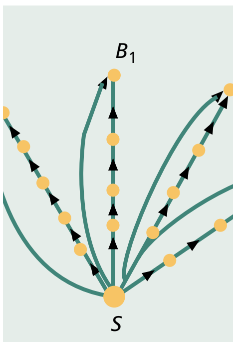
ENERGY CONCERNS IN WIRELESS NETWORKS

ANTHONY EPHREMIDES, UNIVERSITY OF MARYLAND



ABSTRACT

The focused attention of the community on energy issues in the study and design of wireless networks has spurred a great deal of recent research on the subject. In this article a brief overview is provided of what constitutes the major energy efficiency issues in ad hoc networks. Emphasis is placed on key conceptual points, which are then illustrated in the case study of wireless multicasting of connection-oriented traffic.

INTRODUCTION

Although wireless networks have existed for many years already, explicit concern about their energy-efficient operation has emerged only recently. It is quite evident that when the power source is either costly or in short supply, energy efficiency is of paramount importance. In some wireless network applications, energy is actually entirely non-renewable and is thus an overriding constraint for the design and operation of the network.

Although wireless networks have existed for many years already, explicit concern about their energy-efficient operation has emerged only recently. In some wireless network applications, energy is actually entirely non-renewable and is thus an overriding constraint for the design and operation of the network.

There are four major classes of wireless networks, and it is useful to distinguish them as we try to make the case for energy efficiency. The first (and oldest) with which most people are familiar is the class of cellular networks. To some, this is the only type of wireless network. However, cellular networks involve wireless transmission in only the first and/or last segment of a communication path. Thus, they exhibit few of the characteristics of wireless networking. Nonetheless, it is clear that mobile users are concerned about the longevity of the batteries in their handheld devices, even though these can be replaced or recharged. Furthermore, even at the base station there is a desire for low energy consumption, because, despite the availability of energy supply, there are serious concerns about excessive heat generation.

Another class of wireless networks is what people refer to as wireless local area networks (WLANs). These are truly and entirely wireless, but require only single-hop transmission. Typical wireless LANs involve laptops with Bluetooth or 802.11 cards in them that are in close proximity to power supplies. For them, energy efficiency does not seem as pressing, although there can be WLANs that stand alone outside buildings and for which, therefore, long battery life is also important.

The third class consists of networks that utilize satellite links. In these, energy consumption is a serious concern despite the possibility of recharging the solar cells onboard the satellite. And, of course, at the earth stations there are concerns about heat generation (if they are fixed) or battery life (if they are mobile).

The fourth, and most interesting, class is what we call ad hoc networks. The term has evolved to mean any network that can be set up wirelessly without the use of infrastructure. Thus, cellular and satellite networks do not fall into this category, while WLANs may. We distinguish here the meaning of the term to characterize only those infrastructureless networks that require multiple hops for connecting all the nodes to each other. These are networks that until recently were of interest mostly in military applications, but now there is also widespread commercial interest for cases of sensor networks in which multiple micro-embedded devices are interconnected in autonomous systems that are deployed only once. The latter clearly survive only as long as their original batteries are capable of providing the energy they need. But all ad hoc (multihop) networks are critically dependent on the rate of energy consumption. In fact, it is in this class of networks that all aspects of energy efficiency are most clearly displayed; thus, they will be the main focus of the rest of this article.

The main discriminator of ad hoc (multihop) networks that sets them apart from other wireless network classes is the need to *relay* (and, hence route) messages. This is a key difference that makes the ad hoc network an entirely wireless analog of the classical wired one, in the sense that the entire protocol stack from the physical to the application layer must be designed on the basis of no fixed infrastructure.

In addition to mobility and the peculiarity of the wireless transmission medium (fades, noise, interference, etc.), the main characteristics of such ad hoc networks directly related to design and operation are vertical layer integration and criticality of energy consumption.

As far as layer integration is concerned, it is increasingly understood today that the separation of network functions into layers (as done by the International Standards Organization, ISO, through its original introduction of open systems interconnection, OSI), was, in a way, the "original sin" in networking. Although it facilitated initially the comprehension of network operation and simplified somewhat the immense complexi-

1070-9916/02/$17.00 © 2002 IEEE

ty of total network design, it ended up masking the intricate interdependencies among the quantities and variables that resided at separate layers. Thus, it led to suboptimal results.

The "contribution" that the study of ad hoc wireless networks has made is the demonstration of the inevitability of coupling among the layers. In particular, the use of wireless transmission media necessitated the joint consideration of the physical layer along with the medium access control (MAC), link, network, and transport layers. Thus, asynchronous transfer mode (ATM) over wireless, TCP over wireless, mobile IP, and so on represent recent widespread lines of research and investigation that illustrate the need for layer coupling. For example, TCP, as designed for the Internet, interprets the loss of a packet as the result of buffer overflow due to congestion. On wireless links, however, a packet may be "lost" because of errors on the channel. Similarly, the effects of wireless link on quality of service requires modification of the ATM switching fabric. In this article we will not focus on this coupling. Rather, we will focus on the other main characteristic of ad hoc wireless networks: the importance of energy efficiency. It will become apparent, however, that the study of energy consumption involves, in yet another way, the coupling among the layers in the protocol stack.

In developing this brief review we follow this outline. The notion of a wireless link is examined in some detail so as to motivate what follows. We focus on how and where energy savings can be achieved and differentiate the case of energy-efficiency from the case of energy-constrained operation. We focus on a case study. We consider source-initiated multicasts for session traffic in static ad hoc networks. The case of wireless multicasting is an especially interesting one in ad hoc wireless networks. First of all, it includes and extends the case of unicast routing, which has been investigated extensively in recent years. But, more important, it illustrates best the new trade-offs between performance and energy consumption. Thus, a substantial portion of this article will focus on this case study. Finally, we address briefly the notion of "capacity" under energy constraints, and finally draw a few principal conclusions.

THE NOTION OF A WIRELESS NETWORK

We have all learned to draw a graph to depict a communication network, as in Fig. 1. This is a useful and accurate depiction of the network topology when the nodes are interconnected with dedicated wired lines. The tendency has been to do the same when the network under consideration is a wireless one, and that has been the cause of many misconceptions and much fallacious reasoning. If there are no "hard-wired" connections between the nodes, the notion of a "link" between, say, nodes A and B is an entirely relative one. In fact, it is so relative that links in a wireless network should be thought of as "soft" entities that are almost entirely under the control of the network operator.

To see why this is the case, it is important to understand that the radiated energy from

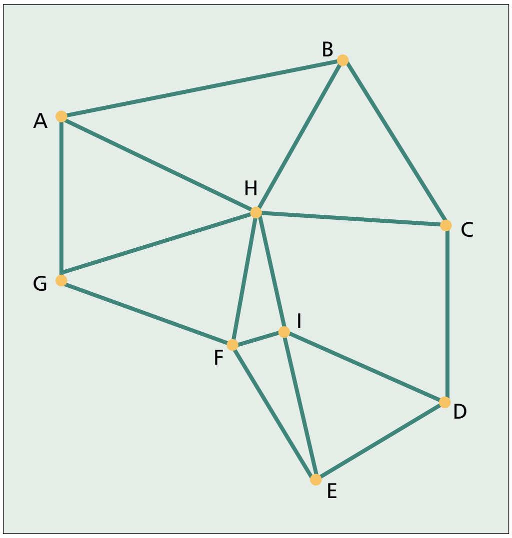


Figure 1. *A network graph.*

the antenna of a transmitting node travels over unlimited distances. As it scatters in the surrounding space, smaller and smaller fractions of it are capable of reaching remotely located receiver antennas. Nonetheless, finite amounts of such energy do indeed reach a node receiver, no matter how far that node is. Thus, it would seem that any wireless network is in fact a fully connected or mesh network (just like a WLAN). Of course, this is not a useful model because the signal strength is severely attenuated as the signal travels away from the transmitter. In fact, it decays nonlinearly according to the formula

