Chapter 1

Related Work and Preliminary Considerations

1.1 Introduction

In this chapter we take a preliminary look at *ad hoc wireless networks*. This is currently a hot research area, especially because there is an increasing need for connectivity 'anywhere' and, in particular, 'anyhow' (with and without a fixed infrastructure). While traditional networks have fixed nodes with wired connections (either optical fibers or copper lines), ad hoc wireless networks can, in general, be described as *multi-hop wireless networks* with *mobile* nodes. However, the mobility condition can be relaxed, and we can identify an ad hoc wireless network as a network where all the nodes are connected through wireless links, and where there is not a central or dominant node – as opposed to, for example, the case of cellular wireless networks where a base station (BS) exists in each cell. All the nodes in an ad hoc wireless network are at the same hierarchical level. In this sense, sensor networks can be regarded as a special case of ad hoc wireless networks.

Communication design in an ad hoc wireless network in a very general and meaningful way is a very challenging and complicated task [1]. The simple fact that the communication design should be sufficiently general to incorporate both the case of fixed nodes and mobile nodes is, in and by itself, a difficult objective to meet. This chapter is a preliminary high-level assessment of the situation, with the aim of understanding 'how' an ad hoc wireless network should be designed. In particular, we are concerned with the capabilities and limitations that the *physical layer* imposes on the network performance. In fact, most of the existing literature focuses on higher layers (such as the network and medium access control (MAC) layers), 'taking for granted' that the lower layers, and in particular the physical layer, can successfully cope with the channel impairments. This assumption is reasonable in networks with very reliable communication links (e.g. fixed optical networks). However, this assumption is much less meaningful in the case of wireless networks, where the radio communication links are very unreliable and subject to weather and environmental conditions. This leads to a more severe channel distortion (e.g. channels with fading, either non-selective or selective). Hence, it is necessary to take into account the channel characteristics in designing an ad hoc wireless network. In particular, it is desirable to come up with an integrated design comprising both the physical, MAC and network layers. This is the goal of the remaining chapters of this book.

Ad Hoc Wireless Networks: A Communication-Theoretic Perspective Ozan K. Tonguz and Gianluigi Ferrari

The remainder of this chapter is organized as follows. In section 1.2, we briefly review the various approaches for the design and analysis of an ad hoc wireless network that appeared in the literature. In section 1.3, we make simple and preliminary considerations for a more meaningful approach to the design of an ad hoc wireless network, taking into account the physical layer. In section 1.4, an overview of the major underlying assumptions considered in this book is presented. Section 1.5 concludes the chapter and provides the reader with an overview of the main philosophy behind the book.

1.2 Related Work

Ad hoc wireless networks have attracted a lot of attention over the last few years, because of the increasing demand for ubiquitous connectivity. As mentioned in section 1.1, the design of ad hoc wireless networks seems to require novel approaches, since they have peculiar characteristics which differ substantially from those of fixed networks or cellular networks, for which well-established design techniques already exist [2]. In the following, we briefly describe the main approaches that have appeared in the literature, indicating the potential limitations that are apparent at a first glance. However, it is fair to say that these limitations are understandable, since there are several constraints and it seems very difficult to take all of them into account simultaneously. Our final goal is to obtain a very general and adaptive model, and the simple considerations in section 1.3 point in this direction. We want to underline that the references considered in the following are by no means complete, and represent just a few samples of the much vaster ensemble that have appeared in the literature.

1.2.1 A Routing-Based Approach

Considering an ad hoc wireless network, a simple and immediate way to visualize it is to consider a set of *nodes* or *dots* distributed over a surface. These dots may be moving. Each node may want to communicate with another node in the network, hence the communication system needs to ensure that a *packet* sent by the starting node, the *source*, will eventually reach the intended node, i.e. the *destination*.

In a fixed network, one strives to find the shortest sequence of segments or *links* – each segment having two nodes at its end points – connecting the source with the destination. In this case, the focus is mainly on *routing*. However, in order for each node of this *route* to know which is the next node to forward a received packet to, it is required that each node has perfect knowledge of the network topology. This vision is very simplistic, and assumes that the transmission on each link corresponds to an error-free transfer of information. If this were the case, then it would be reasonable to only focus on routing. This, unfortunately, is not the case with links constituted by radio channels, hence this approach could be very limited for ad hoc wireless networks.

If the nodes are moving, then it is tempting to simply extend the above approach, focusing on the design of routing strategies which try to track the evolution of the network's topology. For example, extensions of the transport control protocol (TCP) to a mobile environment were proposed [3]. This is also the approach considered in almost all the possible routing protocols presented in [4]. In particular, the solutions proposed in [4] range from *proactive* routing protocols, where an updated description of the network topology is maintained at each node, to *reactive* routing protocols, which dynamically try to adapt to the changing conditions only if needed. In some cases, the authors claim that the numerical results account for realistic radio channel models, but it often seems that the physical layer is simply ignored. In all of these cases, it is assumed that the physical layer makes each link in the network an error-free connection. Hence, each node should only worry about the forwarding of an incoming packet. In this way, the focus again shifts to routing.

