

## LECTURE 12

# Designing High-Capacity Wireless Networks

This lecture discusses the key principles used in designing high-capacity wireless networks. We identify three main ideas that are used in a variety of different networks:

1. *Make every transmission count.* One way of realizing this idea is to work hard to minimize the number of transmissions that don't lead to a useful (error-free) packet transmission. One example of applying this idea is in developing MAC protocols that reduce collision rates.
2. *Control errors.* Because bit errors are a fundamental property, high-capacity designs usually attempt to control errors and erroneous data transmissions.
3. *Maximize spatial reuse.* Maximizing the number of transmissions that can occur concurrently while following the two properties mentioned above improves capacity in wireless networks in practice.

The previous lecture discussed wireless MAC protocols, whose primary goal is to reduce collisions. This lecture starts with a discussion of error control strategies. We then describe the two dominant kinds of wireless network architectures: cellular and ad hoc. We discuss scaling issues in both contexts.

### ■ 12.1 Error Control

The main method to control bit errors is to use a combination of forward error correction (FEC) and link-layer retransmissions (ARQ, which stands for “automatic repeat request”). In addition, modern wireless devices can transmit at multiple different rates using different modulation schemes, adapting their transmission rate to prevailing channel conditions. Such schemes are often called *autorate adaptation* or *rate adaptation* schemes. In general, the idea in these schemes is to observe channel conditions such as signal-to-noise ratios (SNR) and bit or packet error rates, and pick a transmission rate that will increase the

likelihood of successful packet delivery. Higher transmission rates use aggressive modulation schemes (and perhaps less-resilient FEC), making them more vulnerable to bit errors at lower SNR. Hence, when the error rate is high, it may improve packet delivery rates to actually reduce the transmission rate.

One simple approach is as follows: the sender observes the packet error rate over some past window (*e.g.*, the past  $K$  packets or the past  $T$  seconds). If the loss rate (or SNR, or other such quantity) is lower than some threshold (say,  $T_l$ ), then the sender reduces the transmission rate. If it is higher than some other threshold, (say,  $T_h$ ), then increase the transmission rate. Usually, devices are able to transmit at a discrete set of rates. The autorate adaptation scheme used in most WiFi systems is based on this kind of idea (many schemes have been proposed in recent years).

Of course, many wireless devices also have the ability to control their transmit power levels, and may be able to both control errors and achieve high system-wide capacity using a combination of rate adaptation, error control, and power control. A complete system design for a wireless network involving all these “knobs” is still an open question in general.

## ■ 12.2 Maximizing Spatial Reuse

There are two different kinds of wireless network architectures: *cellular* and *ad hoc*. Cellular architectures achieve high capacity by partitioning the network into *cells* and by suitably *provisioning* resources within and between neighboring cells. For example, neighboring cells might be provisioned to use different frequencies, to reduce the likelihood that concurrent transmissions in two nearby cells will interact adversely. Usually, cellular networks require only one wireless hop before packets from a wireless device reach an access point or base station connected to a wireline network infrastructure. Cellular wireless networks are the most common form today.

A very different wireless network architecture might be required when the network has to be deployed quickly in an area where pre-planning and provisioning is not easy or is impossible, or if the network is to be deployed in remote areas. Such *ad hoc* wireless networks usually involve multiple wireless hops between communicating entities, some of which might even be access points connecting devices to a wireline infrastructure. Ad hoc wireless networks, originally proposed for mobile devices in remote areas (*e.g.*, for applications such as disaster relief or the military), are now gaining in popularity in two places: in wireless sensor networks, and in wireless “mesh” networks to bring inexpensive Internet access to users. In both cases, a key feature of these networks is the use of multiple wireless hops and absence of careful provisioning or planning in their deployment.

Interestingly, WiFi deployments in many locations today are set up as one-hop “cellular” systems, but because they are often run by independent parties (*e.g.*, people in different homes or office buildings), they tend to exhibit the same “anarchic” and unplanned characteristics of ad hoc wireless networks. Understanding how to achieve high capacity in these networks is, in many ways, similar to the same goals in multi-hop wireless ad hoc networks.

Cellular wireless networks achieve spatial reuse using two techniques: resource provisioning (*e.g.*, frequency allocation) and power control. The idea is that access points and wireless devices transmit in pre-configured frequencies (or using pre-allocated codes or

time-slots) and at one of a set of pre-defined power levels, so as to control how much they will interact with other transmissions in nearby cells.

## ■ 12.3 Scaling Issues

In general, designers of wireless networks have to consider three scaling issues that might limit their size and growth:

1. *Aggregate impact of far-away nodes.* The issue here is whether the aggregate impact of transmissions from a set of nodes that are each “far” from a particular transmitter-receiver pair can, in aggregate, cause the aggregate interference or noise to reach the point where the particular transmission may not be possible at a high enough rate. Specifically, as explained in the previous lecture, the maximum transmission rate is given by the Shannon capacity,  $B \cdot (1 + \frac{S}{N})$ , and the question is whether the aggregate  $N$  caused by other concurrent transmissions can cause  $N$  to become too high.

