II: Outage Capacity, *IEEE Trans. Info. Theory*, Mar. 2001, pp. 1083–27.

[36] M.-S. Alouini and A.J. Goldsmith, "Capacity of Rayleigh Fading Channels under Different Adaptive Transmission and Diversity-combining Techniques," *IEEE Trans. Vehic. Tech.*, July 1999, pp. 1165–81.

[37] G.J. Foschini and M.J. Gans, "On Limits of Wireless Communications in Fading Environments when Using Multiple Antennas," *Wire. Pers. Commun.*, 1998, pp. 311–35.

[38] E. Teletar, "Capacity of Multiantenna Gaussian Channels," *Euro. Trans. Telecommun.*, Nov. 1999, pp. 585–96.

[39] R. G. Gallager, "Energy Limited Channels: Coding, Multi-access, and Spread Spectrum," Unpublished, appeared as abstract in *Proc. Conf. Info. Sys. Sci.*, Princeton, NJ, Mar. 1988.

[40] S. Verdú, "Recent Results on the Capacity of Wideband Channels in the Low-Power Regime," this issue.

[41] H. Mandyam and A. J. Goldsmith, "Capacity of Finite Energy Channels," *Proc. Allerton Conf. Commun. Cntl. Comp.*, Oct. 2001.

[42] S. B. Wicker, *Error Control Systems for Digital Communication and Storage*, Englewood Cliffs: Prentice Hall, 1995.

[43] G. Ungerboeck and I. Csajka, "On Improving Data Link Performance by Increasing the Channel Alphabet and Introducing Sequence Coding," *IEEE Int'l. Symp. Info. Theory*, June 1976.

[44] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon Limit Error-correcting Coding and Decoding: Turbo Codes," *Proc. IEEE ICC*, June 1993, pp. 1064–70.

[45] *IEEE Trans. Info. Theory*, Special Issue on Codes and Graphs and Iterative Algorithms, Feb. 2001.

[46] S. B. Wicker and S. Kim, *Codes, Graphs, and Iterative Decoding*, Boston: Kluwer, to appear, 2002.

[47] C. Heegard and S. B. Wicker, *Turbo Coding*, Boston: Kluwer, 1999.

[48] C. Berrou and A. Glavieux, "Near Optimum Error Correcting Coding and Decoding: Turbo Codes," *IEEE Trans. Commun.*, Oct. 1996, pp.1261–71.

[49] Buckley, M. E. and Wicker, S. B., "The Design and Performance of a Neural Network for Predicting Decoder Error in Turbo-Coded ARQ Protocols," *IEEE Trans. Commun.*, Apr. 2000, pp. 566–76.

[50] S. Benedetto *et al.*, "Parallel Concatenated Trellis Coded Modulation," *Proc. IEEE ICC*, June 1996, pp. 974–78.

[51] R. Robertson and T. Worz, "Bandwidth-efficient Turbo Trellis-coded Modulation Using Punctured Component Codes," *IEEE JSAC*, Feb. 1998, pp. 206–18.

[52] C. Fragouli and R. Wesel, "Turbo-encoder Design for Symbol-interleaved Parallel Concatenated Trellis-coded Modulation," *IEEE Trans. Commun.*, Mar. 2001, pp. 425–35.

[53] R. McEliece, "Are Turbo-like Codes Effective on Nonstandard Channels," *IEEE Info. Theory Newsletter*, Dec. 2001 (based on Plenary Lecture in *Int'l. Symp. Info. Theory*, June 2001).

[54] A. Bahai, "Challenges of High Speed Wireless Systems: Beyond 3G and WLAN," *EE350 Seminar Series*, Stanford Univ., Jan. 2002.

[55] J. Proakis, *Digital Communications*, 3rd ed., McGraw-Hill, 1995.

[56] M. K. Simon and M.-S. Alouini, "Digital Communications over Fading Channels — A Unified Approach to Performance Analysis," Wiley-Interscience, 2000.

[57] J. Winters, "Smart Antennas for Wireless Systems," *IEEE Pers. Commun.*, Feb. 1998, pp. 23–27.

[58] V. Tarokh, N. Seshadri, and A.R. Calderbank, "Space-time Codes for High Data Rate Wireless Communications: Performance Criterion and Code Construction," *IEEE Trans. Info. Theory*, Mar. 1998, pp. 744–65.

[59] H. El Gamal and A.R. Hammons, "A New Approach to Layered Space-time Coding and Signal Processing," *IEEE Trans. Info. Theory*, Sept. 2001, pp. 2321–34.

[60] S.-T. Chung and A. Goldsmith, "Degrees of Freedom in Adaptive Modulation: A Unified View," *IEEE Trans. Commun.*, pp. 1561-1571, Sept. 2001.