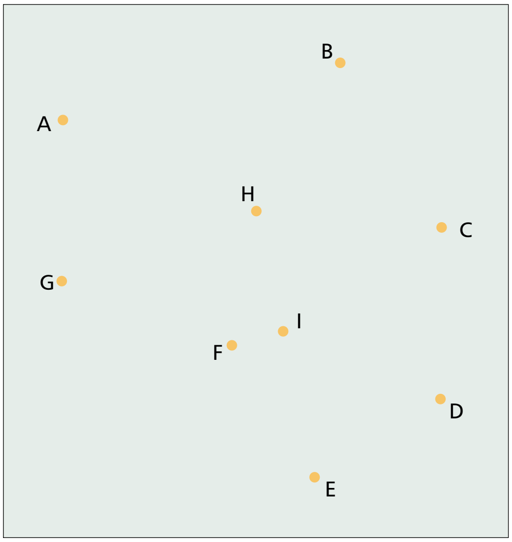
$$S(r) = Sr^{-\alpha}, \qquad (1)$$

where S is the amplitude of the transmitted signal, r is the distance from the transmitter, $S(r)$ is the amplitude of the received signal at distance r, and α is a parameter whose value ranges from 2 to 4. In order to achieve "successful" reception it is first necessary to establish a desired quality of service in terms of the maximum acceptable value of the bit error rate (BER). Typically, 10^{-9} or lower is set for data and 10^{-6} or higher for speech. Depending on the application and the desired fidelity of reception, this value can range from 10^{-2} to 10^{-11}. The way the wireless physical and link layers operate requires choices: of waveforms to represent the digital units (or symbols), modulation/demodulation schemes, coding/decoding schemes, antenna profile, detection structure, and additional signal processing elements at the transmitter and receiver. These choices are elaborate and complex, and constitute the subject matter of digital communication theory. In addition, the achievement of successful reception depends on the amount of noise or other impairments (interference, fades) the channel introduces. And, very importantly, it depends on the bandwidth of the channels and hence on the "rate" of transmission (i.e., the time interval between successively transmitted symbols). Putting this all together, the criterion for successful reception is summarized by the requirement that

$$SINR > \theta, \qquad (2)$$

If there are no "hard-wired" connections between the nodes, the notion of a "link" between, say, nodes A and B is an entirely relative one. In fact, it is so relative that "links" in a wireless network should be thought of as "soft" entities that are almost entirely under the control of the network operator.

Clearly then, energy concerns lead to further coupling among the layers in the protocol stack. It is also clear that under volatile environment conditions where the connectivities amongst the nodes keep changing, there is a need for highly adaptive protocols that continually sense, and react to, these variations.



■ Figure 2. *A wireless network graph.*

where SINR is the received signal-to-interference-plus-noise ratio at the receiver and θ is a threshold that depends on the detector structure, modulation/demodulation, and coding/decoding used. The SINR on the left side of Eq. 2 depends principally on the channel, other-user signals, transmit and receive antennas, and, more important, *RF transmission power* and *transmission rate*. The last two quantities — let us call them *P* and *R*, respectively (the power and bit rate) — determine the amount of signal energy packed in each symbol and are hence highly adjustable. It is far more difficult to modify the signal waveform or modulation, coding, and processing algorithms (although even these can be tunable in the software-defined radios of the future). Power and rate control, on the other hand, are implemented already in a variety of wireless devices.

The bottom line is therefore that whether a link exists or not depends on all these choices, starting from the desireed QoS BER and ending (for a given channel environment, given equipment and SP algorithms) with the chosen values of *P* and *R*. Clearly, boosting *P* may increase the number of links that are feasible from one node in an ad hoc network to other nodes, and reducing that value may decrease that number. Of course, a link that is feasible will become infeasible if the transmission on neighboring feasible links increases the interference on the link of interest and brings the value of SINR below the value of θ.

It should be clear, then, that the existence of a wireless link is a very volatile notion. Thus, the proper way of depicting a wireless network is simply via the location of its nodes, as in Fig. 2. And it should also be clear that the choice of the transmission power *P* may very well determine that existence. As the value of *P* has obvious energy consumption consequences, it is quite clear that energy concerns in a wireless network are closely related to the choice of transmission power, since the latter not only influences the amount of energy consumed on that link but, more important, determines which links are feasible and hence which paths can be used for routing to the final destination. Clearly, then, energy concerns lead to further coupling among the layers in the protocol stack. It is also clear that under volatile environment conditions where, due to mobility, channel fluctuations, jamming, and so on, the connectivities among the nodes keep changing, there is a need for highly adaptive protocols that continually sense and react to these variations. As will be seen, things are sufficiently complicated even without these temporal connectivity fluctuations. Therefore, we will not attempt to address the complications that result from mobility and channel variability in this article. However, we must recognize and point out their critical importance.

Energy Savings

Let us start with listing the ways in which energy is consumed (and hence potentially saved) in a wireless network. Clearly there are three major modes of operation for any wireless node. It is either transmitting, receiving, or simply "on." In the last mode it typically "listens" but is not actively receiving. In the transmitting mode energy is spent in two major ways. The first is in the front-end amplifier that supplies the power for the actual RF transmission. This includes the radiated energy as well as the internal heat losses in the antenna and the amplifier itself. The second is in the node processor that implements all the signal generation, formatting, encoding, modulation, memory access, and other signal processing functions. We call the first transmission energy and the second processing energy.

In the receiving mode, energy is consumed entirely by the processor, including the low-noise amplifier that boosts the output of the receiving antenna to levels suitable for demodulation, decoding, buffering, and so on. That is, in this case the consumed energy is only of the processing type. Finally, in the "on" mode, the energy consumed is again of the processing type (since the voltage controlled oscillator, VCO, is operating to be ready to commence demodulation of an incoming signal, and all circuits remain properly initialized and charged) but also possibly of some transmission type, since a listening device may be required by network protocol to emit periodic beacon signals. The grand total amount of energy spent per unit time while in the "on" mode is quite small compared to the receiving and transmitting modes. However, since in many applications (especially in sensor networks), a node may spend most of its lifetime simply in the "on" position without actually receiving or transmitting, this mode may represent the lion's share of the total consumed energy. In other applications, where a node is actively transmitting during most of the time in which it is not switched off, the amount spent in passive "on" mode is negligible and need not be taken into consideration. All of the aforementioned modes of operation involve all the components (hard and soft) that constitute the node. This includes the hardware units as well as the processing algorithms. Thus, to reduce the amount of energy needed to achieve a given communication task (e.g., the transmission or reception of a packet), it is necessary to examine the role of every one of these components and to also look at the way they interact.

Energy expenditures are affected by design choices made at every one of these levels. For example, the selection of batteries is crucial. They determine the total amount of energy that is available to the node. Interestingly, there is more to this than just the nominal amount of Coulombs stored in them. The pattern of draining energy from the battery actually affects the grand total that can be delivered. In [1] it was shown how the pulsed operation of each cell can increase the yield of a battery. The consequences are important. For example, they imply that a time-division multiple access (TDMA) protocol might be more energy efficient than other MAC protocols (if everything else that affects consumption remains the same). Incidentally, this observation illustrates yet another way in which there is coupling across the layers.

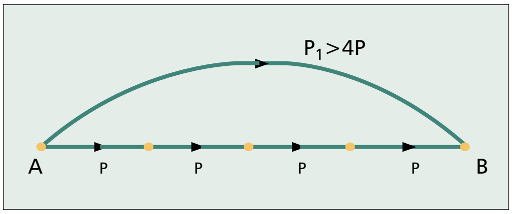
Next, energy can be saved by proper selection of hardware. Power amplifiers are known to be nonlinear. Thus, when driven to saturation (which may be necessary if transmission at maximum power is desired) they are very inefficient and consume a much higher amount of joules per joule delivered to the antenna. On the other hand, operating an amplifier at its linear range represents underutilization of its potential. Thus, amplifiers with better efficiency curves can save energy. In addition, the circuit layout of every chip has notorious energy consumption effects. In [2] there are ample discussion and pointers to energy-efficient circuit architectures. Note that the desire for energy efficiency in the circuit layout is motivated further by the need to reduce the thermal effects and avoid excessive heating of the device.