It is important to observe that some of the proposed protocols are interesting, and the underlying ideas are meaningful. For example, the dynamic source routing (DSR) protocol [5] and the zone routing protocol (ZRP) [6] are totally *on-demand* protocols, and the underlying ideas seem extendible to a more general design of wireless networks based on realistic physical constraints. The concept of associativity-based routing (ABR) [7–9] is also interesting: it indicates that the route to be preferred should not be the shortest one, but the one passing through the densest area of the network. This should ensure the longest possible route lifetime. The concept of *flooding* [10] and on-line local estimation based on a very few observables [11] also seem significant. Geographic random forwarding is considered in [12, 13]. A nice overview of architectures and protocols for ad hoc wireless networks is presented in [14].

1.2.2 An Information-Theoretic Approach

One of the fundamental and most intriguing concepts in information theory is the concept of the *capacity* of a single communication channel, measured in bits per second [15]. In a network, an extension of this concept leads to the *transport capacity* of a network, given by the product between the data-rate (b/s) and the distance (m) through which the bits can be carried. The transport capacity can also be interpreted as a measure of the *goodput* of the network [2]. This is intuitive, since the throughput increases either if the network can transport a few bits for a long range or many bits for a short range. In order to evaluate this theoretical network communication limit, information theorists allow themselves to make some unrealistic assumptions, for example in terms of routing strategy or MAC protocols. In [16], a first approach to the computation of the transport capacity of a network with fixed nodes is considered. The main result is that in a wireless network with N nodes distributed in a finite circle or sphere, with optimal *placement* of the nodes, optimally chosen *traffic pattern* and optimally chosen *transmission range*, the transport capacity is $\Theta(\sqrt{N})$, where the notation $\Theta(\cdot)$ indicates that the transport capacity is asymptotically, i.e. for $N \to \infty$, of the order of $\sqrt{N}$ [17]. This implies that the throughput per node is $\Theta(1/\sqrt{N})$. Hence, in a network with increasing node density – observe that the area where the N nodes fit is finite – the throughput per node goes to zero. This is somewhat obvious, since the number of hops that a generic packet has to make increases without limit. It is clear that this result, besides the optimality conditions mentioned above, does not consider at all the *delay* characterizing a packet transmission. While in [16] the authors claim that mobility should further reduce the transport capacity of a wireless network, in [18] this conclusion is challenged and the opposite is proved true. In [18], however, the authors make some unrealistic assumptions which justify their results. They assume that the buffering capacity of each node is unlimited and that a node perfectly recognizes when it can communicate to the nearest neighbor with a signal-to-interference ratio (SIR) above a given threshold. Moreover, there is no delay constraint. Given these premises, the routing idea is simply implemented by the following two phases. When a source node $\mathrm{n_S}$ wants to transmit a packet, it waits for the first node $\mathrm{n_R}$ passing by and transmits the packet to it. If this node is not the destination, then it becomes a *relay* node. This means that node $\mathrm{n_R}$ stores the received packet and keeps on wandering in the limited area. Whenever it comes near to the desired destination node $\mathrm{n_D}$, it just delivers the packet. As one can see, this is a very efficient communication protocol (they refer to it as *multiuser* *diversity routing*), with the least possible number of hops. However, it is obvious that this is a highly unrealistic communication protocol and could lead to very large delays (presumably not infinite, since the area where the nodes move is finite).

The information-theoretic approaches that recently appeared in the literature consider *ad hoc* conditions in order to maximize the transmission of information in the network. Constraints such as delay, storing capacity, realistic moving patterns (where a node is free to go away), power consumption and the impossibility of knowing the current SIR are simply not considered. While the concept of the capacity of a single-input single-output channel introduced by Shannon is a definite, simple and meaningful concept which represents a useful limit to take into account, it seems that an equivalent meaningful quantity for an ad hoc wireless network has not been clearly identified yet – the concept of *transport capacity*, however, well describes the information transfer in the network. The concept of *capacity per unit cost* [19] might be a possible candidate as well. An interesting information-theoretic perspective on multiaccess channel is presented in [20].

1.2.3 A Dynamic Control Approach

Wireless networks can be modeled as dynamic systems, where many parameters, for example the transmission protocols of each node, need to be dynamically adjusted [21]. In this sense, control theory could provide useful tools for the analysis of the network behavior. An approach of this type is suggested in [22], where the authors propose a routing scheme converging with probability one to the set of approximate *Cesaro–Wardrop equilibria*, which are suitably defined. The proposed adaptive scheme has two components: an iterative delay estimation scheme (DES) and an iterative flow adaptation scheme (FAS). The basic idea is that of associating to each node a particular time-varying flow, and then adjusting the flows from each node based on a few observables. It is possible to derive a set of ordinary differential equations in the flows, whose solution returns the steady-state behavior of the network. From this idea, in [23] a load adaptive routing protocol is proposed. It is arguable that a dynamic control approach could be meaningful in analyzing the convergence of the network flows structure. However, it seems difficult to use this tool to effectively define a communication protocol. It is easy to see that in this case also, a possible network analysis concentrates mainly on routing based on network flows.