First, observe that MAC protocols don’t handle this problem, because they are concerned only with avoiding collisions in the “local” neighborhood. This question is relevant for transmitting nodes that don’t necessarily detect each other’s transmissions when sensing their carrier.

2. *The price of cooperation.* As nodes are added to a network, how much additional data-delivery capacity does the network gain? The question is whether the nodes in the network increasingly use their capacity to *forward* other nodes’ data, leaving little for their own data.
3. *Routing protocol scalability.* As the network grows in size, how does the routing protocol scale? Concerns include the amount of state per node, and the rate of routing traffic required to maintain a consistent routing topology; the latter issue is of particular concern if the nodes are mobile.

### ■ 12.3.1 Aggregate Impact of Far-Away Nodes

The following model is due to Tim Shepard.<sup>1</sup> Consider a radio network in which the impact of any concurrent transmission on any other reception may be modeled as noise (spread spectrum systems with orthogonal codes are an example of such a system). Suppose nodes are laid out in a two-dimensional space at constant density,  $\rho$ . Suppose also that each node is interested in directly communicating with one of its closest neighbors (*i.e.*, if a node wishes to send packets to a destination further away, a routing protocol would have to arrange for that multi-hop wireless data delivery).

It is easy to see that the distance to a nearest neighbor is proportional to  $\frac{1}{\sqrt{\rho}}$ . Call this quantity  $R_0$ . If each transmitter sends data at a power level  $P$ , and if the attenuation at distance  $r$  is proportional to  $r^{-2}$  (true for free space; in other environments, it falls off faster than that, *e.g.*,  $r^{-4}$ ), then the signal strength at the receiver is proportional to  $r^{-2}$ .

Now consider the total noise at this receiver. To calculate this quantity, we will estimate the contribution from all nodes at distance  $r$  from the receiver, and then integrate this

<sup>1</sup>T. Shepard, “A channel access scheme for large dense packet radio networks, Proc. ACM SIGCOMM 1996.

contribution for all values of  $r$  from  $R_0$  (the smallest value of  $r$ ) to infinity. The number of nodes in the annulus at distance  $r$  and width  $dr$  is equal to  $2\pi r\rho dr$ . The contribution of these nodes to the noise at a given receiver is equal to  $\frac{2\pi r\rho dr}{r^2}$ . Integrating that from  $R_0$  to infinity, we find that the aggregate noise is infinite!

That is bad news. Fortunately, we don't live in an infinite world. If we assume that there are  $M$  nodes in all, we find that at node density  $\rho$ , the maximum distance of a node,  $R_{max}$ , is given by  $\pi R_{max}^2\rho = M$ . Solving for  $R_{max}$  and integrating the aggregate noise from  $R_0$  to  $R_{max}$ , we find that the total signal-to-noise ratio falls off as  $\frac{1}{\log M}$ .

This result is good news for large-scale wireless networks, because it says that the SNR from lots of far-away concurrent transmitters falls off pretty slowly (the noise from them grows logarithmically in the number of nodes).

### ■ 12.3.2 The Price of Cooperation

How much of a node's data carrying capacity is used to forward other nodes' packets? It turns out that the answer to this question depends strongly on the workload and communication patterns. If senders are picked at random, and each sender picks a random receiver in the network, then the aggregate network capacity of  $N$  nodes scales as  $\sqrt{N}$ . This result is not good news, because it implies that the per-node network capacity goes as  $\frac{1}{\sqrt{N}}$ . Fortunately, few networks display totally random communication patterns, and in fact exhibit spatial locality. If that occurs, the per-node network capacity scales better: for instance, if all communication is over a constant number of hops, then the per-node capacity scales proportionally with  $N$ .

One way to understand the  $\sqrt{N}$  result is to model each network as a set of regions that can communicate concurrently. The previous analysis that showed that the SNR does not drop appreciably with the size of the network implies that this simplified model is a reasonable one. Hence, if there are  $N$  nodes and they are laid out at some constant density, then the *one-hop* network capacity is proportional to  $N$  (*i.e.*, the number of concurrent transmissions is proportional to  $N$ ), because it is proportional to the area of the space in which the nodes are laid out).

Now, if senders and sender-destination pairs are picked at random, the expected path length between a sender and destination is proportional to  $\sqrt{N}$ . This, in turn, means that the *end-to-end* network capacity goes as  $\sqrt{N}$ , because each end-to-end data transfer requires  $\sqrt{N}$  hops, and the total one-hop capacity is linear in  $N$ . As mentioned before, this result implies that the *per-node* capacity goes as  $\frac{1}{\sqrt{N}}$ : *i.e.*, as nodes are added to the network, each node's usable capacity goes down!

Ultimately, whether forwarding costs dominate and prevent large-scale wireless networks depends on workload and communication patterns. In many cases, it is possible to design networks where spatial communication patterns are more local than purely random.

### ■ 12.3.3 Routing Scalability

We talked a bit about geographic routing in lecture.