Abstract

Transmitter power control can be used to concurrently achieve several key objectives in wireless networking, including minimizing power consumption and prolonging the battery life of mobile nodes, mitigating interference and increasing the network capacity, and maintaining the required link QoS by adapting to node movements, fluctuating interference, channel impairments, and so on. Moreover, power control can be used as a vehicle for implementing on-line several basic network operations, including admission control, channel selection and switching, and handoff control. We consider issues associated with the design of *power-sensitive* wireless network architectures, which utilize power efficiently in establishing user communication at required QoS levels. Our focus is mainly on the network layer and less on the physical one. Besides reviewing some recent developments in power control, we also formulate some general associated concepts which have wide applicability to wireless network design. A synthesis of these concepts into a framework for power-sensitive network architectures is done, based on some key justifiable points. Various important relevant issues are highlighted and discussed, as well as several directions for further research in this area. Overall, a first step is taken toward the design of power-sensitive network architectures for next-generation wireless networks.

Toward Power-Sensitive Network Architectures in Wireless Communications: Concepts, Issues, and Design Aspects

NICHOLAS BAMBOS, STANFORD UNIVERSITY

Wireless networking is rapidly becoming a major component of the modern communications infrastructure and could compete with wireline networking in terms of business volume in the near future. Such developments are mainly driven by strong market demand for personal communications systems and services. The unique aspect of wireless networking is that it provides ubiquitous, tetherless access to the users which, coupled with the high bandwidth provided by wireline networks, can offer a comprehensive solution to modern communication needs.

Contrary to the controlled environment of the wire, the wireless channel may be highly erratic and essentially stochastic; interference is an omnipresent fact one has to cope with. Historically, networking science has evolved based on the wireline network paradigm, where concepts such as routing, admission control, packet and circuit switching, and congestion management first emerged. Modern applications (multimedia) are rapidly driving wireless communication to the point where advanced networking capabilities, previously limited to the wireline world (active congestion control, guaranteed quality of service, connection reliability etc.), need to be readily available. However, they have to be supported by the highly erratic radio channel in the presence of node mobility in urban environments, where sporadic connectivity and unpredictable propagation events (shadowing, multipath fading, etc.) are the rule rather than the exception.

This dichotomy introduces a plethora of very interesting research issues that are unique to wireless networking. In this article we address a few such problems, mainly from the point of view of the network analyst and designer. Perhaps more important, we profile and characterize some additional issues that we see as critical to efficient network design and needing to be further explored by networking research. Our central theme is that of adaptive transmission power control. In this article we explore a framework for design of network architectures that are power-sensitive, as discussed in detail throughout the article.

Research supported by NSF, DARPA, and the Center for Telecommunications at Stanford University.

The Power Control Concept and Its Practical Significance

Communication link/path setup (and reconfiguration) and maintenance of the user-required quality of service (QoS) on it are key functions of network control in any networked communication environment. In wireless networking these functions are heavily dependent on transmitter power control (PC). Associated are some important optimization issues, for example, transmitter power minimization, network capacity maximization, and optimal resource allocation. Essential to PC is online link QoS monitoring for adaptation to changes induced by mobility and channel impairments.

From our point of view, PC is of fundamental importance to the operation of wireless networks for a number of reasons. In a nutshell, by adjusting its transmitter power a communication link interacts with the rest of the network and can get feedback information by monitoring the interference induced on its receiver by the other reacting links. As a result, PC can be used as a vehicle for performing several key dynamic network operations online, such as admission control, link QoS maintenance, channel probing, resource allocation, and handoffs. We articulate this position extensively later in this article, presenting a sequence of basic problems that need to be addressed in order to understand the dynamics of wireless networks and develop high-performance system architectures. A specific approach is proposed, based on some novel ideas and a few key recent results.

Our methodology for designing a power-sensitive network architecture is based on studying the PC dynamics on a canonical network model and then leveraging selected theoretical results in designing efficient algorithms and protocol suites. Optimization problems can then be formulated and analyzed on the canonical network model, reflecting performance trade-offs. The algorithms and protocols should be based on well-justified heuristics. To be practical, they must be sufficiently:

- *Distributed* — allowing autonomous execution at the node or link level, requiring minimal (if any) usage of network communication resources for control signaling
- *Simple* — suitable for real-time implementation with low strain on node/link computation resources
- *Agile* — for fast tracking of channel changes and adaptation to network stretching due to node mobility
- *Robust* — to gracefully adapt to diverse stressful contingencies, rather than stall and collapse
- *Scalable* — to maintain high performance at various network scales of interest

What can one actually expect from PC and overall power-sensitive network architectures which manage power efficiently? A moment of reflection reveals that there can be some immediate practical benefits, including:

- Increased power savings at mobile nodes for prolonged battery life (currently a major problem)
- Better maintenance of stable QoS on wireless links for supporting QoS-sensitive services (multimedia)
- More efficient handling of mobility (handoffs, etc.)
- Various others, aimed at reducing the strain on raw computation and communication resources by network control operations, hence increasing the effective capacity offered to users

Moreover, judicious use of transmission power on a link implies milder interference on other links sharing the channel, leading to increased network capacities.

Basic Paradigms of Wireless Networking: A General Conceptual Model

What is a canonical conceptual model of wireless networking for the purpose of addressing PC issues in the context sketched above? Considering modern wireless networks, as well as those realistically projected to exist in the near future, one can distinguish two basic wireless networking paradigms. Essentially, they are differentiated by the existence or lack of fixed wired infrastructure.

Ad Hoc Networking – The canonical example is that of laptop (mobile) computers equipped with radios, which establish wireless links between them and build a network topology allowing multihop connectivity. Its key characteristics are:

- There is *no fixed infrastructure* (rapidly deployable).
- There is *wireless multihop* communication, dynamically set up and reconfigurable as nodes move around.

Originally conceived for military applications (as was packet radio), it has found significant commercial applications where instant wireless infrastructure is needed (disaster relief, search and rescue operations, instant wireless LANs) and there is no fixed network (isolated areas, nonwired buildings). It is desirable that the network support multimedia traffic.

Cellular Networking – The canonical example is cellular telephony. Its key characteristics are that there is some fixed wired infrastructure, which is always accessible through a single-hop wireless link. A key notion here is that of the *access point*, where the mobile connects via the single-hop link.

At an appropriately high level of modeling abstraction one can consider the network as a collection of interfering wireless links in a channel. Actually, there may be more than one orthogonal (noninterfering) channels where links can be accommodated. In many practical systems, the channels are nonoverlapping frequency bands. Within each band, code-division multiple access (CDMA) transmission is used (either direct-sequence or frequency-hopped) with a number of different codes, taking advantage of the good performance of spread-spectrum transmission in combating interference and especially multipath fading, which is a serious problem in urban propagation environments. Some systems employ power control to mitigate the near/far effect in spread-spectrum transmission, where the intended receiver of a signal is swamped by some other interfering transmitter which happens to be quite close to it. Finally, plain time-division multiple access (TDMA), frequency-division multiple access (FDMA), or hybrid TDMA/FDMA/CDMA schemes are used in various radio communication systems.

Under the conceptual model of a wireless network being a collection of interfering links, the cellular networking paradigm is a *special case* (single-hop up/downlinks) of the far more general ad hoc networking one (wireless multihop); of course, each has its own operational parameters and constraints at lower modeling levels. This is why we prefer to think more in terms of the ad hoc networking paradigm in what follows. Most of the results, however, are directly reducible to the cellular one too, where they actually become simpler. When we need to focus on the cellular paradigm below, this is explicitly mentioned.

Actually, the above conceptual model also applies to satellite networks, as well as to networks of wireless relay stations, since they can also be viewed as collections of links between nodes (satellites, relays). One could argue that these networks form a separate paradigm where the fixed wired infrastructure is replaced by a fixed wireless one. We find the conceptual model useful for formulating and studying various problems addressed below in a unified framework.

Research Background in Power Control

The first extensive effort to bring advanced networking ideas to the wireless communication world was the packet radio project of the '80s which spawned a lot of research activity [1–5]. By then, advances in communication and information theories had provided a good understanding of the potential and limitations of point-to-point wireless communication [6], while random accessing, broadcasting, and channel sharing techniques had been developed [2, 7]. In packet radio networks, which are mainly designed to support packet-switched datagram traffic, several problems have been studied, including how to establish and maintain connectivity between various mobile nodes [8–12], and dynamically allocate slots/codes to various communication links supported at the nodes [13–22]. Power control in packet radio networks was mainly used for adjusting the transmission range to reach various receivers [23–26].