In addition to these, the choice of antenna has energy repercussions. Not only do the actual dimensions, material, and shape determine the energy efficiency of the antenna, but also its (possibly electronically determined) radiation pattern (e.g., directivity) has important effects on energy consumption.

One of the basic cornerstones of digital communications is the realization that the chosen combination of modulation/demodulation and coding/decoding determines the spectral efficiency of the system (i.e., the achievable number of bits per second per hertz). In other words they affect the relationship between the two sides of Eq. 2 by determining the value of θ [3]. Hence, they determine the needed value of received (and, hence, transmitted) signal strength for successful reception and thereby affect the amount of energy consumption.

In a similar vein, the choice of signal processing algorithm implementations — including their software specifications as well as their very large scale integration (VLSI) incarnations — have significant effects on energy expenditure. For example, choosing how to compress a signal, or store and/or retrieve it from memory does have measurable effects on energy consumption.

All these aforementioned observations are fairly well known and have been understood for some time. What is not yet clearly understood is how the simultaneous choice of the different design parameters affects the total energy efficiency in a wireless system. There are many efforts underway (e.g., [4]) that try to assess the interaction between these diverse choices.



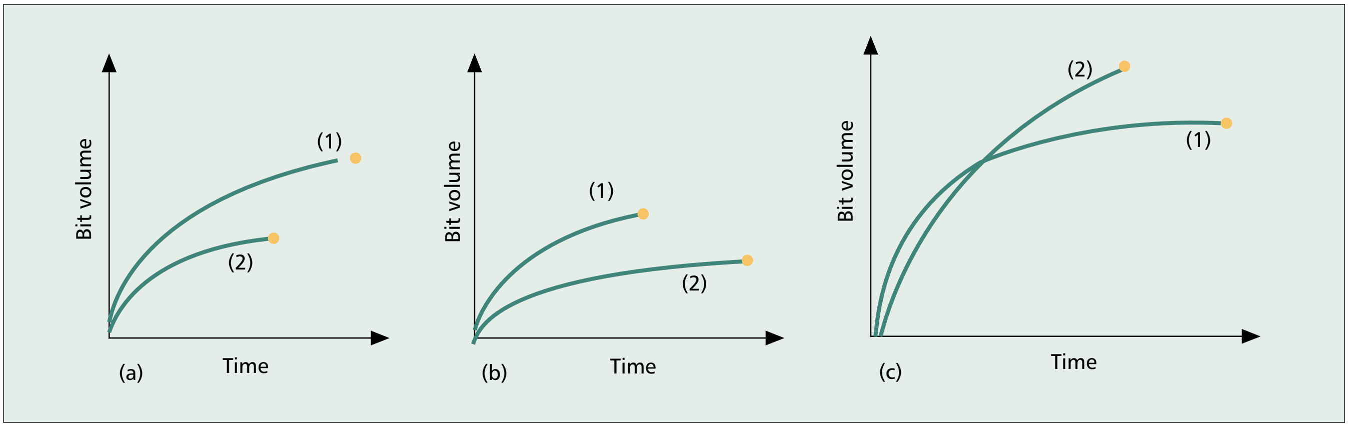
■ Figure 3. *Wireless routing.*

More important, however, it was not realized until very recently that the choice of higher-level protocols (e.g., routing or multicasting) has equally significant effects on energy consumption. It was of course realized that in code-division multiple access (CDMA) cellular systems, power control (which has been studied primarily as a means of boosting system capacity rather than controlling energy consumption [5]) does have an effect on energy expenditure. Yet, to date, there have been few efforts to incorporate power control explicitly in energy-efficient design of MAC protocols. One example is [6], where the authors consider the trade-off of the benefits of capture on throughput against those of time separation (and hence collision avoidance) between packets in contention-based access schemes under energy constraints. Much more remains (and needs) to be done, however, in that area.

Also, fairly recently the choice of schedules for pager devices to be turned on or off has been looked at from the energy conservation point of view [7]. Thus, steps toward exploiting design choices at higher layers of networking for energy savings are already underway. However, these first few timid attempts pale beside the explosive interest in energy-efficient protocol design that has developed in the last couple of years. It should be mentioned that the first observations that routing and MAC protocols have significant impact on energy expenditures were made in the context of an ARL-funded consortium, called Advanced Telecommunication Information-Distribution Research Program (ATIRP) that was one of the Army-sponsored Federated Laboratory Research efforts from 1995 to 2001. In [8] the simple observation was made that since attenuation of radio signals follows the nonlinear curve of Eq. 1, the choice of transmission power for routing a packet from A to B in Fig. 3 by relaying it along a tandem of nodes in a straight line requires less energy than routing it in a single hop. The importance of this observation lies in the fact that it suggests that in ad hoc wireless networks it is energy-efficient to choose long paths along a series of short hops rather than short paths along a series of long hops. Hence, it made plain the coupling between routing and transmission power choice.

However, things are not that simple. First of all, if energy efficiency is the only concern in a communication system, one might as well trans-

The choice of signal processing algorithm implementations has a significant effect on energy expenditure. For example, choosing how to compress a signal or how to store it and/or retrieve it from memory does have measurable effects on energy consumption.



■ Figure 4. *Performance of different strategies.*

mit nothing. Energy reserves will thus remain intact in perpetuity. Clearly communication performance is also of paramount interest. Thus, the choice of how to incorporate energy efficiency in the overall design is far from clear. One approach is to try to minimize energy consumption subject to throughput (or delay) staying above (or below) a certain threshold. Alternatively, one can try to maximize throughput (or minimize delay) per joule of expended energy. Neither of these approaches led to simple precise formulations or easy solutions.

Then we note that energy consumption (as mentioned earlier) does not occur only through transmission, but also through processing. So, if the decision is made to route via nearest neighbor toward any destination, the consequences are that the delay increases (due to the multiple hops) and the processing energy increases (due to the repeated relaying of the signal). In addition, it is far from clear what happens to the *overall* transmission energy, since to implement a nearest neighbor routing policy, significantly augmented overhead control traffic will be required to coordinate the establishment of the routing paths and access control protocols across the entire network.

Therefore, the introduction of energy efficiency considerations complicates the upper layer designs (e.g., routing) despite the simplistic observation in Fig. 3. At a minimum, the use of energy by intermediate relay nodes raises issues of coordination among node transmissions, and of desirability or appropriateness that may depend on the application supported by the network.

In ad hoc (multihop) wireless networks, the problem of routing has (not surprisingly) received more attention than any other design and operation problem. For example, see [9, references therein]for only a segment of the work on ad hoc network routing.[1] So it is no surprise that the lion's share of the recent work on energy efficiency in ad hoc networks seems to be concentrating again on routing [10–13]. However, the community has embraced the notion of energy efficiency in quite a broad way and has made energy concerns one of the cornerstones in overall ad hoc network design. The approaches taken so far and the particular problems chosen by the various researchers on energy efficiency have been very recent and very diverse. The proceedings of IEEE INFOCOM 2001 and 2002 as well as those of recent Mobicom and Mobihoc conferences are the best examples of this diversity. It would be impossible to do them justice by summarizing them here. Additionally, it is too early to detect a common or prevailing thrust in these investigations. They are characterized by the somewhat disorderly and chaotic nature of a multifront and multi-objective area of emerging research.

Before we continue with the examination (in some more detail) of the case of wireless multicasting in the next section, we should make an important distinction. We should differentiate between treating energy as a cost function and treating it as a hard constraint. These are two very different situations. In the former, the viewpoint of the designer, albeit complex in execution, is simple enough in perception. The objective is simply to minimize the amount of energy per communication task. Although the definition of a communication task can quickly become convoluted, the idea is to treat energy as an expensive but inexhaustible resource. For handheld devices, for example, we want the batteries to deliver as much functionality as possible before they are recharged or replaced. But they *can* be recharged or replaced.

