PHYSICAL LAYER CONSIDERATIONS IN BUILDING
A HIGH SPEED AMATEUR RADIO NETWORK

Glenn Elmore N6GN
550 Willowside Rd.
Santa Rosa, CA 95401

ABSTRACT

The high information capacity required by an amateur high speed network requires optimum uses of resources. Use of both directional antennas and the UHF and microwave bands is essential to obtain efficient use of hardware and spectrum resources. The use of shorter point-to-point links and small clusters of local users can achieve dramatic increases in user information throughput. Coordination and cooperation at all levels will be necessary to make a high speed amateur network a reality.

Amateur radio avoided dying in its infancy largely due to the establishment of a low speed digital network. The American Radio Relay League was founded to further this. Now we amateurs find ourselves entering the information age. I believe that if amateur radio is to continue in this age it must offer relevancy and that to do this amateurs must develop and implement a significant high speed information network; the alternative is for the hobby to wither and eventually die.

Although in the past radio amateurs have led the way into new technologies and operations, in recent times we have increasingly tended to adapt our operation and pursuits to existing technologies. Amateur packet radio, which has experienced very rapid growth in the last few years, is an example of this. Our interest in packet was stimulated by seeing similar communications within the industry and military complexes. The name given to the current link layer protocol, AX.25, is itself a variant of the name of a previous commercial protocol, X.25. Not only at lower layers do we see this borrowing of technology. The idea of a worldwide amateur BBS system and the greater dream for a digital amateur network have followed rather than led similar existing information services in the military and commercial sectors.

Certainly it is to be expected that more reuse of existing tools and methods will be required as our society and world get more complex. It is also true that insightful adaptation of methods and technologies often results in tremendous benefits. However, as we amateurs adapt our ways to meet the changing face of technology we need to examine the peculiarities of our applications, along with our strengths and weaknesses in order to achieve the most successful results.

I believe that if we are to succeed in developing and implementing a high speed amateur network that we must examine the fundamentals of communicating information by radio as well as our own resources and strengths and then design our network accordingly.

At the lowest ISO layers, physical and link, we have sought to implement packet communications by adapting existing hardware and protocols. Telephone modem hardware went into and is still

inside most TNCs. Similarly, our link layer operations are tailored after the fashion of an IEEE 802 model. Both of these were originally intended for wire lines, a very different environment from amateur radio. I believe that many of the problems which amateur packet is now experiencing are traceable to this mismatch of solutions and environments.

A high speed amateur network requires a blend of two parts; high speed communication of information and wide area general access. High speed offers the opportunity for a great breadth and depth of applications. By representing and transmitting a large amount of information digitally, the possibility for a wide range of applications exists.

A digital data stream can be used to represent voice, TV, FAX, as well as computer programs, files and data. Digital representation allows general transmission, storage and retrieval of this same data and also allows error detection and correction techniques to be used.

Along with this, a wide area network can allow amateurs to communicate and share resources in new ways. Such communication and shared resources could offer relevancy in the information age and rekindle the fundamental excitement and spirit with which the hobby began. The extent of possible applications of such a network applied to the diverse interests and pursuits of amateurs is truly staggering.

HIGH SPEED DATA

Neither of the above two ingredients has previously been seen in amateur radio. High data speed is necessary for both. Even a network providing low user speed requires high speed data if a large number of users are involved and all communication can not be carried out directly between end users. Any intermediate interconnection facilities become providers of a shared resource. If there is to be equitable sharing among many users then either the end users' data rate must be reduced for such communication or else these interconnections (possibly "backbones") must be capable of greater speeds in order to accommodate more than one user at a time.

A fundamental limitation to communication is noise. If this limitation were not present there would be no need for any particular transmit power, antenna gain or receiver bandwidth between two stations seeking to QSO. Whatever transmit power was recovered by the receiving antenna could simply be amplified as required to allow detection (recovery of transmitted information) to take place.