With the explosion of cellular networks in the '90s PC became very important for improving spatial channel reuse and increasing network capacity [27–42; see also Additional Reading 2, 4–6]. Early work focused on balancing the signal-to-interference ratios (SIRs) of all network users, globally lowering them as the network became congested. Recently, the focus has been on adjusting transmitter powers to maintain a required SIR threshold for each network link using the least possible power. A key algorithm for distributed PC has been introduced in [43, 44] and extended to the asynchronous case in [45]. In [46, 47] an interesting PC scheme has been introduced allowing joint PC and base station assignment to minimize powers. The fundamentals of spread spectrum transmission, as it relates to PC, have been studied in [48–50]. Additional aspects of power control (discrete/constrained power levels, bursty traffic, etc.) have been studied in [51–54]. Error-driven PC has been studied in [55–58] (also discussed later). Finally, the important problem of handoffs in cellular networks has also received considerable attention [59-63].

Adaptive Power Control for Power-Sensitive Network Architectures: Some Key Concepts and a Design Framework.

In this section we present a framework for PC which forms the backbone of a power-sensitive network architecture. It should be kept in mind that the defining goals of such an architecture are:

- To minimize power consumption and prolong battery life of mobile nodes
- To mitigate interference and increase network capacity
- To maintain link QoS by adapting to node movements and channel impairments

Moreover, by controlling its transmitter power each link can *autonomously probe* (interact with) the rest of the network and observe its *collective reaction* by monitoring the interference induced on its receiver. As a result, PC can be used to implement various basic dynamic network operations online, such as admission control, channel selection and switching, and handoff control. This unique opportunity has mostly been overlooked in the past; we will highlight it in what follows. It should actually be at the heart of any power-sensitive network architecture. Critical to the implementation of PC is the online, quick, reliable estimation of link QoS, whose variations trigger PC actions.

Our main goal here is to pinpoint justifiable design concepts and principles rather than specify particular protocols. That should be the subject of extensive further research.

The Wireless Network as a Collection of Power-Controlled Interfering Links

Our first task is to isolate and analyze the *PC dynamics* of the wireless network. A level of modeling abstraction appropriate for our purposes is to consider the network as a collection of interacting links in a channel, adopting the conceptual model discussed earlier. Links correspond to concurrent single-hop transmissions in the channel. In reality, of course, a multihop communication session may comprise several time blocks within each time frame (TDMA), existing in various frequency bands (FDMA), and using various spreading codes within each band (CDMA). Hence, the network may consist of several noninteracting (orthogonal) TDMA/FDMA channels where links may exist. We isolate an arbitrarily chosen channel to study the interaction dynamics, and reduce the problem to a single CDMA channel network [23] with near/far interfering links or a single FDMA channel with only co-channel interference (the multichannel case is examined later).

Suppose there are N links in the channel and let G_{ij} be the power gain (loss) from the transmitter of the ith link to the receiver of the jth one. It involves the free space loss, multipath fading, shadowing, and other radio wave propagation effects, as well as the spreading/processing gain of CDMA transmission [5, 48]. To keep things simple, here we assume that all Gs are deterministic (time-averaged) and do not change with time (no mobility). In reality, Gs suffer rapid stochastic fluctuations with time-varying statistics due to node mobility (we address this case later).

A measure of the link QoS (related to bit and/or packet error rates) is the SIR observed at the receiver [5, 48]. Indeed, the link throughput is an increasing function of the SIR, while the packet queuing delay is a decreasing one. Thus, the link QoS is an increasing function of the SIR overall. To simplify the discussion, we base our formulation of the power control problem directly on the SIR. Let

$$R_i = \frac{G_{ii}P_i}{\sum_{j\neq i} G_{ij}P_j + \eta_i}, \tag{1}$$

be the SIR of the ith link, $i, j \in \{1, 2, 3, \ldots, N\}$. P_i is the ith link's transmitter power, while $\eta_i > 0$ is the thermal noise power at its receiver node. For each link i there is a lower SIR threshold γ_i, reflecting a certain QoS the link has to maintain in order to operate properly. Therefore, we require that

$$R_i \geq \gamma_i, \text{ for every } i \; \{1,2,3,\ldots, N\} \tag{2}$$

in order for the links to happily coexist in the channel, satisfying their QoS demands. The previous setup connecting QoS to SIR is asymptotically accurate for the CDMA channel in the large interferer population limit (Gaussian collective interference) and quite valid in practice [5, 48]. Rewriting the inequalities of Eq. 2 in matrix form, we get $(\mathbf{I} - \mathbf{F})\mathbf{P} \geq \mathbf{u}$ with $\mathbf{P} > 0$ (component-wise), where $\mathbf{P} = (P_1, P_2, \ldots P_i, \ldots, P_N)'$ is the column vector of transmitter powers,

$$\mathbf{u} = \left(\frac{\gamma_1\eta_1}{G_{11}}, \frac{\gamma_2\eta_2}{G_{22}}, \frac{\gamma_3\eta_3}{G_{33}}, \ldots, \frac{\gamma_i\eta_i}{G_{ii}}, \ldots, \frac{\gamma_N\eta_N}{G_{NN}}\right)'$$

is the column vector of normalized noise powers, and $\mathbf{F}$ is the matrix with entries

$$F_{ij} = \frac{\gamma_i G_{ij}}{G_{ii}} \mathbf{1}_{\{i\neq j\}},$$

where $i, j \in \{1, 2, 3, \ldots, N\}$ (matrix of cross-link normalized power gains). $\mathbf{1}_{\{\}}$ is the standard indicator function, that is, $\mathbf{1}_{\{i\neq j\}}$ is 1 if $i \neq j$ and 0 otherwise.

The matrix $\mathbf{F}$ has nonnegative elements and is irreducible (since we are not considering isolated groups of links not interacting with each other). By the Perron-Frobenious theorem [45, 64], the *maximum modulus eigenvalue* ρ_F of $\mathbf{F}$ is real, positive and simple, while the corresponding eigenvector is positive component-wise. Moreover, the existence of a feasible power vector $\mathbf{P} > 0$ satisfying $(\mathbf{I} - \mathbf{F})\mathbf{P} \geq \mathbf{u}$ (hence, Eq. 2) is equivalent to $\rho_F < 1$ and also equivalent to the fact that

$$(\mathbf{I} - \mathbf{F})^{-1} = \sum_{k=0}^{\infty} \mathbf{F}^k$$

exists and is positive component-wise. The power vector

$$\mathbf{P}^* = (\mathbf{I} - \mathbf{F})^{-1}\mathbf{u} \tag{3}$$

is the *Pareto optimal* one satisfying Eq. 2, in the sense that any other $\mathbf{P} > 0$ satisfying it would require at least as much power from every transmitter [45] (i.e., $\mathbf{P} \geq \mathbf{P}^*$ component-wise). Observe that the iteration

$$\mathbf{P}(k + 1) = \mathbf{F}\mathbf{P}(k) + \mathbf{u}, \tag{4}$$

($k = 1, 2, 3, \ldots$), converges to $\mathbf{P}^* = \lim_{k\to\infty}\mathbf{P}(k)$ when $\rho\mathbf{F} < 1$, as can be seen by recursive substitution:

$$\mathbf{P}(k) = \mathbf{F}^k\mathbf{P}(0) + \left[\sum_{i=0}^{k-1}\mathbf{F}^i\right]\mathbf{u}.$$

To conclude, we observe that in principle the goal of our PC strategy should be to set powers at $\mathbf{P}^*$, when it is possible to satisfy the SIR requirements (Eq. 2) of all links simultaneously. There are, however, many other considerations which we examine below.

It should be noted that the Pareto optimal power vector $\mathbf{P}^*$ (Eq. 3), to which the algorithm converges, is basically the *minimal power* operational point for the network of links, for which the SIR (QoS) constraints are satisfied. All the algorithms discussed below also converge to a neighborhood of the Pareto optimal power vector. Therefore, they tend to minimize the power needed to support network operations given the QoS constraints. Minimizing the power is the central theme of this line of research. This is why the we call the architecture *power-sensitive*.