On the other hand, when energy is a hard constraint, it is not a question anymore of using it efficiently because it is expensive. In this case the energy is *not* renewable *at all*. It is all you have, period! Of course, for this reason it is even more imperative to use it economically. But at the same time we want to perform some communications function. The temptation to slide toward the preposterous solution of hoarding energy forever by transmitting nothing (mentioned earlier) is clearly increased. But obviously this solution does not make sense. So, what is then the right approach?

Even at the conceptual level, the finite energy case is far more complicated. One must question what the fundamental objectives are. Clearly, longevity of the network is an important one.[2] But some form of communication performance must also be an objective. Since by definition the network will only last a finite amount of time, what is a reasonable communication performance objective? The throughput? The total data volume delivered? Consider, for example,

[1] Again, to give proper credit, it is necessary to mention that the very first attempt to design and analyze an ad hoc wireless network intelligently and systematically can be traced back to the design of an Intra-Task-Force (ITF) network by a team at NRL in the early '80s [14–15]. The programs that were initiated by the Defense Advanced Research Projects Agency (DARPA) on low-cost packet radio and survivable radio networks followed soon thereafter.

[2] A complicating issue when dealing with a network of wireless nodes is to define when the network "dies;" is it when the first node runs out of energy? Is it when all nodes do? (If you think about it there may be one node, the last, that will not die since all others have died, and hence it can't spend energy anymore). Or is it when a fraction of them dies?

the hypothetical plots in Figs. 4a, 4b, and 4c. They depict the performance of two alternative strategies for constructing multicast trees in the same wireless network under the same traffic load. These strategies are referred to as protocols (1) and (2). The performance is measured by the total volume of bits cumulatively delivered by all nodes of their destinations over the lifetime of the network. In each figure the total bit volume is sketched against time for the two alternative protocols. In Fig. 4a it is clear that strategy 1 is superior because it outperforms strategy 2 in every reasonable sense. It has higher volume (not only at the end but also at every intermediate time instant), it causes the network to last longer, and its instantaneous throughput is always superior.

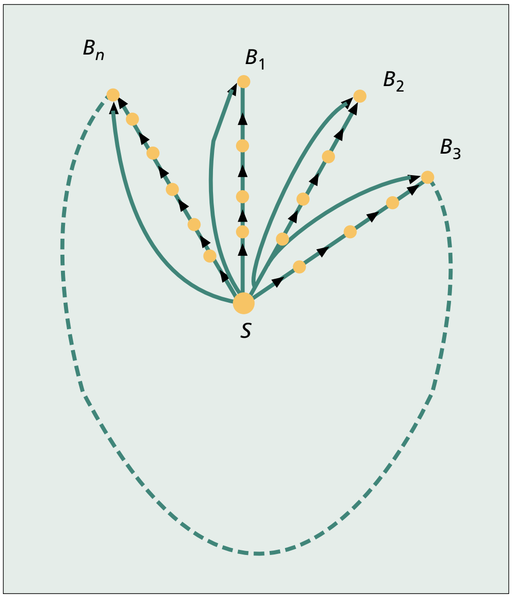
In Fig. 4b, it is a little harder to decide. Strategy 1 delivers more data at the end but the network dies earlier. In sensor networks, for example, longevity is very important. Which strategy is preferable? And, in Fig. 4c, it is even more complicated. Now the volume curves cross each other. So strategy 1 provides a more vigorous network in the beginning that outperforms strategy 2 until a crossover time, when the network under strategy 1 turns sluggish (although it does live longer), but strategy 2 delivers a higher traffic volume overall. Again, which strategy would one choose?

Lest the reader think this is an artificial, hypothetical scenario, it should be added that we have actually encountered all three profiles shown in Fig. 4 during our studies of alternative strategies for multicasting with an energy constraint. As seen in detail in the next section, the definition of bit volume in multicasting counts the total number of bits *delivered* rather than transmitted. Clearly, a single bit transmission in a wireless environment may result in multiple bit receptions since it can be received simultaneously by several neighboring nodes.

In conclusion, there is a need for some careful thinking in formulating objectives for the design and operation of energy-efficient wireless networks, especially in the case of a hard constraint on available energy.

Review of a Case Study: Multicasting

To illustrate some of the points made so far, and to add some specificity to our remarks, we will outline the study of a concrete example, that of source-initiated multicasting. The reason we chose multicasting is twofold. First, it is a generalization of the routing problem, which has received a great deal of attention; second, it introduces a new twist to the trade-offs involved in energy efficiency. Recall the argument made in connection with Fig. 3, where it became apparent that, at least as far as transmission energy of the payload traffic is concerned, it is advantageous to route over contiguous, short, multiple hops. But in multicasting, as depicted in Fig. 5, if the intended receivers are B_1, B_2, …, B_n, the grand total energy needed to transmit via short hops may very well exceed that to transmit via a single hop (assuming all destinations can be



■ Figure 5. *Wireless multicasting.*

reached via a single transmission). In other words, the disadvantage of nonlinear attenuation is mitigated by the advantage of simultaneously reaching multiple destinations. At the outset it is unclear which is the best way to resolve this trade-off. Thus, in addition to the complexities outlined in the previous section, we have a new dimension of complication. Also, note that when the multicasting problem is fully immersed in the actual wireless network environment, it is complicated further by MAC considerations. It is far from clear whether multiple short hops or a single long hop create more interference. And interference adds its own energy consumption consequences.

Note that here we do not consider the information theoretic case of broadcasting, in which different signals are transmitted to the source node's neighbors over the individual channels between the source and these neighbors. This is an entirely different and very interesting case in which the main objective is the determination of the region of achievable rates at which the source can communicate simultaneously to its neighbors. This case is subsumed somewhat in a brief discussion in a later section.

In considering the multicasting problem we choose the source-initiated case rather than the usual Internet-like case in which the source is transmitting continuously and receiving nodes decide to join at will. The latter is more appropriate for entertainment or basic data distribution like weather, stock prices, and so on. In a wireless ad hoc network the applications we have in mind are more of the sensor or battlefield type. In both cases different nodes generate at random instants new information that is targeted toward specific subsets of network nodes. Thus, a multicast tree path is needed from each source to its destinations.

Just to be sure that the problem under consideration is clearly understood, let us define explicitly what we mean by multicasting. We assume that every node is a potential source of

Note that when the multicasting problem is fully immersed in the actual wireless network environment, it is complicated further by MAC considerations. It is far from clear whether the multiple short hops or the single long hop create more interference. And, interference adds its own energy-consumption consequences.

Note that with session-oriented traffic, bandwidth resources need to be reserved and dedicated along the tree path of the session for its entire duration. These resources are released upon termination of the session. This is generally not the case in connectionless or non-real-time traffic.

messages intended for a set of destination nodes that depends on the source node and the message. Thus, upon generation of each message, we need to decide the set of nodes to which the source needs to transmit this message in one hop. Hence, we need to decide its transmission power. Then we need to specify which of the nodes so reached need to relay this message further toward the intended destinations. Therefore, we need to decide the power at which they need to transmit and which of the nodes their transmissions reach need to relay this message further, and so on and so forth. That is, we need to construct a wireless transmission tree.

At this point an important distinction is needed as to the type of traffic we consider. We may have image or video or speech transmissions that require session-based operation, or datagram data that do not require bandwidth reservation and connection establishment but rather a store-and-forward mode of operation. Both cases are interesting and important. Our main focus will be first on connection-oriented traffic because it permits crisper crystallization and understanding of the trade-offs, and because some interesting (unintended and unanticipated) consequences arise. Note that with session-oriented traffic, bandwidth resources need to be reserved and dedicated along the tree path of the session for its entire duration. These resources are released at termination of the session. This is generally not the case in connectionless or non-real-time traffic.

So our model is the following. The node locations are randomly distributed, and each node independently generates (in a Poisson fashion) requests for establishing multicast trees to subsets of network nodes (which are independently chosen according to a uniform distribution). The objective, then, is to create an energy-efficient tree for each request. And, of course, the energy efficiency objective comprises the entire network and the long-term duration of its operation. Clearly, then, some objective functions need to be defined.