[61] A. J. Goldsmith and P. P. Varaiya, "Capacity of Fading Channels with Channel Side Information," *IEEE Trans. Info. Theory*, Nov. 1997, pp. 1986–92.

[62] S. Kandukuri and N. Bambos, "Power Controlled Multiple Access (PCMA) in Wireless Communication Networks," *Proc. IEEE INFOCOM*, Mar. 2000, pp. 386–95.

[63] T. Holliday and A. Goldsmith, "Wireless Link Adaptation Policies: QoS for Deadline Constrained Traffic with Imperfect Channel Estimates," to appear, *Proc. IEEE ICC*, Apr. 2002.

[64] M.-S. Alouini, X. Tang, and A.J. Goldsmith, "An Adaptive Modulation Scheme for Simultaneous Voice and Data Transmission over Fading Channels," *IEEE JSAC*, May 1999, pp. 837–50.

[65] M.-S. Alouini and A.J. Goldsmith, "Area Spectral Efficiency of Cellular Mobile Radio Systems," *IEEE Trans. Vehic. Tech.*, July 1999, pp. 1047–66.

[66] X. Qiu and K. Chawla, "On the Performance of Adaptive Modulation in Cellular Systems," *IEEE Trans. Commun.*, June 1999, pp. 884–95.

[67] S. A. Jafar and A. J. Goldsmith, "Adaptive Multicode CDMA for Uplink Throughput Maximization," *Proc. IEEE VTC*, May 2001, pp. 546–50.

[68] S. Kandukuri and S. Boyd, "Simultaneous Rate and Power Control in Multirate Multimedia CDMA Systems," *IEEE Intl. Symp. Spread Spec. Tech. Appl.*, Sept. 2000, pp. 570–74.

[69] X. Tang and A. Goldsmith, "Admission Control and Adaptive CDMA for Integrated Voice and Data Systems," *Proc. IEEE VTC*, May 2001.

[70] C. Chien et al., Adaptive Radio for Multimedia Wireless Link," *IEEE JSAC*, May 1999, pp. 793–819.

[71] S. Kallel, "Analysis of Memory and Incremental Redundancy ARQ Schemes over a Nonstationary Channel," *IEEE Trans. Commun.*, Sept. 1992, pp. 1474–80.

[72] X. Qiu et al., "Performance Comparison of Link Adaptation and Incremental Redundancy," *Proc., IEEE. WL Commun. Net. Conf.*, Sept. 1999, pp. 771–75.

[73] P. Jung, P. W. Baier, and A. Steil, "Advantages of CDMA and Spread Spectrum Techniques over FDMA and TDMA in Cellular Mobile Radio Applications," *IEEE Trans. Vehic. Tech.*, Aug. 1993, pp. 357–64.

[74] T.S. Rappaport, *Wireless Communications — Principles and Practice, 2nd ed*, Prentice Hall, 2002.

[75] S. Verdú, *Multiuser Detection*, Cambridge Univ. Press, 1998.

[76] N. Abramson, "The Aloha System — Another Alternative for Computer Communications," *Proc. Fall Joint Comp. Conf.*, 1970, p. 37.

[77] D. Bertsekas and R. Gallager, *Data Networks*, 2nd ed., Prentice Hall, 1992.

[78] A. Chockalingam and M. Zorzi, "Energy Consumption Performance of a Class of Access Protocols for Mobile Data Networks," *Proc. IEEE VTC*, May 1998, pp. 820–24.

[79] P. Karn, "MACA: A New Channel Access Method for Packet Radio," *Proc. Comp. Net. Conf.*, Sept. 1990, pp. 134–40.

[80] V. Bharghavan et al., "MACAW: A Media Access Protocol for Wireless LAN," *Proc. ACM SIGCOMM '94*, Aug. 1994, pp. 212–25.

[81] C.-K. Toh et al., "MARCH: A Medium Access Control Protocol for Multihop Wireless Ad Hoc Networks," *Proc. IEEE MILCOM*, 2000, pp. 512–16.

[82] D.A. Maltz, J. Broch, and D.B. Johnson, "Lessons from a Full-scale Multihop Wireless Ad Hoc Network Testbed," *IEEE Pers. Commun.*, Feb. 2001, pp. 8–15.

[83] Z. J. Haas, J. Deng, and S. Tabrizi, "Collision-free Medium Access Control Scheme for Ad Hoc Networks," *Proc. IEEE MILCOM*, 1999, pp. 276–80.

[84] S.-L. Wu, Y.-C. Tseng and J.-P. Sheu, "Intelligent Medium Access for Mobile Ad Hoc Networks with Busy Tones and Power Control," *IEEE JSAC*, Sept. 2000, pp. 1647–57.

[85] A. B. MacKenzie and S. B. Wicker, "Selfish Users in Aloha: A Game-Theoretic Approach," *Proc. IEEE VTC*, Oct. 2001, pp. 1354–57.