Distributed Power Control: The Concept of Active Link Protection

Using the iteration (Eq. 4) for distributed power control in wireless networks was first conceived by Foschini and Miljanic [43] of Bell Laboratories. Indeed, Eq. 4 can be written as $P_i(k + 1) = (\gamma_i/G_{ii})(\Sigma_{j\neq i}G_{ij}P_j(k) + \eta_i)$, or

$$P_i(k+1) = \frac{\gamma_i}{R_i(k)}P_i(k) \quad (5)$$

for every link $i \in \{1, 2, 3, \ldots, N\}$. Note that Eq. 5 can be implemented in a distributed manner, where $P_i(k)$ is the transmitter power of the ith link in the kth time slot (step). Each link measures autonomously its current SIR, $R_i(k)$, and tries to achieve its target γ_i in the next step, by boosting its power when the current SIR is below its target γ and lowering it otherwise. We call Eq. 5 the standard distributed power control (DPC) scheme; from Eq. 4 we see that it converges to the Pareto optimal P* (when that exists). Interesting extensions have been obtained by Mitra (asynchronous implementation) [45] and Yates (constrained powers, joint power control and access point assignment) [46, 47].

As pointed out by Foschini and Miljanic [44], the DPC algorithm, Eq. 5, and its extensions may allow the link SIRs to dive below γ during their evolution; hence, the transmitter powers may fluctuate erratically in the transient phase before convergence. When new links try to access the channel, established ones may inadvertently be dropped because of temporary SIR (QoS) degradation below acceptable levels, even if the new links could eventually be accommodated. If they are not admissible at all (no feasible power vector) all SIRs will degrade, while powers will escalate uncontrollably. This is a serious problem from the network admission control point of view. Indeed, in practice the channel will mostly be in a transient state, as new links try to access it all the time. It should block and reject inadmissible links, and not inadvertently drop established ones already operating in it.

To solve this problem, we have introduced [65, 66] an alternative PC scheme which protects the currently active (operational) links, maintaining their SIRs (QoS) above the thresholds γ_i at all times, as new links try to access the channel and gain admission to the network. Moreover, if the latter cannot be accommodated, they are simply suppressed and rejected, without hurting the active links in the process. The new PC algorithm works as follows, updating the powers in every step $k = 1, 2, 3, \ldots$. Let $\mathcal{L}$ be the set of all links. We define the set of *active* or *operational* links during the kth step to be $\mathcal{A}_k = \{i \in \mathcal{L}: R_i(k) \geq \gamma_i\}$, and the set of *new* or *inactive* ones (trying to gain admission by raising their SIRs above the required thresholds, but not having achieved it yet) to be $\mathcal{B}_k = \{i \in \mathcal{L}: R_i(k) < \gamma_i\} = \mathcal{L} - \mathcal{A}_k$. Finally, let $\delta > 1$ be a control parameter, arbitrarily chosen at this point. As seen later, we shall be working with δs slightly higher than 1. The new algorithm, Distributed Power Control with Active Link Protection (DPC/ALP), updates the transmitter powers as follows:

$$P_i(k+1) = \frac{\gamma_i\delta}{R_i(k)}P_i(k)\mathbf{1}_{\{i\in\mathcal{A}_k\}} + \delta P_i(k)\mathbf{1}_{\{i\in\mathcal{B}_k\}}, \quad (6)$$

or, equivalently, $P_i(k + 1) = \gamma_i\delta/G_{ii}I_i(k)\mathbf{1}_{\{i\in\mathcal{A}_k\}} + \delta P_i(k)_{\{\mathbf{1}\in\mathcal{B}_k\}}$, where $I_i(k) = \Sigma_{j\in\mathcal{A}_k\cup\mathcal{B}_k-\{i\}}G_{ij}P_j(k) + \eta_i$ is the interference (plus noise) measured at the ith link's receiver during the kth update. Note that for $i \in \mathcal{B}_k$ we have $P_i(k) = \delta^k P_i(0)$, where $P_i(0)$ is the initial power. $\mathbf{1}_{\{\}}$ is the standard indicator function.

Note that under DPC/ALP (Eq. 6), the active links play the standard DPC game with enhanced targets $\delta\gamma_i$, while the new ones power up gradually at a constant factor rate δ. DPC/ALP introduces the following new ideas into the classical PC approach:

- *Protection margin* — It artificially raises the SIR target of each active link $i \in \mathcal{A}_k$ to $\delta\gamma_i$, so as to provide a protection margin $\varepsilon = \delta - 1 > 0$ for its actual requirement γ_i. This lets active links absorb the jolts that new links induce on them by powering up, while SIRs stay above γ_i.
- *Controlled power-up* — It forces each new link $i \in \mathcal{B}_k$ to power up gradually (in a guarded manner), inducing a bounded degradation on each already active link at each step, which is absorbed by the SIR protection margin.

It can be proved that the DPC/ALP scheme has some interesting desirable properties, especially concerning network admission control. We briefly present them below (proofs can be found in [65, 66]). Specifically, for any fixed $\delta < \infty$ and any step $k \in \{0, 1, 2, 3 \ldots\}$, we have:

- *SIR protection of active links* — For any active link i we have $R_i(k) \geq \gamma_i \Rightarrow R_i(k + 1) \geq \gamma_i$; that is, $i \in \mathcal{A}_k \Rightarrow i \in \mathcal{A}_{k+1}$ or, equivalently, $\mathcal{A}_k \subseteq \mathcal{A}_{k+1}$ and $\mathcal{B}_{k+1} \subseteq \mathcal{B}_k$. Hence, once active a link never becomes inactive (new) again.
- *Bounded power overshoot* — For every active link $i \in \mathcal{A}_k$, we have $P_i(k + 1) \leq \delta P_i(k)$.
- *SIR improvement for new links* — For any new link $i \in \mathcal{B}_k$, $R_i(k) \leq R_i(k + 1)$; that is, the SIRs of new links increase with time.

It is now clear how a set of links evolves under DPC/ALP. New links power up gradually and increase their SIRs. If they ever meet their target γ they switch to active and remain so forever (or until they decide to depart). Note that these dynamics are ideally suited for admission control. We discuss this matter later.

We now look at a couple of provable facts about the DPC/ALP dynamics in two extreme cases (proofs can be found in [65, 66]), which provide some useful intuition. Let $\mathcal{A}_0 \neq \varnothing$ and $\mathcal{B}_0 \neq \varnothing$ be the sets of *initially* active and new links correspondingly. The following can then be shown:

- *Totally inadmissible new links* — If no new link ever becomes active (e.g., $\mathcal{A}_k = \mathcal{A}_0$ and $\mathcal{B}_k = \mathcal{B}_0$ for every $k \in \{1, 2, 3, \ldots\}$), then $\lim_{k\to\infty}R_i(k) = \gamma_i$ for every $i \in \mathcal{A}_0$ and $\lim_{k\to\infty}R_i(k) \leq \gamma_i$ for every $i \in \mathcal{B}_0$. Moreover, the powers of all links explode geometrically and
$$\lim_{k\to\infty}\frac{P_i(k)}{\delta_k} = \alpha_i^* < \infty.$$
- *Fully admissible links* — If there exists a feasible power vector such that all links $\mathcal{A}_0 \cup \mathcal{B}_0$ can be accommodated in the channel at SIR requirements γ_i, then under the DPC/ALP scheme all new links will become active in finite time. That is, there is a finite k_o such that $\mathcal{A}_k = \mathcal{A}_0 \cup \mathcal{B}_0$ for every $k > k_o$.

The second fact shows that if all links are compatible, all new ones become active eventually; hence, DPC/ALP does what is expected of it! Apparently it does it better than plain DPC in that it has some key desirable properties. However, there is a catch. Obviously, if all links $\mathcal{A}_0 \cup \mathcal{B}_0$ can be accommodated with enhanced SIR targets $\delta\gamma_i$ (hence, γ_i too), the powers converge to $\mathbf{P}^* = (1 - \delta\mathbf{F})^{-1}\mathbf{u} < \infty$ and the SIRs to $\delta\gamma_i$ for every $i \in \mathcal{A}_0 \cup \mathcal{B}_0$. On the contrary, if all links can be accommodated at SIR thresholds γ_i but *not* at $\delta\gamma_i$, then (although all links will become active at some finite time) their powers will explode to infinity. This is because the links are shooting for the unattainable SIR targets δ_{γ_i}. Exploding powers is an unacceptable situation, even though the links are active. Hence, we see that under DPC/ALP the channel capacity is slightly reduced compared to plain DPC. This reduction is very small for the operational region of δ in practice (1.01 to 1.10). It is clearly overcompensated by the DPC/ALP benefits, due to the three key properties. Actually, the lost capacity can be fully recaptured by dynamically relaxing $\delta \to 1$ [66].