Before selecting such objective functions (also referred to as global performance measures), we must address the question of mobility. Wireless networks of most types are expected to be mobile and dynamic. However, to capture the effects of mobility as we initiate this investigation is rather difficult (if not impossible). For one thing, mobility will mask the trade-offs we have alluded to, and for another, there are some wireless networks that are indeed static (like the single deployment of stationary sensor devices). Thus, initially we consider a static wireless network (with the intent, of course, to eventually include mobility in the model).

Suppose that during the network's operation there is a total of N multicast requests. Let the ith such request have n_i intended receivers. Then a tree must be constructed to reach those n_i receivers from the source node. Constructing a tree, unlike the case of nonwireless networks where the links are well defined by the network graph, implies that every node on the tree must choose its transmission power and hence the set of links that it will establish toward its selected set of neighbors. The usual treatment of multicast tree construction in *wireline* networks is greatly simplified by the fact that every link has a well-defined metric. Thus, the problem becomes simply one of choosing a tree that reaches its destinations with minimum total cost (defined as the sum of the link costs on the tree). By the way, this easily formulated problem is a combinatorial complex one that is NP-complete [16] and has been extensively studied in the literature [17–19]. Due to its complexity, only heuristic techniques exist for its solution.

In the wireless context, however, things are worse than this. We do not have a well-defined link metric. For example, link bit error rate, delay, residual energy, and residual capacity are all appropriate quantities to serve as link metrics; but we do have the usual global objectives, such as total delay, throughput, blocking probability, total energy consumption, and so on. If we could find link metrics that map accurately to the global objectives, the wireless case would become similar to the wired one, but as we shall see, this is not easy (or even possible). Thus, we have some truly uncharted territory to explore.

Before proceeding further we must observe one additional complication. As the multicast requests keep coming in and different trees are generated, to actually carry out the transmissions we need to consider the MAC problem. If we assume contention-based protocols, TDMA-based scheduled transmissions, or CDMA signals, we have to deal with seriously intertwined variables and a situation that will be very difficult to resolve. In the literature there have been attempts to consider joint MAC and routing protocols (e.g., [20, 21]), but, first, these have not been considered in the context of multicasting, second, they have not taken energy efficiency into consideration, and third, they have been fairly ad hoc choices (which is fine). We find that things become somewhat more tractable if we assume channelized transmission — that is, frequency-division multiple access (FDMA) — and, in addition, transceiver-limited nodes. What this means is that we do not have interference problems as long as each node is transmitting at a different frequency, and we allow each node to be capable of transmitting and/or receiving multiple signals at the same time by virtue of having multiple transceivers. So, if a node has T transceivers, it is capable of participating in up to T different sessions, assuming there are at least T different frequencies available to it. If the total number of channels is F, the maximum number of distinct sessions each node can participate in is $\min(F, T)$. Although every node has its own T transceivers to control, the F channels are in a common pool and are accessible by every node.

As long as F and T are finite it is possible that either some multicast sessions will be blocked entirely or that some of the intended n_i destinations may be unreachable because of unavailability of unoccupied frequencies and/or transceivers. Thus, for the ith multicast session we denote by m_i the number of destinations that can actually be reached. Clearly, $m_i \leq n_i$. In addition to the complexity of resolving the frequency reuse problem by appropriate assignment of channels, in the wireless network case we have the additional problem that, as argued in an earlier section, every node receives residual energy from a signal transmitted by any other node, no matter how large their geographical separation. Thus, the frequency scheduling problem must be made on the basis of Eq. 2. Although in principle this can be done, it complicates the problem even further. Note, for example, that the analysis of channel assignment and frequency reuse in cellular networks initially avoided this problem by neglecting interference from more than a small number of cells away. Although this assumption is inaccurate, it permitted an approximate approach to the problem. In an ad hoc network, however, we do not have a cellular structure; thus, there are no natural boundaries of what might constitute no-interference zones.

One way to avoid this complication is to initially assume that both F and T are large enough so that every multicast request can be accommodated (so the energy-efficiency trade-off can be understood) and then to relax this assumption by first assuming T finite (to permit partial or total blocking due to lack of available resources) and then both T and F finite (in which case the scheduling of the available frequencies must be resolved).

Let us now return to the issue of global objectives. One such objective is simply the total expended energy. If the durations of the multicast sessions are independent and identically distributed, the expended energy is, on average, proportional to the total utilized power. If the sum of the transmission powers of all nodes that relay the signal on the multicast tree of the ith request is P_i (i.e., P_i is the total power needed to support the ith multicast signal), that objective function would be simply the sum of all the P_is as i varies from 1 to N. We also need, however, an objective function that will address the communication performance issue. Otherwise,

$$\sum_{i=1}^{N} P_i$$

can be minimized by rejecting every call request and thus spending zero energy. (By the way, although admission control can be, in principle, exercised for better overall performance, initially we assume that the network will not try to anticipate future gains and thus will try to accommodate every call request. Therefore, a call will be blocked — partially or totally — only if there are not available resources).

Another objective function (that addresses the issue of communication performance) is the total *throughput*. For session-oriented traffic we know that throughput (i.e., the average number of accepted calls per unit time) is maximized when the *blocking probability* is minimized. But in a multicast environment, where a call may be partially blocked, the notion of blocking probability is not well defined. The natural generalization of this notion is what we call *multicast efficiency* defined by

$$e = \frac{1}{N}\sum_{i=1}^{N}\frac{m_i}{n_i}$$

(in unicast routing, $n_i = 1$ and $m_i = 0$ or 1; hence, the efficiency reduces to the blocking probability).

Having these two functions, we now have the difficulty common to any multi-objective optimization problem: how to optimize both functions. As mentioned earlier, it is possible to impose a constraint on one and try to optimize the other or, equivalently, try to optimize a linear combination of the two (although theoretically the two approaches are equivalent, there are serious practical difficulties in deciding the values of the coefficients in the linear combination case since the two quantities have vastly different value ranges).

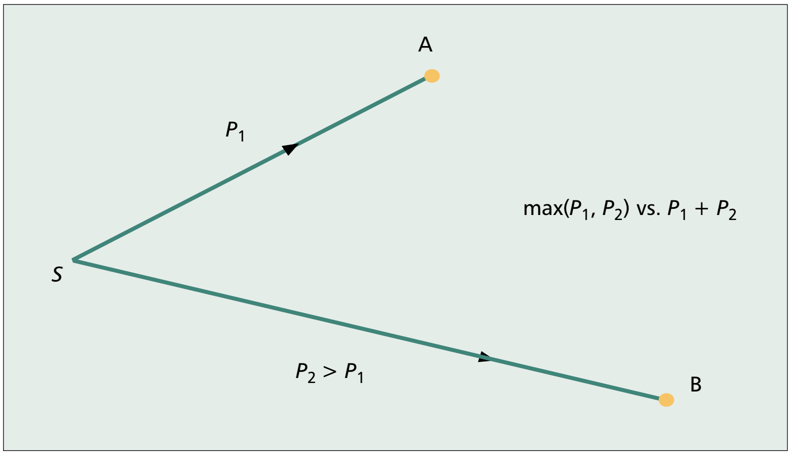
Alternatively, one can attempt to "fuse" the two objectives in a single function. There are numerous ways of doing that. We chose what we called the *yardstick* measure (since we might use it as the sole yardstick for comparing different multicast tree const ruction algorithms). The yardstick is defined as

$$Y \triangleq \frac{1}{N}\sum_{i=1}^{N}\frac{m_i}{n_i}\frac{m_i}{p_i};$$

that is, it is the average value (over all multicast requests) of the products of individual tree efficiencies and individual tree "throughput-per-joule" values. Note that m_i/n_i is a measure of the efficiency of the ith tree (in accordance with the definition we gave earlier) and m_i/P_i, which can be written as $m_i t_i/P_i t_i$ (where t_i is the duration of the ith multicast), represents the number of bits this multicast delivers (if t_i is measured in bit interval units) divided by the energy expended for that delivery. We do not wish to argue too strongly about the merits of this function. It just represents one choice (of many possible) that captures both throughput efficiency and transmission energy. Note that we are indeed ignoring processing energy, overhead traffic, and energy expended while being on. However, it is a straightforward step to include these in the formulation. This is done in [22–24] where many of the details of this work along with various generalizations of it can be found.