1.2.4 A Game-Theoretic Approach

To date, it seems that there is no complete game-theoretic approach for the design of ad hoc wireless networks or, more generally, communication networks. The game which somehow seems to be more related with a communication network is the *maximum flow game*. Given a directed graph and identifying a source and a destination, the maximum flow associated with a particular subset of nodes is given by all the source–destination arc-distinct paths which can be obtained with the considered set of nodes. This game belongs to the class of combinatorial optimization games [24], as shown in [25]. The solutions of a game are generally related to the concept of the *core* of a game, which is a well-defined set of real vectors associated with the game structure. In [26] the authors show that proving that a real vector is not a core member of the maximum flow game is NP-complete. This, in turn, is equivalent to saying that finding the ensemble of flows in all the links which attain the maximum total flow is NP-complete.

It is still not very clear how the concept of a maximum flow game could be used in designing a communication protocol for an ad hoc wireless network. However, it is important to observe that most of the techniques used in combinatorial optimization games reduce to integer linear programming techniques. In this sense, linear and nonlinear programming techniques [27] could be very important tools in fixing the network parameters. In fact, given a meaningful objective function with variables representative of the ad hoc wireless network, and given a set of meaningful constraints, optimization theory leads to the solution of this problem in many cases.

1.3 A New Perspective for the Design of Ad Hoc Wireless Networks

In section 1.2, a quick overview of the main approaches recently reported for the design and analysis of an ad hoc wireless network was given. One can see that none of these approaches explicitly considers the *physical layer*, which plays a fundamental role in the case of radio channels. While in fixed networks it is reasonable to leave to higher levels the task of reconstructing the transmitted stream of information, for example requesting retransmission of a damaged packet (e.g. using automatic repeat request, ARQ), in ad hoc wireless networks the physical layer probably plays a fundamental role in combating the channel impairments. Moreover, most of the approaches considered in section 1.2 do not take into account (deliberately or inadvertently) some 'real' constraints in ad hoc wireless networks, such as:

- battery power consumption: this is a major limitation, since once a node has exhausted its power, it cannot support any communication. Hence, its power consumption affects the entire community, not only the node itself;
- network area: most of the results in the literature assume that all the nodes are confined to a precise surface with finite area. It is obvious that if the nodes move too far part, then radio communication is impossible. However, the communication protocol should accommodate a very general topological situation, where some nodes may go away;
- throughput and delay: the evaluation of these parameters, and in particular of their ratio, is fundamental. None of the results in section 1.2 clearly considers this performance parameter.

We now consider a simple approach, trying to figure out what happens when a source node, n_S, comes into an area where there *may* be other nodes, and needs to link itself to the network. We assume that there is global addressing, i.e. each node is associated with a unique address (for example, this will be the case in version 6 of the internet protocol, IPv6), and each node knows the desired destination address. We assume that each node is equipped with an omnidirectional antenna. We also assume that it is possible to quantify the *spatial density* of the nodes, and we define this parameter as ρ_S. A node may not be given this parameter. As soon as a node needs to communicate, it starts looking around to see if there are nodes in its proximity. For example, it could send a 'hello' message, or even the first packet of the information it needs to transmit – this implicitly assumes packetized data transmission. Given the omnidirectionality of the antennas, one can visualize the propagation of a message as an expanding *bubble* and we will refer to the transmission of a packet from a node as *bubbling*.

When the transmit power is depleted by the channel and the information is by and large unrecoverable, we assume that the bubble 'blows up'. If another node is reached by the bubble before it blows up, then this node can return an acknowledgment (ACK), i.e. bubbles back, to notify its presence. This analogy of bubbling can be formalized in the context of the theory of *continuum percolation* [28–30], which represents a statistical tool to analyze and characterize planar random processes. In [31], the authors consider *broadcast percolation* and in [32] this theory is applied to evaluate the impact of the use of BSs in sparse ad hoc wireless networks, to improve the likelihood of percolation. Note that this theory pertains to other scientific areas, besides engineering (e.g. molecular biology, disease spread study, etc.). In the case of ad hoc wireless networks, and in particular in the communication-theoretic approach proposed in the next chapters, this theory could play a major role in the *route discovery* (or *joining*) phase. After node $\mathrm{n_S}$ has sent its request to set up a route, if no node replies to the sent message, node $\mathrm{n_S}$ can proceed in other ways.

- The original transmit power per node, P_{t}, might be insufficient. If the node knew ρ_{S}, it could assume $P_{\mathrm{t}} \propto 1/\rho_{\mathrm{S}}$, i.e. the larger the number of nodes in the region, the lower the power that a node needs to reach its nearest neighbor. In any case, $\mathrm{n_S}$ bubbles with $P_{\mathrm{t}}' > P_{\mathrm{t}}$ and sees what happens.
- Given that the node can move, after reaching a predefined transmission power threshold $P_{\mathrm{t}}^{\mathrm{th}}$, the node can decide to move, following a pattern which brings it over a circle centered on its original position. It then moves repeatedly over the circle, bubbling from each new position. It can then transfer to an external circle and repeat the search. After a prespecified number of moves it stops.