Autonomous Online Admission Control: The Voluntary/Forced Dropout Concept

The fact that after a new link becomes active it remains so throughout its communication life naturally introduces the notion of *network admission*, occurring at the time when the link becomes active. Consider now a set of links, having evolved according to DPC/ALP for a number of steps k_o, and assume that none of the currently new links $\mathcal{B}_{k_o}$ will ever become active in the future. According to the earlier discussion, each new link's SIR will saturate below its target γ. Imagine then that we choose some new link at random and *kill* it (set its power to zero). This automatically reduces the interference on all other links, giving a second chance to some new ones to reach their targets γ, gaining network admission.

In a real network we have a dynamic environment; indeed, links arrive all the time, experience some delay until gaining network admission, transmit for some time (service time), and then depart. It turns out that it pays for a new link to *voluntarily* drop out when its objective of reaching the target γ does not seem realizable. This mitigates congestion and may help other new links gain admission. The dropout link can retry for admission later. We call this concept *voluntary dropout* (VDO). How can VDO be implemented autonomously by each new link online? Two possible strategies are the following, each based on a different idea for autonomously gauging network congestion in the link's vicinity and forming a belief about the likelihood of gaining admission soon. The key DPC/ALP property making that possible is the *SIR increasingness of new links*. When a new link i arrives to the network (appears initially as a request for establishing a single-hop connection between two nodes), it starts its local clock (counting steps k), begins transmitting at a very low power $P_i(0)$ at $k = 0$ (low enough so as not to kill any active link[1] in its vicinity, e.g.,, of the order of the thermal noise power), and powers up according to the DPC/ALP scheme, implementing either of the following two strategies:

- *Timeout-based VDO* — The new link initially sets a *timeout horizon* T. If it has not gained admission by time T, it computes a *dropout horizon* D as a decreasing function of $(\gamma - R(T))$, for example, $D = \lfloor Ae^{-\alpha(\gamma - R(T))} \rfloor$ where $A, \alpha > 0$, and $\lfloor\,\rfloor$ is the lower integer part. The link keeps trying for admission until time $T + D$. Note that links closer to their SIR target tend to try longer. Throughout D the new link hopes that some other link dropping out or naturally departing might boost its chance of admission. If that has not occurred by $T + D$, the link voluntarily drops out, setting its power to zero.
- *SIR-saturation-based VDO* — The new link retains M-step memory of its previous SIR values at all times; it also sets a threshold ΔR of significant anticipated SIR improvement in the last M steps. At the beginning of each step k, the link checks whether significant SIR improvement has occurred in the last M steps (e.g., $R(k) - R(k - M) \geq \Delta R$). If so, it simply updates its power according to DPC/ALP. If not, it thinks that its SIR has saturated due to network congestion and admission is currently impossible; hence, it attempts to drop out for a while. To implement that, it flips a coin which gives 1 with probability p^{drop} which is an increasing function of $(\gamma - R(k))$, for example, $p^{drop} = 1 - e^{-\beta(\gamma - R(k))}$. If the outcome is 1 the link drops out immediately; if not, it updates its power according to DPC/ALP and starts the process all over. Note that the closer the link to its target the less likely it is to drop out.

After a new link drops out (backs off) it lies dormant for some time B and then retries for admission by starting power-up once again from the initial lowest power $P_i(0)$. Observe that the saturated SIR of a new link may suddenly climb significantly due to the dropping out of a neighboring new link or the natural departure of an active one. If the memory window M is too narrow, the link may react to a short-lived lack of SIR improvement by dropping out unnecessarily. On the contrary, if the window is too wide the link may persist too long before dropping out and congest the network excessively. Clearly M is a design parameter that needs to be optimized, as are the ΔR, β, T, D, A, α, B, for which analogous considerations hold. Actually, T, D, B, and so on may even be random, say, exponentially distributed.

---

[1] *Regarding the initial power level $P(0)$ (from which new users start powering up when trying to access the channel) the question is how large can that be so as not to kill active links. For simplicity, let us assume that all links have the same γ and the same thermal noise power η. The network operates under DPC/ALP with some δ, so the protection margin is $(\delta - 1)\gamma$. Assume now that there are some active users in the channel, which have converged to their enhanced SIRs $\delta\gamma$ (recall that convergence is geometrically fast), and a new link starts powering up from power level $P(0)$. Let G_{max} be the maximum power gain G between the transmitter of the new link and all the receivers of the active ones. That is, G_{max} is the minimum power attenuation from the new link to the active ones, or the maximum interference/interaction strength between new and active links. The active links being at their enhanced SIRs $\delta\gamma$ and enjoying a protection margin $(\delta - 1)\gamma$, can tolerate an initial power (as simple calculations show)*

$$P(0) < \frac{(\delta - 1)\eta}{G_{\max}}$$

without any of them dropping below the SIR threshold γ and becoming inactive. That way, the new link enters the channel smoothly and without killing any active one. (If we have at most N_{max} new links starting power-up simultaneously, the denominator in the previous formula becomes $N_{max}G_{max}$.) The issue now is to estimate an upper bound on the G_{max}. Consider first the cellular network paradigm, where downlinks are set up in a different channel than uplinks, so they do not interact. Assuming that the base station is at some height, say 30 m, from the ground where the mobile users move, we have G_{max} of the order 30^{-4} (assuming that the power attenuation follows an inverse-fourth-power law, as is typically the case in urban propagation environments). Taking $\delta = 1.1$ we get roughly $P(0) < 10^5\,\eta$. The point is that for any application scenario we can compute a G_{max} (this is typically much less than 1), and get an upper bound on $P(0)$. A very conservative estimate of a power from which to start is just the thermal noise one, $P(0) = \eta$. In simulations we observe that one can be much more aggressive with the startup power without seeing any active link diving below γ and being killed.

Finally, note that each link acts autonomously, so the whole admission/rejection process is distributed. Actually, each link may set its parameters individually.

In the previous discussions, we have always assumed that the transmitter powers are unlimited. In reality, however, there is a power ceiling P^{max} that no transmitter may exceed (actually, each transmitter may have a different power ceiling). As a new link is powering up, active ones respond by also raising their powers according to DPC/ALP; hence, some active link i may be pushed against the power ceiling. Since it cannot exceed P^{max}, it will see its SIR dropping below its threshold γ as the new link keeps ramping up its power; hence, it will cease being active (QoS dives) and may be dropped. This would defeat the primary purpose of DPC/ALP, which is to protect the QoS of active links at all times. Fortunately, this can be remedied due to the second DPC/ALP property of *bounded power overshoots*, which ensure that $P_i(k + 1) \leq \delta P_i(k)$ for every active link $i \in \mathcal{A}_k$. Therefore, as long as $0 \leq P_i(k) \leq P^{max}/\delta$, the link knows that $P_i(k + 1) \leq P^{max}$ (i.e., it can stay active in the next step). However, the critical case is the following:

Forced dropout (FDO) under bounded powers — When an active link's power enters the zone $P^{max}/\delta < P_i(k) < P^{max}$, the link realizes that it may not survive in the next step, because it may not have enough power to react. At this point it transmits a *distress signal* announcing its situation. All new links receiving it (above some power level) drop out immediately. These are the new links in the vicinity of the active one, which are precisely those stressing the latter.

Observe that, since in practice δ is slightly higher than 1, the interval $[P^{max}/\delta, P^{max}]$ is quite slim and provides a zone of alert (red zone). The bounded power overshoot property of DPC/ALP assures us that an active link can go from $[0, P^{max}/\delta]$ to $[P^{max}, \infty)$ only by crossing $[P^{max}/\delta, P^{max}]$, where the distress signal is transmitted. On the contrary, under the plain DPC, a link's power can jump directly from $[0, P^{max}]$ to $[P^{max}, \infty)$ and the link cannot autonomously detect when its SIR is about to drop below its threshold γ.

Actually, from simulation experiments [66] we see that when VDO is used the aggregate network throughput increases more than 10 times (!) compared to when no dropout is allowed; this highlights the importance of VDO in mitigating congestion by diffusing it temporarily. Also, from simulations of a wide variety of networking scenarios [66] we find that in practice δ should slightly larger than 1 (between 1.01 and 1.10, providing 1–10 percent protection margin) in order to minimize delay and maximize throughput. We discuss some ideas for dynamically choosing δ later.

Quick Noninvasive Channel Probing and Monitoring: The Probing Concept

In the previous discussion new links drop out only after they have tried aggressively to gain admission and failed. In the process, they congest the channel by raising the interference level, hence, reduce the effective channel capacity. Fortunately, it turns out that the DPC/ALP scheme can be used to autonomously ***probe the channel*** and get a reasonable prediction as to whether a link is admissible in a quick and fairly noninvasive manner. This is based on the following fact, which can be shown for the DPC/ALP dynamics [67].