No matter which of these global performance measures we select, there is no way in which we can identify a link measure that will map perfectly to it. The basic reason for this is that no individual link-related quantity can capture the blocking probability (or the multicast efficiency we are using here). This is why in the long history of telephone network routing only heuristic solutions were developed for call routing (unlike the case of data networks where link delay relates directly to total delay and for which a variety of solid optimal algorithms have been developed). Blocking results from unavailability of resources to form a path from the source to the destination(s). The availability of such resources on a single link does not relate to total blocking in a simple and direct quantitative way. By contrast, total delay along a path is simply the sum of the individual delays along the links of the path. Thus, routing for connection-oriented traffic that minimizes blocking probability remains to a large extent an art, since the

Taking advantage of the node-centric (rather than link-centric) nature of ad hoc networks, we may decide on a node metric. A simple choice is simply the minimum power needed for a node to transmit successfully to a particular neighbor or the minimum power needed for a node to reach all of its children nodes on the tree.



■ Figure 6. *The wireless multicast advantage.*

methodologies of optimal routing that are used in data traffic are not applicable.

Still, to develop some algorithms for multicast tree construction (albeit suboptimal and heuristic), we do need to decide on a link metric. Alternatively, taking advantage of the node-centric (rather than link-centric) nature of ad hoc networks, we may decide on a node metric. A simple choice (that maps reasonably well to the total energy consumption metric) is simply the minimum power needed for a node to transmit successfully to a particular neighbor (that would yield a link metric) or the minimum power needed for a node to reach all of its children nodes on the tree (that would correspond to a node metric). Note the important distinction between the way these costs accumulate in the case of wireless networks and the way they do in nonwireless networks. In Fig. 6 we see that the cost (power) needed for S to reach A is p_1 and the cost (power) to reach B is p_2. If the multicast tree includes S and the downstream neighbors A and B, the total cost for this inclusion in a nonwireless network would be $p_1 + p_2$, while in the wireless network it would be max(p_1, p_2). This important difference (identified essentially at the beginning of this section) is termed the *wireless multicast advantage*, and forms the basis of our approach.

So, with these considerations as the background, we now propose and explain a simple centralized principle for constructing multicast trees. We focus only on the construction of a single tree (which is equivalent to assuming that all call requests are accepted, which in turn implies that the values of T and F are arbitrarily large), and we assume that the intent is to reach all other nodes in the network (i.e., we want to construct a broadcast tree).

Looking at Fig. 7 we start from the source node S and keep increasing its power until it reaches the closest of its neighbors (say, A_1); the power needed is p_1. Then we let S's power increase until it reaches its second closest neighbor (say, A_2). It takes power $p_2 > p_1$ to do that. We then compare p_2–p_1 (i.e., the incremental power needed to reach one additional neighbor) to the power p_3 that A_1 requires in order to reach A_2. If

$$p_3 > p_2 - p_1,$$

the beginning of the tree has S as the source and A_1 and A_2 as its children (at total "cost" p_2). If

$$p_3 < p_2 - p_1,$$

the beginning of the tree has S, A_1, and A_2 in tandem with total cost $p_1 + p_3$. This is the incremental power cost principle. The subsequent steps continue along similar lines. That is, we keep increasing the powers of all three nodes until a fourth node is reached; then we compare the increments on the power levels of all three nodes required to reach that fourth node and select (as an addition to our tree) the branch that requires the smallest increment. Clearly, as this process continues, it is possible to find that a choice at step k supercedes some earlier choices and reshapes the tree accordingly. For example, if in Fig. 7 we opted for the S, A_1, A_2 tandem and the fourth node is reached most economically (in the incremental power sense) by increasing the power of S to level p_4 (that manages to reach A_1 and A_2 as well), the tandem is replaced by a star centered at S with total "cost" p_4 . This procedure will yield a broadcast tree and is termed the Broadcast Incremental Power (BIP) protocol. It was first proposed and developed in [25].

If the destination node set is a subset of the network node set (i.e., if we want to execute a multicast rather than a broadcast), the heuristic we just described can be adjusted by first constructing a broadcast tree and then proceeding to prune it by removing all leaf nodes that are not destinations (including intermediate nodes that do not have destination nodes among their children). This procedure is called the Multicast Incremental Power (MIP) protocol (also described first in [25]). There are many ways in which this protocol can be modified and/or augmented. First, there is a "sweep" operation that can be performed by which, after we complete a tree, we revisit the intermediate nodes and see whether subsequent node additions supercede their role and thus eliminate them from the tree (or adjust their power downward). Then we can consider the case of finite values of F and T, and add frequency assignment heuristics to the tree construction, while at the same time we start registering partial (or total) call blockings. Finally, we can consider truly alternative tree construction methods for comparison purposes. A simple alternative is to consider each multicast call as the superposition of several unicast calls. Then, using the transmission power metric on each link, we may find the least cost unicast tree (by using a Bellman-Ford-type algorithm), and then construct the multicast tree by superposing the unicast trees and consolidating the needed transmissions (and therefore the costs) whenever possible. These variations and alternatives have been explored in [22–25] at considerable length and will not be expanded on here. It is worth noting, however, that these investigations and related performance evaluations have yielded interesting observations that reveal some of the dynamics that govern the energy-related trade-offs in ad hoc networks. Although no definitive conclusions can be drawn yet, plenty of food for thought has been generated. For one thing, the possibility of implementing these algorithms in a distributed way is currently being explored. Also, the possibility of incorporating mobility in the model (based on a distributed implementation) appears to be feasible now. To begin with, the relative notion of a link (as explained in an earlier section) permits considerable elasticity in network connectivity. This means that as nodes move away from other nodes that are transmitting to them, an increase of the transmission power enables the continued satisfaction of Eq. 2 and thus maintenance of the link. Conversely, when nodes move closer to each other, transmission powers can be lowered so that the existing connectivity can be maintained at lower energy cost. Alternatively, the rate can be actively (at the transmitter) or passively (at the receiver) adjusted toward the same end. Of course, as nodes move, alternative connectivities may become preferable to existing ones. How to best take advantage of this possibility is not clear. However, the node mobility (or channel fluctuations, for that matter) need not disrupt the connectivity of a network, nor necessitate emergency measures (unless the resulting changes are dramatic). Such elasticity permits the maintenance of ongoing sessions in the presence of moderate mobility.

If we turn our attention now to the case where the available energy is totally nonrenewable (i.e., hard-constrained), clearly the formulation outlined above is not satisfactory. Although the algorithms we described will continue to favor energy-efficient paths (on average), there is no provision for reacting to, or avoiding, disconnections due to exhaustion of the energy reserves at individual nodes. We mentioned earlier the difficulties associated with selecting appropriate global performance measures (since longevity becomes a new important component of the overall picture in this case). These difficulties are more acute here. If we insist on using some link- (or node-)based costs in the construction of the trees, the link (or node) metric must be revised. One way of doing this is by modifying the cost from A to B (for the link AB) or from A to its children (for the node-based cost case) to include the quantity $E_A(0)/E_A(t)$, where $E_A(t)$ represents the residual energy reserves of node A at time t, $t \geq 0$. That is, the cost quantity will be given by

$$P_A\left[\frac{E_A(0)}{E_A(t)}\right]^{\beta},$$

where P_A is the needed power with which A can reach B (or all its children) and β is a parameter that weighs the relative importance of the residual energy factor against that of the transmission power. Clearly, when $\beta = 0$, this metric reduces to the one we used before. In [26] there is an exhaustive discussion and analysis of the constrained energy case. What is worth mentioning here is the fact that when the new metric is used, all nodes tend to die (i.e., run out of energy) within a very short time of each other. In other words, this choice enforces a sort of load balancing that brings about a behavior that may or may not be desirable. In [27], where a similar problem was formulated, it was shown