The joining phase indicated above (repeated bubbling and moving) is limited by the finite amount of energy available at a node. If a node cannot keep contact with any other node, then it just stops and waits for some other node to pass by.

After the joining phase, we assume that $\mathrm{n_S}$ has at least a neighboring node. We assume that each node may keep track of its nearest neighbors only, defined as those reachable with a bubble – a node cannot have knowledge of the entire network topology, in order to account for any mobility pattern. Let us assume that $\mathrm{n_S}$ wants to communicate with another node $\mathrm{n_D}$, which is not one of its neighbors. A *percolation* process should start. One can visualize this as a sort of progressive bubbling, as shown in Figure 1.1: the nodes hit by the bubble generated by $\mathrm{n_S}$ generate new bubbles, and the external nodes hit by these bubbles in turn repeat the process, and so on. Hopefully, at some point $\mathrm{n_D}$ is hit by one of these bubbles. At this point, $\mathrm{n_D}$ starts sending back an acknowledgment, which should hopefully propagate back to $\mathrm{n_S}$ (along the discovered route), creating what could be defined as a *communication tube*, rather than a specific route. In Figure 1.2, a possible communication tube is shown. As one can see, it is not a fixed route, but nodes are allowed to exit and enter the communication tube, which can 'bend' in order to preserve connectivity.

Before describing the route maintenance phase, we briefly comment on the *multi-hop* nature of the packet transmission in an ad hoc wireless network. In the classical view of a fixed network [2], the communication links are assumed to be almost error-free. Hence, when a node receives a packet and forwards it to the next node, it is implicitly assumed that the integrity of the packet is preserved. However, this is far from obvious in an ad hoc wireless network, where usually each radio communication link rapidly degrades the quality of the transmitted modulated symbols. *Regeneration* at each node becomes fundamental. In order to perform regeneration, a forward error correction (FEC) strategy is very attractive.

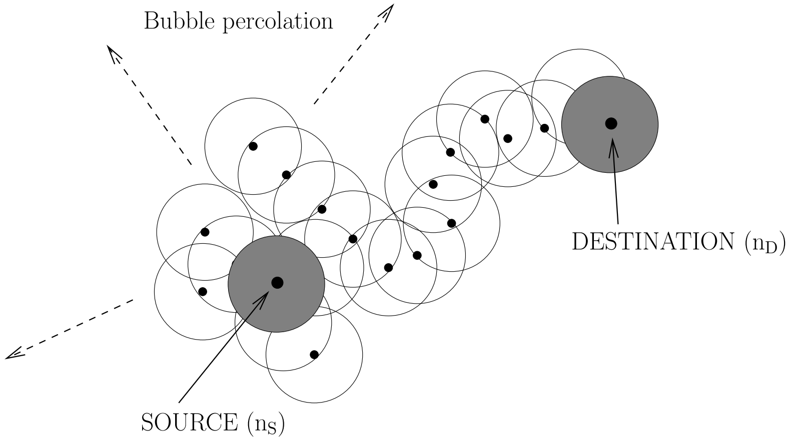

Figure 1.1 *Bubble percolation* for route creation.

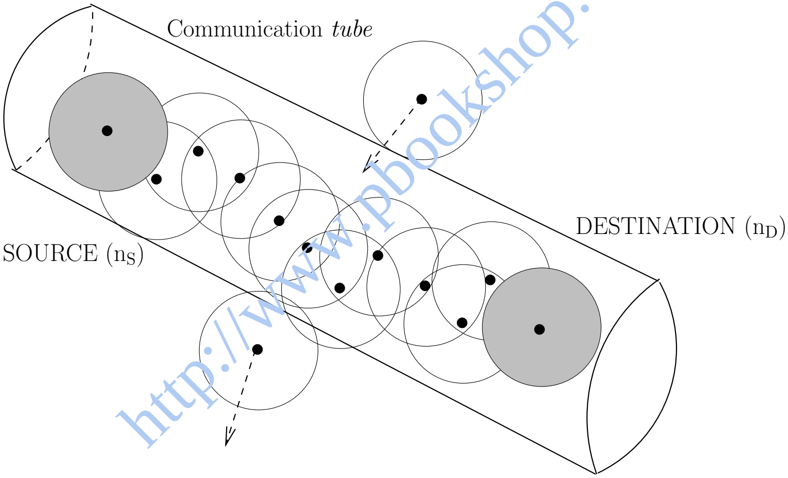

Figure 1.2 Communication tube between source and destination.