Suppose we have a set of active links $\mathcal{A}_0$ on the network at time $k = 0$ and a single new link, which has just started powering up in the channel (according to the DPC/ALP scheme) with initial power $P(0)$. We assume that:

- The channel can accommodate every initially active link $i \in \mathcal{A}_0$ at SIR threshold $d\gamma_i$.
- Initially, the SIRs of all active links $i \in \mathcal{A}_0$ are such that $\gamma_i \leq R_i(0) \leq \delta\gamma_i$ (i.e., powers and SIRs have previously stabilized).

Then, while the new link remains inactive, the following bounds can be established for the evolution of SIR values $R(k)$:

$$\frac{1}{\frac{X(1)}{\delta^k}+Y(1)} \leq R_i(k) \leq \frac{1}{\frac{X(\delta)}{\delta^k}+Y(\delta)} \tag{7}$$

for every $k \in \{0, 1, 2, ...\}$(for which admission has not yet been achieved). The parameters $X(\delta)$, $Y(\delta)$ can be explicitly computed [67], and turn out to have the following nice property:

$$\lim_{\delta\to 1} X(\delta) = X(1) = X^* \quad \text{and} \quad \lim_{\delta\to 1} Y(\delta) = Y(1) = Y^*. \tag{8}$$

Note that as $\delta \to 1$ (recall that $\delta > 1$), the lower and upper bounds in Eq. 7 come closer and eventually coincide. Actually, this behavior extends to the case of multiple new links; each has its own X^*, Y^* then. The above are mathematically provable facts [67].

Motivated by the previous results, and looking into their engineering meaning, we see that for $\delta \approx 1$ we can write for the SIR of the new link

$$R(k) \approx \frac{1}{\frac{X^*}{\delta^k}+Y^*} \tag{9}$$

for every step k during which the link has not yet become active. Note that the necessary and sufficient condition that the new link never become active (is inadmissible) is simply $\lim_{k\to\infty} R(k) = 1/Y^* \leq \gamma$ (recall that $\delta > 1$ and the SIRs of new links always increase under DPC/ALP). On the contrary, if $\gamma < 1/Y^*$, it is easy to see (solving $R(k^*) \approx \gamma$ for k) that the link will be admitted after $k^* \approx (\log\delta)^{-1}(\log X^* - \log(\gamma^{-1} - Y^*))$ steps, at power $P^* \approx P_o\delta^{k^*}$. (10)

The interesting thing is that the new link can predict its future evolution (hence, whether it is admissible or not) by autonomously estimating X^* and Y^* from the first two power-up steps only. Indeed,

$$X^* \approx \frac{\delta}{\delta-1}\left[\frac{1}{R(0)}-\frac{1}{R(1)}\right] \quad \text{and} \quad Y^* \approx \frac{1}{\delta-1}\left[\frac{\delta}{R(1)}-\frac{1}{R(0)}\right] \tag{11}$$

Since the first couple of power-up steps happen at very low powers, $P(0)$, $P(0)\delta$, this probing process minimally upsets the other links in the channel. If the link determines (predicts) that it is inadmissible it drops out immediately so as not to stress the active links unnecessarily. Hence, the whole probing process is quick, rather noninvasive, and obviously fully distributed and autonomous.

The previous results extend naturally to the case where there are several new links trying for admission. Each new link can again predict in two steps its SIR evolution in the channel, assuming that no new link drops out. However, a link which predicts that is currently not admissible may eventually be admissible if some other new link drops out. Hence, to drop out new links should use some randomized mechanism, for example, flip a coin having probability of kicking them out which is an increasing function of $\gamma - 1/Y^*$.

We hasten to add that the above are true under the special setup introduced by the assumptions made in Eq. 7. However, because of the simplicity of the probing, one wonders whether it could be applied to the actual dynamic network situation to performance. Indeed, we have checked this out by simulation, and some preliminary results have been very encouraging [67]. In some cases, the throughput increases almost 40 percent over the timeout-based VDO. This is due to the noninvasiveness of the process, which induces minimal congestion. Therefore, probing seems to have a large practical potential, and could be a key aspect of power-sensitive network architectures.

Associated Research Issues for Power-Sensitive Architectures

Presented above is a justifiable approach to PC and the design of wireless network architectures, rather than a specific technology. A few salient concepts emerge, as well as some associated design principles. In a nutshell, it pays to:

- *Protect active links* in order to suppress fluctuations of QoS and inadvertent dropping of ongoing calls, as well as smooth out the transient network dynamics (triggered by random call arrivals)
- *Voluntarily drop out* when admission does not seem feasible and stressing is detected; this streamlines the link interactions and ultimately increases network capacity and reduces admission delay
- *Probe noninvasively* and be aware of your channel and resource environment to support autonomous decision making; interesting things to know include what resources are out there (access points, free bandwidth) and at what cost these are available (power needed, etc.)

The DPC/ALP scheme can support these functionalities autonomously at the link level. The SIR values measured at the link receiver, which are needed for PC at the transmitter, can be either transmitted back on some separate control channel (minimal control traffic) or appended to the message acknowledgments and so forth. Actually, the basic intuition behind the DPC/ALP scheme implies that it can be implemented in practice, not necessarily using the SIR as a link QoS measure, but some other QoS measure $Q_i(P_1, P_2, P_3, ..., P_i, ..., P_N)$ for links i. The key point is that Q_i should be increasing in P_i and decreasing in P_j for every $j \neq i$. For example, we could take as a QoS measure the bit/packet/message error rate observed at the link transmitter.

In this section we focus on highlighting some key research issues associated with PC which need to be further researched and understood in order to develop and implement efficient power-sensitive network architectures. One can think of this section as a blueprint or plan for further research in PC. Some novel ideas are introduced, some fundamental performance tradeoffs exposed, and certain solution approaches to the emerging problems are proposed.

Online Adaptation of DPC/ALP to Congestion

How does one choose the power-up factor δ so as to optimize DPC/ALP performance? Simulation of the network scenario at hand is not quite satisfactory because it is highly case-dependent. Actually, δ could be dynamically estimated and should adapt to network congestion instead of being constant. Here is the intuition. There are a few measures of congestion that can be observed online, such average admission delay, average power of active links, and PC relaxation dynamics. The quantity δ should adapt to channel congestion (evaluated online over intervals long enough for reliable estimation). The idea is that when congestion decreases δ should increase, making DPC/ALP more agile in converging to enhanced QoS targets $\delta\gamma$ (feasible due to low congestion). On the contrary, when congestion builds up δ should decrease, making DPC/ALP power up more gently and converge to $\delta\gamma$ (very close now to γ; cannot be generous under high congestion), becoming sluggish (just like plain DPC). This should be done in conjunction with an online dynamic search (descent algorithm) to obtain the minimum congestion δ. Updates of δ can be broadcast on a separate control channel monitored by the links in a mobile computing network or through the wired infrastructure in cellular networks. These ideas should be explored more for online optimization of DPC/ALP with respect to its operational parameters.

The Multichannel Case: Channel Selection and Switching

We have always assumed above that there is only one channel. In reality, however, there may be several noninterfering (orthogonal) channels (TDMA/FDMA), each corresponding to a group of time slots in some frequency band(s). For CDMA channels a spreading code has to be allocated to each user. Each node can tune into any of the channels/codes to transmit or receive. A link of some SIR (QoS) target γ may be admissible in several channels at *different powers*, depending on other coexisting links. Upon arrival, each new link can probe several channels (randomly selected or round-robin) and choose to access that where it can be admitted at the lowest power (thus the fastest). The speed (two steps) and noninvasiveness of probing makes it possible to examine a significant number of channels without substantially adding to the admission delay or overall congestion.

A link initially admitted to some channel can intermittently probe alternative ones, hunting for another where it can achieve its target QoS at significantly lower power (because congestion has relaxed in that other channel since its admission due to link departures). If it succeeds, it switches to the lowest-power channel (following a smooth transition protocol), trying to dynamically balance the load (congestion) among the channels.

In preliminary simulation studies of probing-based channel selection and switching [67] (only two channels) we have observed up to 55 percent reduction of delays (normalized) and 40 percent of average power. The potential and particulars of probing-based channel selection and switching should be studied more on practical designs for realistic networking scenarios.