[86] I. Katzela and M. Naghshineh, "Channel Assignment Schemes for Cellular Mobile Telecommunication Systems — A Comprehensive Survey," *IEEE Pers. Commun.*, June 1996, pp. 10–31.

[87] K. K. Parhi and R. Ramaswami, "Distributed Scheduling of Broadcasts in a Radio Network," *Proc. IEEE INFOCOM*, Mar. 1989, pp. 497–504.

[88] B. Selman, "Stochastic Search and Phase Transitions: AI Meets Physics," *Proc. Int'l. Joint Conf. Artl. Intell.*, 1995.

[89] C. P. Gomes et al., "Connection Between Phase Transitions in Complexity and Good Decoding," *Int'l. Symp. Inform. Theory Appl.*, Honolulu, HI, Nov. 5–8, 2000.

[90] D. Goodman et al., "Packet Reservation Multiple Access for Local Wireless Communications," *IEEE Trans. Commun.*, Aug. 1989.

[91] E. Uysal-Biyikoglu, B. Prabhakar and A. El Gamal , "Energy-Efficient Transmission of Packets in a Wireless Network," to appear *IEEE Trans. Net.*, also *Proc. IEEE INFOCOM*, Mar. 2002.

[92] R. D. Yates, "A Framework for Uplink Power Control in Cellular Radio Systems," *IEEE JSAC*, Sept. 1995, pp. 3141–47.

[93] G.J. Foschini and Z. Miljanic, "Distributed Autonomous Wireless Channel Assignment Algorithm with Power Control," *IEEE Trans. Vehic. Tech.*, Aug. 1995, pp. 420–29.

[94] L. Kleinrock and J. Silvester. "Optimum Transmission Radii for Packet Radio Networks or Why Six is a Magic Number," *Proc. IEEE Nat'l. Telecommun. Conf.*, Dec. 1978, pp. 4.3.1–5.

[95] P. Gupta and P. R. Kumar, "Critical Power for Asymptotic Connectivity," *Proc. IEEE Conf. Dec. Contl.*, Dec. 1998, pp. 1106–10.

[96] V. Rodoplu and T. H. Meng, "Minimum Energy Mobile Wireless Networks," *IEEE JSAC*, Aug. 1999, pp. 1333–1400.

[97] B. Krishnamachari et al., "Critical Density Thresholds in Distributed Wireless Networks," to appear, *Festschrift for Ian Blake*, 2002.

[98] E. Royer and C.-K. Toh, "A Review of Current Routing Protocols for Ad Hoc Mobile Wireless Networks," *IEEE Pers. Commun.*, Apr. 1999, pp. 46–55.

[99] M. R. Pearlman, Z. J. Haas, and S. I. Mir, "Using Routing Zones to Support Route Maintenance in Ad Hoc Networks," *Proc. IEEE WL Commun. Net. Conf.*, Sept. 2000, pp. 1280–84.

[100] A. Tsirigos and Z. J. Haas, "Multipath Routing in the Presence of Frequency Topological Changes," *IEEE Commun. Mag.*, Nov.2001, pp. 132–38.

[101] S.-J. Lee and M. Gerla, "Dynamic Load-aware Routing in Ad Hoc Networks," *Proc. IEEE ICC*, June 2001, pp. 3206–10.

[102] J.-C. Cano and P. Manzoni, "Evaluating the Energy-consumption Reduction in a MANET by Dynamically Switching-off Network Interfaces," *Proc. IEEE Symp. Comp. Commun.*, 2001.

[103] W. Ye, J. Heidemann, and D. Estrin, "An Energy-efficient MAC Protocol for Wireless Sensor Networks," to appear, *Proc. IEEE INFOCOM*, Mar. 2002.

[104] M. B. Pursley, H. B. Russell, and J. S. Wysocarski, "Energy-efficient Routing in Frequency-hop Networks with Adaptive Transmission," *Proc. IEEE MILCOM*, 1999, pp. 1409–13.

[105] A. Michail and A. Ephremides, "Energy Efficient Routing for Connection Oriented Traffic in Ad-hoc Wireless Networks," *Proc. IEEE PIMRC*, Sept. 2000, pp. 762–66.

[106] L. Subramanian and R. H. Katz, "An Architecture for Building Self-configurable Systems," *Mobile Ad Hoc Net. Comp.*, 2000.

[107] R. Jain, A. Puri, and R. Sengupta, "Geographical Routing Using Partial Information for Wireless Ad Hoc Networks," *IEEE Pers. Commun.*, Feb. 2001, pp. 48–57.

[108] L. Zhou and Z. J. Haas, "Securing Ad Hoc Networks," *IEEE Net.*, Nov/Dec. 1999, pp. 24–30.

[109] R. Karri and P. Mishra, "Energy Management of Secure Wireless Sessions," preprint.