In particular, since the packets are expected to be short, powerful block codes with simple decoding techniques (e.g. threshold decoding [33]) could be used (for instance, short length turbo codes are not effective). ARQ seems to be locally feasible, as will be explained below. An intriguing question is how to use FEC capabilities. If no correction is considered for a significant number of hops, at some point it becomes useless to consider error correction, since the level of degradation is already too high. Hence, error correction could be considered in an early phase of the transmission. Error detection is also important, since the use of error detection motivates the activation of an ARQ protocol. A block code's error detection and correction capabilities are concisely described by its *minimum* or *free* distance d_{f}, which

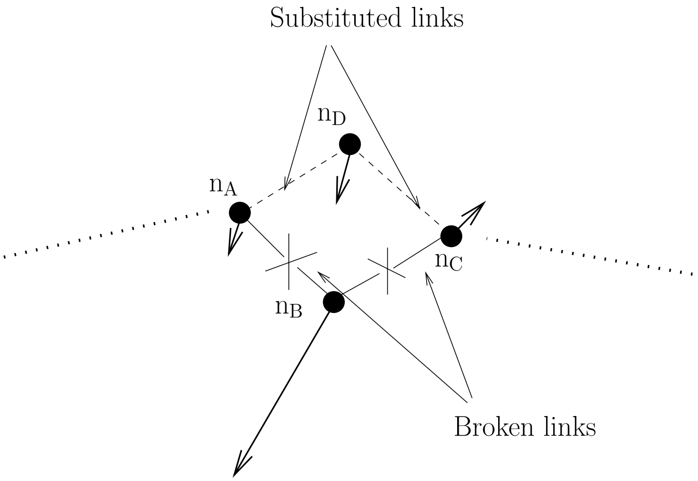

Figure 1.3 Local *opportunistic* route maintenance.

represents the minimum Hamming weight over all possible codewords – a block code can detect up to $d_f - 1$ errors and correct up to $\lfloor (d_f - 1)/2 \rfloor$ errors. Indicating by n_h^{max} the maximum tolerable number of hops, one could intuitively assume $d_f \propto n_h^{max}$ (the larger the number of hops, the more powerful the FEC code should be) and $d_f \propto 1/\rho_S$ (if the concentration of nodes is large, then the nodes are close to each other, and each transmission link is reliable). Since n_h^{max} is clearly proportional to ρ_S (the larger the concentration is, the larger the number of hops a packet should make to reach its destination), we can conclude that $d_f \propto \max\{\rho_S, 1/\rho_S\}$.

Assuming that a communication 'tube' has been established between the source and the destination, we focus on a possible maintenance strategy. Keeping in mind that the topology may not be fixed and may be rapidly varying, we can picture a *cloud* of nodes in the communication tube, which can move and change position. Hence, local and *opportunistic* route maintenance should be considered. In particular, it is reasonable to consider local ARQ, in the sense that a node can ask for retransmission from the immediate previous node, but it is unrealistic to use ARQ from the destination to the source in a multi-hop wireless network scenario. Moreover, due to a possibly changing network topology, links between two consecutive nodes are likely to break down. As a simple example, in Figure 1.3 we consider an intermediate portion of the route in the communication tube, with three nodes n_A, n_B and n_C. Node n_B is acting as a *relay node*, forwarding to n_C the packets it receives from n_A. However, it may happen that n_B is moving away from n_A and/or n_C, hence this relay is going to be torn down. Node n_B needs to be replaced by a substitute, if available. For example, when n_A, n_B and n_C realize that the topology is changing – the return time of an ACK increases – they can communicate in order to find a substitute. Node n_D may be available, slowly moving and hence reliable. A sort of *hand-off* should happen, in such a way that node n_D replaces node n_B as a relay between n_A and n_C. One of the quantities which the nodes should transmit together with the information packet is their *speed*, if available. This could help in understanding when a link is likely to break, as is the case with a fast moving node. Hence, *local route maintenance* is fundamental.

The speed seems to play a very important role in an ad hoc wireless network. As mentioned above, if a node could communicate its speed to its neighbors, this would significantly help in predicting the reliability of a route. Moreover, given an average speed $\overline{v}$ of the nodes in the network, the average dimension $\overline{L}$ of the packets to be transmitted should satisfy $\overline{L} \propto 1/\overline{v}$ (the faster the nodes are, the shorter the packets in order for opportunistic routing to be feasible). This intuition will be confirmed by numerical results shown in Chapter 6. Up to this point, we have implicitly referred to a scalar speed v. However, in order to fully characterize the motion of a node, the angle θ with respect to the west–east oriented direction could be considered. Hence, the mobility of a node can be described by two time-varying random processes: $v(t)$ and $\theta(t)$. It seems reasonable to assume that the speed process and the angle process are independent stochastic processes. However, if the speed v is large, then it is unlikely that a sudden change of angle can happen. We can formulate this by considering the autocovariance function $C_\theta^{(\overline{v})}(\tau)$ of the process $\theta(t)$ (assumed to be stationary, for simplicity) parameterized by the average speed $\overline{v}$. In particular, we can assume

$$C_\theta^{(\overline{v})}(\tau) \simeq 0 \quad \text{if} \quad |\tau| > b_{\overline{v}}$$

where $b_{\overline{v}}$ limits the autocorrelation and depends on $\overline{v}$. The dependence of b on $\overline{v}$ could be such that if $\overline{v}_1 > \overline{v}_2$, then $b_{\overline{v}_1} > b_{\overline{v}_2}$ (the faster the average speed, the more correlated are successive angular directions).