The Minimum-Power Routing Problem in Multihop Wireless Networking

Until now, we have been talking about single-hop links. In the ad hoc networking paradigm, however, the network supports multihop connections on source-destination paths, comprising many links of certain QoS distributed over various channels. Given a set of required source-destination connections of specified QoS, the global minimum routing problem is to choose the links, and the channels in which to establish each one (and its spreading code for CDMA channels), to either minimize the total power needed to support the connections on the network (version 1) or minimize the maximum power spent at any node (i.e., maximize the time until some node dies of battery exhaustion) (version 2). This is an intractable combinatorial optimization problem for any sizable network; hence, one has to search for justifiable heuristics to obtain good practical solutions. There is some good intuition to guide us. For example, the channel/code reuse distance should be as high as possible, and channels should be load-balanced with respect to congestion.

A more tractable and practical case is the individual minimum-power routing problem. The objective is to chose the links/channels/codes to route a newly arrived connection request, minimizing the total power spent on its links (or the maximum spent on any of them) without reshuffling already established links of previously arrived connections. By each node probing the channels to obtain power estimates, the problem reduces to finding the shortest path from the source to the destination, where the arc lengths are the powers needed for establishing the links. Given the importance of power management in mobile networking, it is worth devoting significant effort searching for practical solutions to cope with the inherent complexities.

Node Mobility, Network Stretching and Reconfiguration, and Probing-Based Handoffs

In all previous discussions we have assumed that the network nodes are stationary. In reality, however, several nodes are mobile and roaming around. PC updates the transmitter powers adapting to changing power gains G_{ij} (due to node mobility), trying to maintain the link QoS requirements and preserve the network topology (links/paths) under structural stretching (*elastic* network). As a result, some powers may increase excessively and links become too expensive to maintain. Actually, there may be some other link/channel/code configuration supporting the same end-to-end connections at much lower power. At this point, some links may need to be reconfigured to lower their power. In the *special cellular* network paradigm, where all links are single-hop, this is simply the base station (access point) handoff problem, which can be handled by probing. Indeed, while transmitting to (or receiving from) a base station, a mobile periodically probes several links/channels/codes and estimates the power needed in each case to achieve its target QoS, registering the best alternatives. When its current connection degrades significantly (spends too much power) the mobile switches to its best registered alternative. Quick noninvasive probing makes this probing-based handoff control scheme practical. Of course, in order for PC to be agile and quickly adaptable to QoS changes due to mobility, the *PC relaxation timescale* (time for PC algorithm to freely stabilize) should be significantly smaller than the *mobility timescale* (time in which enough link stretching has occurred to cause significant detectable QoS change). This is another issue that must be further explored, especially with respect to the potential of probing-based handoff control for handling mobility in wireless networking.

The Stochastic Basis of Power Control and Quick Online Estimation of Link Quality

Any adaptive PC scheme basically reacts to link QoS changes trying to maintain the QoS targets. In order to be agile and respond quickly to changes (node mobility, channel impairments) it is essential that these are detected as early as possible after they occur. The problem is that QoS is a statistical quantity estimated online via sequential sampling of some stochastic process (e.g., SIR). Indeed, in reality power gains G_{ij} are random processes fluctuating in time, as are also received power, interference, and SIR. PC should not respond to random fluctuations of SIR in time, but only to average trends reflecting true QoS changes. Identifying trends in the presence of random fluctuations is an inherently slow process involving time averaging. This is a key problem in estimating QoS in terms of bit/packet error rates, time-averaged SIRs, and so on. To allow for reliable QoS estimation the *PC updating timescale* (time between consecutive power updates) should be significantly larger than the *transmission timescale* (time to transmit a bit/packet/message). Traditionally, the QoS estimation problem (SIRs, bit/packet error rates) has been treated within the framework of sequential analysis of stationary processes [68, 69], assuming that within an estimation interval the statistics do not change much (are stationary). Within this framework there is a rich literature on channel estimation in communication systems. Another way is to take a more fundamental approach, using the framework of the quickest detection problem introduced by Shiryayev [70–74]. Briefly speaking, we sample a stochastic sequence X_1, X_2, X_3, ..., X_n, ... At some random time T a structural change occurs, so $X_1, X_2, X_3, ..., X_T$ are distributed according to $F_1(x)$, but $X_{T+1}, X_{T+2}, X_{T+3}$, ... according to $F_2(x)$. For example, Xs could be the SIRs and T the random time when the mobile turns around a corner (15–25dB drop in power gain). We want to detect the change as soon as possible after it has occurred. A solution can be obtained under certain conditions [70, 71, 74] on the statistics of T. To analyze our problem á la Shiryayev, we need to use the theory of Optimal Stopping Rules [75], adapting it to QoS estimation. To the best of our knowledge this approach has not been tried before; we expect it to give some useful results.

The Power Manager's Dilemma: To Transmit or Wait?

There is a fundamentally different twist to the PC problem which we have to consider. Of considerable theoretical interest, it is also of great practical importance in maximizing the mobile's battery life.

Consider a single communication link operating under stochastic interference with stationary statistics. Time is slotted. Let:

- I_n be the interference seen at the receiver during time slot n
- P_n the power the transmitter uses to transmit a message (bit, packet, etc., depending on the slot definition) during time slot n
- $s(p, i)$ the probability the message is received successfully during the current time slot, given that it is transmitted at power $P_n = p$ and the channel interference is $I_n = i$

The function $s(p, i)$ is increasing in p and decreasing in i. Messages arrive at the transmitter according to a Bernoulli process of average rate $\bar{r}$ (i.e., at most one message in each time slot with probability $\bar{r}$) and are queued up in a FIFO queue with infinite buffer, waiting to be successfully transmitted across the link to the receiver. The problem is to design a power selection (control) policy to:

- Choose the transmitter power P_n during time slot n
- Based on past and present interference values $\{I_k, k \leq n\}$ and past chosen power levels $\{P_k, k < n\}$ (the interference is measured at the beginning of the time slot right before the power is chosen)
- So as to minimize the average power consumption

$$\bar{P} = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} Pn$$

- While keeping the transmitter queue stable with average message delay less than $\bar{d}$

given 1) the statistics of $\{I_n\}$ (e.g., a Markov chain switching between various interference levels) and 2) the success probability $s(p, i)$ (reflecting the channel model). Hence, the objective is to maintain the QoS, as expressed by throughput $\bar{r}$ and delay $\bar{d}$, using the minimum possible average power.

For example, consider the simplest possible class of policies, where a threshold I^* is chosen and the link transmits at power P^* (always the same) if $I_n < I^*$, and idles otherwise. Intuitively, we expect it to idle during interference highs, waiting to transmit when the lows appear using lower power; however, too much waiting increases the average message delay and can cause the queue to explode. Surprisingly, this problem has been little studied in the past, despite its obvious importance for power savings. We have recently looked at some of its elementary aspects and have seen that even with simple policies we can achieve substantial power savings [76, 77; see also

Additional Reading 8, 9]. It is clear that much more needs to be done in this research direction.

Error-Driven Power Management

Another interesting approach to power management at the link level has been pursued by Zorzi and Rao [55–58], focusing on error control under given fading statistics of the wireless channel. The objective here is to maximize the mean amount of data delivered per unit of energy consumed. Modeling the channel fading statistics as having Markovian structure, they have analyzed extended automatic repeat request (ARQ) schemes which tend to cease transmitting when they sense that the channel is bad so as not to spend power in vain when the chances of successful reception are slim. To do that, when the transmitter does not receive successful reception acknowledgments for a while it infers that the channel is bad and switches to a dormant "probing mode." In that mode it intermittently probes the state of the channel by transmitting short test messages which spend little power and also add little to congestion; by the acknowledgment response the transmitter infers the channel state. As long as the channel is bad (deep fading) the transmitter stays in the probing mode. As soon as it estimates that the channel is again up it switches back to the normal ARQ transmission mode. The cycle is repeated indefinitely as long as there are data (packets) to be sent. This approach is somewhat related to that presented in the previous subsection, but there are some significant differences regarding the optimization goals and metrics. It is interesting to note, however, that preliminary results under both approaches seem to agree quite well and be compatible. This points to the fact that there is an opportunity here for multiobjective optimization, which needs to be investigated much more deeply since it has the potential to provide significant power savings at the transmission link level.