[110] T. Cover and J.A. Thomas, *Elements of Information Theory*, Wiley Interscience, 1991.

[111] M. Grossglauer and D.N. Tse, "Mobility Increases the Capacity of Ad Hoc Wireless Networks," *Proc. IEEE Infocom*, Mar. 2001, pp. 1360–69.

[112] P. Gupta and P. R. Kumar, "Towards an Information Theory of Large Networks: An Achievable Rate Region," *Proc. IEEE Intl. Symp. Info. Theory*, June 2001, p. 159.

[113] A. Alwan et al., "Adaptive Mobile Multimedia Networks," *IEEE Pers. Commun.*, Apr. 1996, pp. 34-51.

[114] M. Mirhakkak, N. Schult, and D. Thomson, "Dynamic Bandwidth Management and Adaptive Applications for a Variable Bandwidth Wireless Environment," *IEEE JSAC*, Oct. 2001, pp. 1985–97.

[115] R. Ramanathan and R. Hain, "An Ad Hoc Wireless Testbed for Scalable, Adaptive QoS Support," *IEEE WCNC*, Nov. 2000, pp. 998–1002.

ADDITIONAL READING

[1] L. Bao and J. J. Garcia-Luna-Aceves. "Collision-free Topology-dependent Channel Access Scheduling," *Proc. Milt. Commun. Conf. (MILCOM)*, 2000, pp.507–11.

[2] *IEEE Pers. Commun.*, Special Issue on Advances in Mobile Ad Hoc Networking,Feb. 2001.

[3] *IEEE Proc.*, Special Issue on Packet Radio networks, Jan. 1987.

[4] V.O.K.~Li L.C.~Pond, "A Distributed Time-slot Assignment Protocol for Mobile Multi-hop Broadcast Packet Radio Networks," *Proc. Milt. Commun. Conf. (MILCOM)*, 1989, pp. 70–4.

BIOGRAPHIES

ANDREA GOLDSMITH (andrea@wsl.stanford.edu) received her B.S., M.S., and Ph.D. degrees in electrical engineering from the University of California at Berkeley in 1986, 1991, and 1994, respectively. From 1986 to 1990 she was affiliated

While there is still work to be done on improving link, medium access, network, and application protocols, the interactions across these different protocol layers invite novel cross layer designs that exploit these interdependencies.

Recent breakthroughs in ad hoc wireless network design have come from thinking outside the box: the box of layered protocol designs, the box of wireline protocols, the box of guaranteed quality-of-service for demanding applications, and the box of application-agnostic networks.

with Maxim Technologies, where she worked on packet radio and satellite communication systems, and from 1991 to 1992 she was affiliated with AT&T Bell Laboratories, where she worked on microcell modeling and channel estimation. She was an assistant professor of electrical engineering at the California Institute of Technology from 1994 to 1998, and then joined the Electrical Engineering Department at Stanford University where she is now an associate professor. Her research includes work in capacity of wireless channels and networks, wireless communication theory, adaptive modulation and coding, multiantenna channels, communications and control, and adaptive resource allocation for cellular systems and ad hoc wireless networks. She is a Terman Faculty Fellow at Stanford and a recipient of the Alfred P. Sloan Fellowship, a National Science Foundation CAREER Development Award, the Office of Naval Research Young Investigator Award, a National Semiconductor Faculty Development Award, an Okawa Foundation Award, and the David Griep Memorial Prize from the University of California at Berkeley. She is an editor for *IEEE Transactions on Communications* and *IEEE Wireless Communications*.

STEPHEN B. WICKER (wicker@ee.cornell.edu) received his B.S.E.E. with High Honors from the University of Virginia in 1982. He received an M.S.E.E. from Purdue University in 1983 and a Ph.D. degree in electrical engineering from the University of Southern California in 1987. He is professor and associate director for research of the School of Electrical and Computer Engineering at Cornell University. He is the author of *Codes, Graphs, and Iterative Decoding* (Kluwer, 2002), *Turbo Coding* (Kluwer, 1999), and *Error Control Systems for Digital Communication and Storage* (Prentice Hall, 1995). He is also co-editor of *Reed-Solomon Codes and Their Applications* (IEEE Press, 1994), and was Associate Editor for Coding Theory and Techniques for *IEEE Transactions on Communications*. He recently concluded his second term as a member of the Board of Governors of the IEEE Information Society. He teaches and conducts research in wireless information networks, self-configuring systems, and error control coding. His current research focus is on the use of probabilistic models and game theory in the development of highly distributed, adaptive, wireless networks. He was awarded the 1998 Cornell College of Engineering Michael Tien Teaching Award and the 2000 Cornell School of ECE Teaching Award. He has served as a consultant in wireless telecommunication systems, error control coding, and cryptography for various companies and governments in North America, Europe, and Asia.