We now briefly discuss a possible game-theoretic perspective in the planning of an ad hoc wireless network. In section 1.2, the maximum flow game [25] was mentioned. In general terms, an ad hoc wireless network is an association of nodes that cooperate. Hence, in this cooperative scenario, each node should try to act for the benefit of the entire network, rather than for its own benefit. However, it should also preserve some resources for personal use. For example, each node could reserve up to 50% of its transmission power to act as a relay node for other nodes, while keeping 50% just for its own communication needs. This implies that once the power reserved for the network has been consumed, the node then ‘hides’ itself, refusing to act as a possible relay node. Moreover, if a cooperative communication structure can be defined (game theory could be a valuable tool for evaluating performance), this could be beneficial for security purposes. In fact, it is highly likely that a ‘bad’ node will act in a non-cooperative way, leading to possible isolation from the community of ‘good’ nodes.

Finally, linear and nonlinear programming techniques [27] could be extremely useful in a preliminary quantification of the important parameters of an ad hoc wireless network. In fact, once a meaningful objective function and important constraints have been formulated, the solution of a suitable linear/nonlinear program could give significant hints on the specification of the major network parameters. In particular, the transmit power P_t, the average speed $\overline{v}$ of the nodes, the spatial density ρ_S, the free distance d_f of a possible channel code and other seemingly uncorrelated parameters could be simultaneously taken into account. Some of these relationships will be derived in subsequent chapters.

1.4 Overview of the Underlying Assumptions in the Following Chapters

In the remaining chapters of this book, a novel communication-theoretic framework for the analysis of ad hoc wireless networks is presented. This framework is characterized by a bottom-up approach, where the impact of the physical layer on MAC and routing layers is evaluated. In the following, we outline the fundamental underlying assumptions used in the remainder of this book. Note that these assumptions will appear across the following chapters, although some of them might be 'latent' and not immediately clear.

- Peer-to-peer (P2P) multi-hop radio communication is considered.
- Two types of switching will be considered. If a source node, in need of communicating with a destination node, first reserves a series of intermediate relay nodes and then starts transmitting, we will refer to this as *reservation-based* (RB) switching. In this case, a node cannot serve as a relay in more than one route. If, instead, no intermediate node is uniquely reserved, i.e. a node can act as a relay for more than one route, we will refer to this as *non-reservation-based* (NRB) switching. A comparison between these two switching schemes is the subject of Chapter 7.
- Static networks with a grid node distribution are considered in the first part of the book. The impact of speed is the subject of Chapter 6 and clustering is partially analyzed in the final chapter.
- We do not consider how to build and maintain a route. In other words, we assume that route creation is 'magically accomplished'.
- In a scenario with RB switching, once a route has been reserved, then relay nodes, by definition, cannot generate new packets. This assumption can be reinterpreted by saying that if the number N_{c} of nodes wanting to transmit is larger than the maximum number of simultaneously active routes N_{R} then $N_{\mathrm{c}} - N_{\mathrm{R}}$ nodes will have to wait without generating packets.
- Once a route is formed, our analysis does not change regardless of how long the source node keeps the route. In other words, fairness is not explicitly considered in the proposed framework.
- As a benchmark, the case of ideal performance without inter-node interference (INI) is considered. In reality, this would correspond to the use of perfectly orthogonal spreading codes in each multi-hop route or to the use of disjoint frequency bands in the active multi-hop communication routes.
- Stability is not an issue in the considered communication model. In fact, the assumption of generation of information only by nodes with an activated route guarantees the absence of any instability phenomena.
- Denoting by λ_{g} the average packet *generation* rate (dimension [pck/s]), by L the size of each packet (dimension [b/pck]) and by R_{b} the transmission rate on the channel (dimension [b/s]), a necessary condition for the network to work properly is that $\lambda_{\mathrm{g}} L \leq R_{\mathrm{b}}$. This can be interpreted in terms of total traffic generated and transmitted. In fact, if N_{R} is the number of active communication routes (and generating nodes), the network-wide generated traffic is $N_{\mathrm{R}} \lambda_{\mathrm{g}} L$ (dimension [b/s]) and the total rate of transmission of information is $N_{\mathrm{R}} R_{\mathrm{b}}$ (dimension [b/s]). The condition that the transmitted traffic is larger than the generated traffic can be written as $N_{\mathrm{R}} \lambda_{\mathrm{g}} L \leq N_{\mathrm{R}} R_{\mathrm{b}}$, i.e. as $\lambda_{\mathrm{g}} L \leq R_{\mathrm{b}}$. In the rest of the book, we will also refer to the average packet *transmission* rate λ_{t}. In a steady-state network communication scenario, we will assume (in some chapters) that $\lambda_{\mathrm{g}} = \lambda_{\mathrm{g}} = \lambda$. This assumption, besides being reasonable, leads to a simplified performance analysis.