Power-Sensitive Wireless Network Architectures

The previously discussed concepts snd ideas are essential elements of a power-sensitive network architecture as described in this article. There are, however, many other important related considerations, including:

- Network startup and synchronization (after being turned on, how nodes find out about each other and self-organize into a network, maintaining some acceptable level of synchronization)
- Node-bandwidth allocation (how to dynamically allocate time slots at nodes to various communication sessions given the traffic demands, CDMA signal acquisition times, and successful reception probabilities)
- Routing and path reconfiguration (how to choose connection paths through the network and reconfigure them when necessary)

In the limit of very high mobility (speed), network synchronization and organization are very difficult to maintain and communication is better supported on an asynchronous random-access (Aloha-like) basis. Under reasonable mobility, however, we can set up a synchronous network structure to reap several benefits related to throughput, QoS, PC, network intelligence, and so forth. The detailed design and fine-tuning of an efficient network architecture is of course a far-from-trivial matter [Additional Reading 1, 3, 7]. It would heavily depend on the networking scenario at hand (ad hoc, cellular, high/low mobility, etc.). Discussing specific architectures is beyond the scope of this article; its objective is to lay out some general concepts and coherent principles rather than address specific scenarios. The latter is the subject of intense further research, some of which is currently underway.

Concluding Remarks

We have identified a number of key concepts and issues concerning the potential of adaptive power control in wireless networking. A synthesis of these concepts has led to a justified framework for power-sensitive architectures based on the DPC/ALP/VDO/FDO/Probing suite of algorithms. Important related issues of quick QoS estimation, power conservation, and so on have also been highlighted. This is a first step toward designing power-sensitive wireless network architectures. Much further research is needed to reach good technological solutions for power management in the next-generation wireless networks.

Acknowledgments

We would like to thank Prof. L. Kleinrock, Prof. M. Gerla, and Prof. G. Pottie of UCLA, Prof. D. Cox of Stanford, and Dr. D. Mitra of Lucent for useful discussions on the subject of power control and network design. The support of this research by DARPA, NSF, the Center for Telecommunications at Stanford, and Silicon Valley companies is gratefully acknowledged.

References

[1] A. Ephremides, J. E. Wieselthier, and D. J. Baker, "A design concept for reliable mobile radio networks with frequency hopping signaling," *Proc. IEEE*, vol. 75, no. 1, 1987.

[2] A. Ephremides and S. Verdu, "Control and optimization methods in communication network problems," *IEEE Trans. Auto. Contr.*, vol. 34, no. 9, 1992.

[3] M. M. Leiner, D. L. Nielson, and F. A. Tobagi, "Issues in packet radio network design," *Proc. IEEE*, vol. 75, no. 1, 1987.

[4] L. Kleinrock and J. Silvester, "Spatial reuse in multihop Packet radio networks," *Proc. IEEE*, vol. 75, no. 1, 1987.

[5] M. B. Pursley, "The role of spread spectrum in packet radio networks," *Proc. IEEE*, vol. 75, no. 1, 1987.

[6] A. J. Viterbi and Jim K. Omura, *Principles of Digital Communication and Coding*, McGraw-Hill, 1979.

[7] R. Rom and M. Sidi, *Multiple Access Protocols: Performance and Analysis*, Springer, 1990.

[8] D. J. Baker and A. Ephremides, "The architectural organization of a mobile radio network via a distributed algorithm," *IEEE Trans. Commun.*, vol. 29, no. 11, 1981.

[9] D. J. Baker, "Distributed control of broadcast radio networks with changing topologies," *Proc. IEEE INFOCOM*, 1983.

[10] E. S. Sousa and J. A. Silvester, "Spreading code protocols for distributed spread-spectrum packet radio networks," *IEEE Trans. Commun.*, vol. 36, no. 3, 1988.

[11] L. Hu, "Topology control for multihop packet radio networks," *IEEE Trans. Commun.*, vol. 41, no. 10, 1993.

[12] I. Cidon and M. Sidi, "Distributed assignment algorithms for multi-hop packet radio networks," *IEEE Trans. Commun.*, vol. 38, no. 10, 1989.

[13] L. Tassioulas and A. Ephremides, "Stability properties of constrained queuing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE Trans. Auto. Cont.*, vol. 37, no. 2, 1992.

[14] L. Tassiulas and A. Ephremides, "Dynamic server allocation to parallel queues with randomly varying connectivity," *IEEE Trans. Info. Theory*, vol. 39, 1993.

[15] L. Tassioulas and A. Ephremides, "Jointly optimal routing and scheduling in packet radio networks," *IEEE Trans. Info. Theory*, vol. 38, no. 1, 1992.

[16] L. Hu, "Distributed code assignments for CDMA packet radio networks," *IEEE/ACM Trans. Networking*, vol. 1, no. 6, 1993.

[17] R. Nelson and L. Kleinrock, "Spatial TDMA: A collision free multihop packet channel access protocol," *IEEE Trans. Commun.*, vol. 33, no. 9, 1985.

[18] J. A. Silvester and L. Kleinrock, "On the capacity of multihop slotted ALOHA networks with regular structure," *IEEE Trans. Commun.*, vol. 31, no. 8, 1983.

[19] B. Hajek and G. Sasaki, "Link scheduling in polynomial time," *Trans. Info. Theory*, vol. 34, 1988.

[20] J. A. Silvester and L. Kleinrock, "On the capacity of single-hop slotted ALOHA networks for various traffic matrices and transmission strategies," *IEEE Trans. Commun.*, vol. 31, no. 8, 1983.

[21] K. Meier-Hellstern, G. P. Pollini, and D. J. Goodman, "Network protocols for the cellular packet switch," *IEEE Trans. Commun.*, vol. 42, nos.2/3/4, 1994.

[22] I. Chlamtac and A. Farago, "Making transmission schedules immune to topology changes in multi-hop packet radio networks, *IEEE/ACM Trans. Networking*, vol. 2, no. 1, 1994.