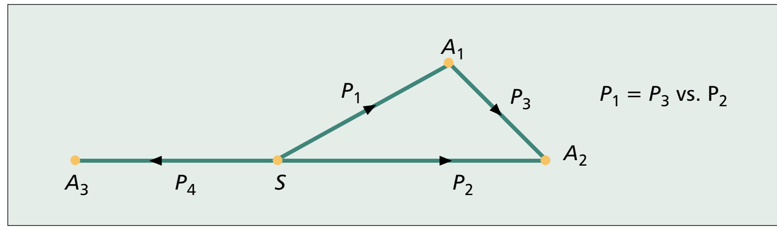


Figure 7. *The incremental power principle.*

that network longevity is maximized if and only if the longevity of the node that dies first is maximized. A consequence is that under optimal control of energy expenditures, the nodes tend to die almost all at once. Thus, our observations are at least consistent with one aspect of optimal behavior. Another thing worth mentioning is that when we compared alternative tree construction algorithms until the network died (and we used several definitions of network death), we observed behavior similar to that shown in Fig. 4, which reinforces the observation that the choice of what constitutes an appropriate global performance measure is not obvious at all.

An extension of this study has also been done recently, in which the antennas are assumed to have a controllable amount of directivity [28]. New degrees of freedom appear in the tree construction process, and there is (as expected) significant improvement in performance. A novel aspect of the reach vs. energy use trade-off is that when the beam becomes very narrow, a great deal of energy is saved, but the wireless multicast advantage is also reduced since there are fewer receiving nodes within the diminished beamwidth.

The case of unicast routing, although a special case of multicasting, deserves separate treatment, because it involves no wireless advantage. This has been looked at in [29] following the principal lines of the approach described for the multicast case. As mentioned earlier, there is a growing body of work concerning energy-efficient routing, but not for session-oriented traffic. A by-product of the proposed point of view in considering connection-oriented routing is that, by using link metrics, it becomes possible to solve blocking probability optimization problems by means of Bellman-Ford-type algorithms that have very low complexity. That alone is a remarkable achievement, even if the resulting solutions are only suboptimal.

Since much of the attention to energy-efficient routing in ad hoc networks has concentrated on data networks, it is useful to make a few closing comments on how the approach described so far for session traffic can be modified to apply to the connectionless case. The main difference is that we need not select multicast trees that require commitment of resources (bandwidth and transceivers). Packets can be stored at each node as in ordinary datagram routing. The case of unicast routing simply requires the determination of a path that has minimum cost. This requires that an appropriate link metric can be determined. The case of data multicast routing is, on the surface, similar to

Intuition seems to suggest that there must be a tight upper bound to the bit volume of an energy-constrained network. Some preliminary results of invariance of bit volume with respect to channel and transceiver resources [36] seem to support that suggestion. However, we do not yet know how to define such an upper bound, let alone how to compute it.

that of session multicast routing, in the sense that, again, we want to determine a least-cost tree for an appropriate link (or node) metric. The difference is that any sensible such metric must involve the queuing delay on each link. Unlike the unicast case where these delays can be determined and updated dynamically as the packet travels along its path (requiring only a next-hop decision at each node), in the multicast case we cannot do such a determination since Bellman-Ford-type philosophies cannot be used when each packet has multiple destinations.

Nonetheless a new link metric is needed no matter what. For the energy-efficient case (without a hard constraint on energy reserves) one can use the following metric:

$$P_A \cdot d_A^{\alpha},$$

where P_A is the needed power (to either one or multiple downlink nodes), d_A is the expected delay of each packet (directly related to the number of packets ahead of it in its storage queue), and α is a parameter for relative weighting. In the energy-constrained case we can modify this metric as

$$P_A \cdot d_A^{\alpha} \left[\frac{E_A(0)}{E_A(t)} \right]^{\beta},$$

where the new factor is identical to the one used before and β is another weighting coefficient. The data case has not been adequately investigated yet, but it does present some interesting new challenges with respect to tree construction.

It is hoped that this brief review of the multicasting case illuminates sufficiently the new, interesting, and very complex trade-offs that arise when energy efficiency is considered in an ad hoc network. It is also hoped that the inevitability of layer coupling in this case has become apparent.

Capacity Issues

We would be remiss in our review if we did not consider the notion of network capacity when energy-efficiency is desired. First of all, let us clarify the use of the term *capacity*. In the study of any kind of networks, the notion of Shannon-theoretic capacity has not been formalized yet [30]. Some recent work [31] has pointed the way toward formulating this notion, but still, not in the sense of delay-sensitive transmissions or bursty users. Thus, our discussion will shy away from Shannon capacity. Before doing so, however, it should be observed that some earlier work on capacity per unit cost has introduced the interesting notion of the number of bits per second per cost unit that can be transmitted if a separate cost function is provided. When expended energy is the cost, it is tempting to consider extending (or using) the work in [32–33] to define similar per-unit-cost capacities. However, this is just a preliminary possibility that has not been carefully developed.

Let us return, then, to the case of other capacities. In the study of networking it is possible to define the *transmission capacity* of a network as the maximum throughput it can sustain when the rates of transmission on each link are fixed. The maximum throughput is the grand sum of all sustainable source-destination transmission rates. While for a nonwireless network this is not a difficult quantity to compute, in a wireless network it becomes quite murky. Recent asymptotic results in [34, 35] gave a formal definition of this transmission capacity for ad hoc networks and showed some interesting properties in their asymptotic behavior with and without mobility that go beyond the scope of this article.

Nonetheless, there is one observation that is useful to make here that, in a way, combines the notion of transmission capacity with that of energy awareness. If the energy stored at each node is finite, the number of bits per second that can be transmitted can vary, but the total number of bits it can transmit over its lifetime is predetermined. If the link is fixed and the transmission over it is isolated from the rest of the network, the amount of energy per bit (as argued in an earlier section) is fixed, and thus the total number of bits is also fixed.

In a network, however, the node *chooses* over which link to transmit. The question, then, arises whether the total volume of bits the network can deliver over its lifetime is also predetermined or otherwise bounded. The answer is not clear. Intuition seems to suggest that there must be a tight upper bound to the bit volume of an energy-constrained network. Some preliminary results of invariance of bit volume with respect to channel and transceiver resources [36] seem to support that suggestion. However, we do not yet know how to define such an upper bound, let alone how to compute it.

Conclusions

It is by now clear that energy-aware, energy-efficient, or energy-constrained operation of wireless networks is not only desirable and important, but a basic requirement. It is also clear that energy can be saved by judicious design within and across any networking layer. Especially for the case of ad hoc wireless networks, protocol design at the networking layer (i.e., layer 3) in conjunction with transmission power choice (a choice that affects the physical and the MAC layers a well) offers the possibility of substantial performance improvement. In addition, the analysis of how energy expenditure, power of transmission, and route selection interact presents a whole set of novel conceptual and trade-off questions. From the practical viewpoint, the effects on performance are clearly of paramount importance. From the intellectual viewpoint, the new questions open up exciting investigation possibilities.

It is hoped that this general review of the broader energy issues and the somewhat more detailed review of the case of wireless multicasting will educate some readers and stimulate others to contribute to this exciting new area.

References

[1] C. F. Chiasserini and R. R. Rao, "Energy-Efficient Battery Management," *IEEE JSAC*, vol. 19, no. 7, July 2001, pp. 1235–45.

[2[Panel Report, "Wireless Technologies and Information Networks," Int'l. Tech. Res. Inst. (World Tech.y Div.), Baltimore, MD, July 2000.