- One of the key performance metrics used to evaluate the network performance consists of the evaluation of the *effective transport capacity*, which corresponds to the total bandwidth–distance product carried by the network. This quantity is analyzed in detail in Chapter 5.
- In the case with INI, we make use of a novel bit-level (rather than packet-level) interference analysis.
- As examples of possible MAC protocols with random access, we will consider two simple protocols, defined as ‘reserve-and-go’ (RESGO) and ‘reserve-listen-and-go’ (RESLIGO). In the case of NRB switching, the key idea of the first MAC protocol is that a node, after reserving a multi-hop route to its destination, starts transmitting, regardless of the activity of the other nodes in the network. In a scenario with the RESLIGO MAC protocol, a node, after discovering a route to its destination, first listens: if no other node is transmitting, then it starts transmitting. These two MAC protocols are described in detail in Chapter 3.

1.5 The Main Philosophy Behind the Book

While a lot of different performance metrics have been proposed for ad hoc wireless networks (and sensor networks), among these the most prominent one is probably the transport capacity, which is equal to the maximum theoretical bit rate–distance product which can be supported in the network [16]. Thus, transport capacity gives a measure of the amount of information carried in such networks as well as the distance over which this information is carried in the network (hence, the unit [(b m)/s]). There is no doubt that transport capacity $\triangleq$ bit rate $\times$ distance is a meaningful metric for overall network performance in ad hoc wireless networks. While this is true, we take the viewpoint in this book that this information, in and by itself, might not be sufficient to describe the overall performance and dynamics in the network.

To explain this somewhat counterintuitive point, consider the following example: suppose we know that the transport capacity of an ad hoc wireless network is 10 (Mb m)/s. This information, in and by itself, does not tell us whether the network can carry 1 Mb/s for 10 hops (assume each hop is of length 1 m to simplify the example) or whether it can carry 10 Mb/s for one hop. Obviously, we need to know a lot more about the connectivity properties of a network (e.g. the average number of hops per route and the average number of simultaneously active routes) to have a better feel about the implications of a transport capacity of 10 (Mb m)/s. Another concern is that transport capacity, in and by itself, does not provide explicit information about the conventional network performance metrics such as delay and throughput (and goodput). In this book we take the viewpoint that to quantify the network performance of ad hoc wireless networks in terms of delay and throughput is important.

We further argue that a more realistic network performance characterization can only be obtained by taking the physical (PHY), MAC and network layer characteristics into consideration and by studying the relationships and couplings between these different layers of the protocol stack.

To that end, we show in this book that the throughput–delay performance will be affected by the capabilities and limitations of the PHY layer which, in turn, is affected by the choice of the MAC protocol used in the specific ad hoc wireless network under consideration.

Our results clearly show that the choice of MAC protocol affects the interference between nodes, thus affecting the end-to-end bit error rate (BER) and packet error rate (PER) of a route. In fact, the choice of the MAC protocol can affect both the transport capacity and the PHY layer performance in terms of the link signal-to-noise ratio (SNR) and, hence, the end-to-end BER/PER.

It is important to understand that the wireless links in an ad hoc wireless network are very unreliable and due to several potential reasons (such as interference, fading and shadowing effects due to mobility, exhaustion of battery power, etc.) a deterioration in the end-to-end route BER is quite possible. Such a degradation of the PHY performance can, of course, be combated via coding and ARQ techniques. This, however, has an adverse effect on the delay–throughput performance of the ad hoc wireless network. In addition, several retransmissions or FEC techniques would decrease the lifetime of the specific node and, consequently, the longevity of the network. These considerations clearly illustrate the strong coupling between PHY, MAC and network layers.

As mentioned before, transport capacity does not provide explicit information about the connectivity of an ad hoc wireless network or sensor network. We show in this book that this is a crucial component of network's behavior and without this component, critical pieces of information such as the average route length (in terms of the number of hops) and the average number of simultaneously active routes in an ad hoc wireless network cannot be accurately predicted.

The theoretical framework we develop in this book shows that connectivity exhibits a bimodal behavior: in other words, there is a critical node spatial density of the network above which the network is fully connected and below which the connectivity level in the network rapidly decreases to zero. These predictions are also endorsed by percolation theory. The communication-theoretic framework developed in this book also establishes the relationship between connectivity and the PHY layer (end-to-end route BER) as well as the relationship between connectivity and route discovery, and routing and forwarding.

It is important to emphasize that connectivity is ultimately related to the topology and the mobility of ad hoc wireless and sensor networks. Topology, among other things, could be heavily influenced by the level of mobility in the network. For example, a high level of mobility in the network may lead to a random topology which, in turn, could adversely affect the connectivity in the network. Increased mobility will also impact PHY layer performance by changing the end-to-end BER/PER performance.

Mobility of nodes could also lead to several other undesired consequences such as shadowing and/or fading and also to the clustering of nodes which, again, might have a bearing on the level of connectivity that can be sustained by the network. If the shadowing or fading effects on certain links (hops) of a multi-hop route are strong and deleterious, then our framework shows that some form of power control might be necessary. Likewise, severe clustering effects might necessitate more intelligent (adaptive) MAC protocols through the use of power control mechanisms. Measures such as power control or adaptive MAC protocols will clearly involve heavy exchange of signaling and control messages, thereby affecting the throughput–delay and transport capacity of the network.