[23] T. C. Hou and V.O.K. Li, "Transmission range control in multihop packet radio networks, *IEEE Trans. Commun.*, vol. 34, no. 1, 1986.
[24] E. S. Sousa and J. A. Silvester, "Optimum transmission ranges in a direct-sequence spread-spectrum multihop packet radio network," *IEEE JSAC*, vol. 8, no. 5, 1990.
[25] H. Takagi and L. Kleinrock, "Optimal transmission ranges for randomly distributed packet radio terminals, *IEEE Trans. Commun.*, vol. 32, no. 1, 1984.
[26] M. Zorzi and S. Pupolin, "Optimum Transmission ranges in multi-hop packet radio networks in the presence of fading," *IEEE Trans. Commun.*, vol. 43, pp. 2201–5, 1995.
[27] J. Zander, "Performance of optimum transmitter power control in cellular radio systems," *IEEE Trans. Vehic. Tech.*, vol. 41, no. 1, 1992.
[28] J. Zander, "Distributed cochannel interference control in cellular radio systems," *IEEE Tran. Vehic. Tech.*, vol. 41, no. 3, 1992.
[29] J. Zander, "Transmitter power control for co-channel interference management in cellular radio systems," *Proc. 4th WINLAB Wksp. Third Generation Wireless Info. Networks*, 1993.
[30] A. M. Viterbi and A. J. Viterbi, "Erlang capacity of a power-controlled CDMA system," *IEEE JSAC*, vol. 11, no. 6, 1993.
[31] J. C. I. Chuang and N. R. Sollenberger, "Performance of autonomous dynamic channel assignment and power control for TDMA/FDMA wireless access," *IEEE JSAC*, vol. 12, no. 8, 1994.
[32] H. Alavi and Ray W. Nettleton, "Downstream power control for a spread spectrum cellular mobile radio system," *Proc. IEEE GLOBECOM '82*, 1982.
[33] S. Ariyavisitakul, "SIR-based power control in a CDMA system," *Proc. GLOBECOM '92*.
[34] T. Fujii and M. Sakamoto, "Reduction of cochannel interference in cellular systems by intra-zone channel reassignment and adaptive transmitter power control," *Proc. IEEE VTC*, 1988.
[35] K. S. Gilhousen *et al.*, "On the capacity of a cellular CDMA system," *IEEE Tran. Vehic. Tech.*, vol. 40, no. 2, 1991.
[36] W. C. Y. Lee, "Power Control in CDMA," *Proc. IEEE VTC*, 1991.
[37] R. W. Nettleton, "Traffic theory and interference management for a spread spectrum cellular mobile radio system," *Proc. ICC '80*, Seattle, WA, 1980.
[38] R. W. Nettleton and H. Alavi, "Power control for a spread-spectrum cellular mobile radio system," *Proc. IEEE VTC*, 1983.
[39] M. Almgren, H. Andersson, and K. Wallstedt, "Power control in a cellular system." *Proc. IEEE VTC*, 1994.
[40] S. Grandhi et al., "Centralized power control for cellular radio systems," *IEEE Trans. Vehic. Tech.*, vol. 42, no. 4, 1993.
[41] S. Grandhi and J. Zander, "Constrained power control in cellular radio systems," *Proc IEEE VTC*, 1994.
[42] S. A. Grandhi, R. Vijayan, and D. J. Goodman, "A distributed algorithm for power control in cellular radio systems," *Proc. Allerton Conf. Commun., Control and Comp.*, 1992.
[43] G. J. Foschini and Z. Miljanic, "A simple distributed autonomous power control algorithm and its convergence, *IEEE Trans. Vehic. Tech.*, vol. 42, no. 4, 1993.
[44] G. J. Foschini and Z. Miljanic, "Distributed autonomous wireless channel assignment with power control," preprint, 1994.
[45] D. Mitra, "An asynchronous distributed algorithm for power control in cellular radio systems," *Proc. 4th Winlab Wksp. Third Generation Wireless Info. Network*, Rutgers Univ., 1993.
[46] R. Yates and C. Y. Huang, "Integrated power control and base station assignment," *IEEE Trans. Vehic. Tech.*, vol. 44, no. 3, 1995.
[47] R. Yates, "A framework for uplink power control in cellular radio systems," *IEEE JSAC*, vol. 13, no. 7, 1995.
[48] S. V. Hanly, "Information capacity of radio networks," Ph.D. thesis, University of Cambridge, U.K., 1993.
[49] S. V. Hanly, "Information theoretic capacity of multi-receiver networks," *Telecommun. Sys.*, vol. 1, no. 1, 1993.
[50] S. Hanly, "Capacity and power control in a spread spectrum macro diversity radio network," Bell Labs tech. rep.
[51] M. Andersin, Z. Rosberg, and J. Zander, "Soft admission in cellular PCS with constrained power control and noise," Tech. rep. TRITA-IT R 94-21, Radio Commun. Sys. Lab, Royal Inst. Technology, Sweden, 1995.
[52] M. Andersin, Z. Rosberg, and J. Zander, "Gradual removals in cellular PCS with constrained power control and noise," Tech. rep. TRITA-IT R 94-22, Radio Commun. Sys. Lab, Royal Inst Technology, Sweden, 1995.
[53] D. Mitra and J. A. Morrison, "A distributed power control algorithm for bursty transmissions in cellular, spread spectrum wireless networks," *Proc. 5th Winlab Wksp. Wireless Info. Networks*, Rutgers Univ., 1996.
[54] D. Mitra and J. A. Morrison, "A novel distributed power control algorithm for classes of service in cellular CDMA networks," *Proc. 6th Winlab Wksp. Wireless Info. Networks*, Rutgers Univ., 1997.
[55] M. Zorzi and R. Rao, "Error control and energy consumption in communications for nomadic computing," *IEEE Trans. Comp.*, 1997.
[56] M. Zorzi and R. Rao, "Energy constrained error control for wireless channels," *Proc. GLOBECOM '96*, London, U.K., 1996.
[57] M. Zorzi and R. Rao, "Energy management in wireless communication," *Proc. 6th Winlab Wksp. Wireless Info. Networks*, Rutgers Univ., 1997.
[58] M. Zorzi and R. Rao, "Energy constraints error control for wireless channels," preprint, 1996.
[59] P. S. Kumar and J. M. Holtzman, "Analysis of handoff algorithms using both bit error rate and relative signal strengths," *Proc. ICUPC*, 1994.
[60] A. Sampath and J. M. Holtzman, "Adaptive handoffs through the estimation of fading parameters," *Proc. Int'l. Conf. Commun.*, 1994.
[61] R. Vijayan and J. Holtzman, "A model for analyzing handoff algorithms," *IEEE Trans. Vehic. Tech.*, 1993.
[62] A. J. Viterbi *et al.*, "Soft handoff extends CDMA cell coverage and increases reverse link capacity," *Proc. Int'l. Zurich Conf. Digital Commun.*, 1994.
[63] N. Zhang and J. M. Holtzman, "Analysis of handoff algorithms using both absolute and relative measurements," *Proc. IEEE VTC*, 1994.
[64] E. Seneta, *Non-Negative Matrices*, George Allen and Unwin, 1973.
[65] N. Bambos, S. C. Chen, and G.J. Pottie, "Radio link admission algorithms for wireless networks with power control and active link quality protection," *Proc. IEEE INFOCOM '95*, Boston, MA, 1995.
[66] N. Bambos, S. C. Chen, and G. Pottie, "Channel access algorithms with active link protection for wireless communication networks with power control," Tech. rep. UCLA-ENG-95-114, UCLA Eng., 1995, submitted for publication.
[67] N. Bambos, S. C. Chen, and D. Mitra, "Channel probing for distributed access control in wireless communication networks," *Proc. IEEE GLOBECOM '95*, Singapore, 1995.
[68] H. V. Poor, *An Introduction to Signal Detection and Estimation*, Springer, 1988.
[69] D. Siegmund, "Sequential analysis: tests and confidence intervals," *Springer Series in Statistics*, 1985.
[70] A. N. Shirayev, "On optimum methods in quickest detection problems," *Theory of Prob. Appl.*, vol. 8, 1963.
[71] A. N. Shirayev, "On Markov sufficient statistics in non-additive Bayes problems of sequential analysis," *Theory of Prob. Appl.*, vol. 9, 1964.
[72] G. V. Moustakides, "Optimal stopping times for detecting changes in distributions," *Ann. Stat.*, vol. 14, 1986.
[73] M. Pollak, "Average run lengths of an optimal method of detecting a change in distribution," *Ann. Stat.*, vol. 13, 1985.
[74] M. Pollak, "Average run lengths of an optimal method of detecting a change in distribution," *Ann. Stat.*, vol. 15, 1987.
[75] A. N. Shiryayev, *Optimal Stopping Rules*, Springer, 1978.
[76] J. M. Rulnick and N. Bambos, "Mobile power management for maximum battery life in wireless communication networks," *Proc. IEEE INFOCOM '96*, 1996.
[77] J. M. Rulnick and N. Bambos, "Mobile power management for wireless communication networks," *IEEE/ACM J. Wireless Networks*, vol. 3, no. 1, 1997.

Additional Reading

[1] M. Gerla *et al.*, A distributed, mobile, wireless infrastructure for multimedia applications, *Proc. 5th Winlab Wksp. Wireless Info. Networks*, Rutgers Univ., 1995.
[2] S. C. Chen, N. Bambos, and G.J. Pottie, "Admission control schemes for wireless communication networks with adjustable transmitter powers," *Proc. IEEE INFOCOM '94*, Toronto, Canada, 1994.
[3] A. Alwan *et al.*, "Wireless, mobile, multimedia networks," *IEEE J. Pers. Commun. Sys.*, vol. 3, no. 2, 1996.
[4] S. C. Chen, N. Bambos, and G.J. Pottie, "On distributed power control for radio networks," *Proc. ICC '94*, New Orleans, LA, 1994.
[5] N. Bambos and G. J. Pottie, "On power control in high capacity cellular radio networks," *Proc. IEEE GLOBECOM '92*, 1992.
[6] S. C. Chen, N. Bambos, and G. J. Pottie, "On power control with active link quality protection in wireless communication networks," *Proc. Princeton Conf. Info. Sys.*, Princeton, NJ, 1994.
[7] K. Scott and N. Bambos, "The self-organizing wireless adaptive network (SWAN) protocol for communication among mobile users," *Proc. GLOBECOM '95*, Singapore, 1995.
[8] J. M. Rulnick and N. Bambos, "Power control and time division multiple access," *Proc. INFOCOM '97*, Kobe, Japan, 1997.
[9] J. M. Rulnick and N. Bambos, "Performance evaluation of power managed mobile communication devices," *Proc. ICC '96*, Dallas, TX, 1996.

Biography

NICHOLAS BAMBOS received a Ph.D. degree in electrical engineering and computer science in 1989 from the University of California at Berkeley. From 1989 to 1996 he was an assistant and later associate professor in the Electrical Engineering Department of the University of California at Los Angeles. Since 1996 he has been an associate professor at Stanford University in the Department of Engineering-Economic Systems and Operations Research and (by courtesy) the Department of Electrical Engineering. He received the NSF National Young Investigator Award in 1992. His current research interests are in modeling, analysis, and design of communication networks and distributed systems.