[3] J. A. Proakis, *Digital Communications*, McGraw Hill, 1983.
[4] W. Stark *et al.*, "Low Energy Wireless Communication Network Design," *Proc. 38th Annual Allerton Conf. Commun., Control, Comp.*, Urbana, Illinois, Oct., 2000, pp. 785–805.
[5] S. Verdu, *Multiuser Detection*, Cambridge Univ. Press, 1998.
[6] W. Luo and A. Ephremides, "Power Levels and Packet Lengths in Random Multiple Access," to be published, *IEEE Trans. Info. Theory*.
[7] S. Dolev, E. Korach, and D. Yukelson, "The Sound of Silence: Guessing Games for Saving Energy in Mobile Environment," *Proc. IEEE INFOCOM*, New York, NY, Mar. 1999.
[8] R. Berry *et al.*, "Minimum Energy and Delay Routing in Packet Radio Networks," *Proc. 3rd Annual ARL FEDLAB Symp.*, Feb. 1999.
[9] M. Steenstrup, *Routing in Communication Networks*, Prentice Hall, 1995.
[10] V. Rodoplu and T. H. Meng, "Minimum Energy Mobile Wireless Networks," *IEEE JSAC*, vol. 17, Aug. 1999.
[11] S. Singh, M. Woo, and C. S. Raghavendra, "Power-aware Routing in Mobile Ad Hoc Networks," *Proc. ACM/IEEE Mobicom*, Oct. 1998, pp. 181–90.
[12] I. Stojmenovic and X. Lin, "Power-Aware Localized Routing in Wireless Networks," *Proc. IEEE Symp. Parallel and Dist. Processing Sys.*, May 2000.
[13] R. Ramanathan and R. Rosales-Hain, "Topology Control of Multihop Wireless Networks Using Transmit Power Adjustment," *Proc. IEEE INFOCOM*, Mar. 2000, pp. 404–13.
[14] D.J. Baker and A. Ephremides, "The Architectural Organization of a Mobile Radio Network via a Distributed Algorithm," *IEEE Trans. Commun.*, vol. 29, Nov. 1981, pp. 1694–1701.
[15] A. Ephremides, J. Wieselthier, and D. Baker, "A Design Concept for Reliable Mobile Radio Networks with Frequency Hopping Signaling," *IEEE Proc.*, vol. 75, no. 1, Jan. 1987, pp. 56–73.
[16] K. Makki, N. Pissinou, and O. Frieder, "Efficient Solution to Multicast Routing in Communication Networks," *J. Mobile Networks (MONET)*, vol. 1, 1996, pp. 221–32.
[17] M. H. Ammar, G. C. Polyzos, and S. K. Tripathi, "Special Issue on Network Support for Multipoint Communications," *IEEE JSAC*, vol. 15, Apr. 1997.
[18] L. Zhang *et al.*, "RSVP: A New Resource Reservation Protocol," *IEEE Net.*, Sept. 1993, pp. 8–18.
[19] C. Diot, W. Dabbous, and J. Crowcroft, "Multipoint Communication: A Survey of Protocols, Functions, and Mechanisms," *IEEE JSAC*, vol. 15, Apr. 1997, pp. 277–90.
[20] V. D. Park and M. S. Corson, "A Highly Adaptive Distributed Routing Algorithm for Mobile Wireless Networks," *Proc. IEEE INFOCOM*, 1997, pp.1405–13.
[21] S. Basagni *et al.*, "A Distance Routing Effect Algorithm for Mobility," *Proc. ACM/IEEE MOBICOM*, 1998, pp. 76–84.
[22] J. E. Wieselthier, G. D. Nguyen, and A. Ephremides, "Energy-Efficient Multicasting of Session Traffic in Bandwidth- and Transceiver-Limited Wireless Networks," to appear, *J. Cluster Comp.*
[23] J. E. Wieselthier, G. D. Nguyen, and A. Ephremides, "Energy-Efficient Broadcast and Multicast Trees in Wireless Networks," to appear, *J. Mobile Networks (MONET)*.
[24] J. E. Weiselthier, G. D. Nguyen, and A. Ephremides, "Algorithms for Energy-Efficient Multicasting in Static Ad-Hoc Wireless Networks," *J. Mobile Networks (MONET)*, vol. 6, 2001, pp. 251–63.
[25] J. E. Wieselthier, G. D. Nguyen, and A. Ephremies, "On the Construction of Energy-Efficient Broadcast and Multicast Trees in Wireless Networks," *Proc. IEEE INFOCOM*, Mar. 2000, pp. 585–94.
[26] J. Wieselthier, G. Nguyen, and A. Ephremides, "Energy Efficiency, Energy Constraints, and Admission Control in Wireless Networks," *Proc. 38th Annual Conf. Commun., Comp., and Cont.*, Urbana, IL, Oct. 2000, pp. 41–46.
[27] J. H. Chang and L. Tassiulas, "Energy Conserving Routing in Wireless Ad-Hoc Networks," *Proc. IEEE INFOCOM 2000*, Mar. 2000, pp. 22–31.
[28] G. Nguyen, J. Wieselthier, and A. Ephremides, "The Effects of Multicast Group Size and Antenna Directivity on Energy-Constrained Multicasting," *Proc. Conf. Info. Sci. and Sys.*, Johns Hopkins Univ., Mar. 2001.
[29] A. Michail and A. Ephremides, "Energy-Efficient Routing of Connection-Oriented Traffic in Ad-Hoc Wireless Networks," *Proc. 11th IEEE PIMRC*, London, U.K., Sept. 2000.
[30] A. Ephremides and B. Hajek, "Information Theory and Communication Networks: An Unconsummated Union," *IEEE Trans. Info. Theory*, vol. 44, no. 6, Oct. 1998, pp. 2416–34.
[31] P. Gupta and P.R. Kumar, "Towards on Information Theory of Large Networks: An Achievable Rate Region," *Proc. IEEE Int'l. Symp. Info. Theory*, Washington, DC, June 2001, p. 159.
[32] R. G. Gallager, "Energy Limited Channels; Coding, Multiaccess, and Spread Spectrum," *Proc. Conf. Info. Sci. Sys.*, Princeton, NJ, Mar. 1988.
[33] S. Verdu, "On Channel Capacity per Unit Cost," *IEEE Trans. Info. Theory*, vol. 36, no. 5, Sept. 1990, pp. 1019–30.
[34] P. Gupta and P. R. Kumar, "The Capacity of Wireless Networks," *IEEE Trans. Info. Theory*, vol. 46, no. 2, March 2000, pp. 388–404.
[35] M. Grossglauser, D. Tse, "Mobility Increases the Capacity of Ad-Hoc Wireless Networks," *Proc. IEEE INFOCOM*, Anchorage, AK, Apr. 2001.
[36] J. Wieselthier, G. Nguyen, and A. Ephremides, "An Insensitivity Property of Energy-Limited Wireless Networks for Session-Based Multicasting," *Proc. IEEE Int'l. Symp. Info. Theory*, Washington, DC, June 2002, p. 292.

BIOGRAPHY

ANTHONY EPHREMIDES (tony@eng.umd.edu) received his B.S. degree from the National Technical University of Athens (1967), and M.S. (1969) and Ph.D. (1971) degrees from Princeton University, all in electrical engineering. He has been at the University of Maryland since 1971, and currently holds a joint appointment as professor in the Electrical Engineering Department and the Institute of Systems Research (ISR). He is co-founder of the NASA Center for Commercial Development of Space on Hybrid and Satellite Communications Networks established in 1991 at Maryland as an offshoot of the ISR. He was a visiting professor in 1978 at the National Technical University of Athens, Greece, and in 1979 at the EECS Department of the University of California, Berkeley and INRIA, France. During 1985–1986 he was on leave at MIT and ETH in Zurich, Switzerland. He has also been the director of the Fairchild Scholars and Doctoral Fellows Program, an academic and research partnership program in satellite communications between Fairchild Industries and the University of Maryland. He has been President of the Information Theory Society of the IEEE (1987), and served on the Board of the IEEE (1989 and 1990). His interests are in the areas of communication theory, communication systems and networks, queuing systems, signal processing, and satellite communications.

Especially for the case of ad hoc wireless networks, protocol design at the networking layer (i.e., layer 3) in conjunction with transmission power choice (a choice that affects the physical and MAC layers as well) offers the possibility of substantial performance improvement.