The communication-theoretic perspective we develop in this book also provides important hints and directions for incorporating other network layer issues (such as fairness) and application layer issues (such as security) into the large framework. In other words, the theoretical framework we develop also lends itself nicely to quantifying fairness and security concerns in ad hoc wireless networks. In addition, we believe that the framework provided

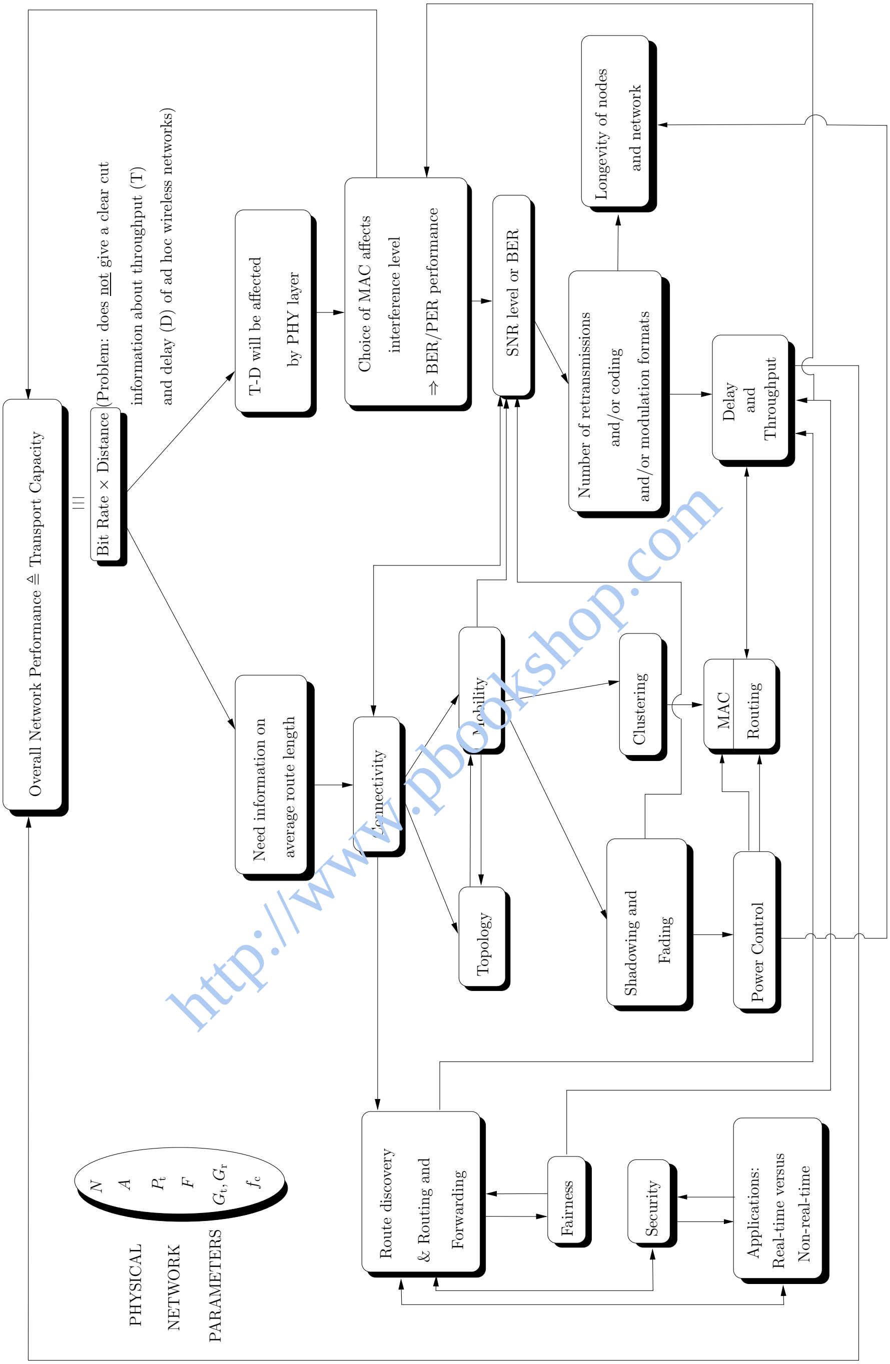

Figure 1.4 Overall communication-theoretic network perspective.

in this book is also conducive to evaluating, for given network parameters, which type of real-time and non-real-time applications can be supported by such networks.

Finally, it is worth mentioning that one of the virtues of the communication-theoretic framework developed in this book is the fact that most important performance predictions are given by closed-form expressions which are functions of key physical network parameters such as N (the number of nodes in the network), A (the coverage area of the network), P_t (the transmit power of each node), F (the noise figure of each receiver at each node), f_c (the carrier frequency the network is operating at), and G_t, G_r (the antenna gains of transmitting and receiving antennas, respectively). This fact enables the reader to explicitly see and appreciate the impact of each physical network parameter on the network's performance. Such insights might be very useful for researchers and engineers involved in designing real-world sensor networks and ad hoc wireless networks.

The overall philosophy of the book is illustrated in Figure 1.4.