In actuality, signal power must be sufficient to allow separation of data from the noise power. Noise power at a receiver is

$$N = KT_eB$$

$$K = 1.38 \times 10^{-23} \frac{\text{watts}}{Hz \times \text{degrees}}, \text{Boltzmanns constant}$$

$$T_e = \text{effective system noise temperature, degrees Kelvin}$$

$$B = \text{Bandwidth, Hz}$$

While good receiver design can reduce T_e, for terrestrial links the receive antenna is almost always receiving direct radiation from an earth which is approximately 290 degrees K. Combined with imperfections in the receive amplifiers, losses in the antenna system and QRM/QRN, the effective unwanted signal (noise) level of the system is usually a little and sometimes a lot greater than this. Deep space links can effectively maintain lower system noise temperatures at the earth end but these are at present out of consideration for the bulk of radio amateur networking use and even if used, one end of such a link is earthbound and represents a high temperature noise source to the other.

Higher speed communication requires proportionately more signal power than lower speed. The Shannon limit[1], from information theory, sets the maximum channel capacity to be

$$C = B\log_2\left(\frac{S}{N} + 1\right)$$

$$C = \text{maximum channel error free data rate, bits/sec using}$$
$$\text{a correct coding scheme}$$

$$S = \text{Signal Power}$$

This capacity is an ideal, common modulation methods may require signal power at least 10 dB greater than this. Also, unless signaling systems (modulation techniques) with a greater number of states M, (where M = bits/symbol = Bps/baud) are used then these faster systems always require greater bandwidth and incur an increased amount of system noise which must be overcome by similar increases in minimum recovered signal power.

The Nyquist theorem indicates that all information may be recovered by sampling a channel at a rate, R_b, no greater than

$$R_b = 2B, \text{symbols per second (baud)}$$

In this optimum case therefore C = R_bM and equating with the Shannon limit above:

$$C = \frac{R_b}{2}\log_2\left(\frac{S}{N} + 1\right)$$

And the maximum number of useful bits/symbol, M_{max} is

$$M_{max} = \frac{\log_2\left(\frac{S}{N} + 1\right)}{2}$$

In high S/N cases, the "excess" signal power can be used to purchase larger M and a resulting higher capacity at approximately $\log_2(S)$. Using power to allow an increased bandwidth for a given S/N is more effective since channel capacity then increase linearly instead of logarithmically.

However, whether or not complex signaling methods allowing more bps/baud are used or not, eventually additional bandwidth is required because there are practical limits to available recovered signal to noise ratio. This ultimately requires greater bandwidth as data rate is increased.

Since total information transfer is equal to an average transfer rate over a time interval, the product of rate and time, it can be seen that the total amount of information communicated is ultimately dependent upon the amount of energy transferred, the product of average power and time, between the transmitter and receiver. For the normal case of uniform mean noise, this energy must be enough greater than the noise energy in the same interval to allow successful data recovery.

The problem of efficient use of resources in designing and implementing an amateur high speed digital network then involves finding and utilizing the most efficient techniques for conveying (transmitting) this information-carrying energy. In an environment of limited resources of funding and spectrum, amateurs must use available resources optimumly if we are to build and operate a high speed network.

Signal Propagation by Radio as a Function of Carrier Frequency and Distance

As it is usually presented, the so-called pathloss equation shows what portion of the transmitted power gets to a distant location which is separated from it by a distance in free space. The assumption is that the distance is great enough that both transmitting and receiving antennas are in the far field regions of the other. The definition of far field distance, D_f is often considered to be:

$$D_f > \frac{2D_a^{\,2}}{\lambda}$$

$$D_a = \text{maximum antenna dimension, wavelengths}$$

$$\lambda = \text{wavelength} = \frac{c}{\text{Frequency}}$$

In its most common form, the pathloss equation describes two isotropic antennas separated by a distance D. An isotropic antenna is an antenna which radiates uniformly in all directions. It is not physically realizable but is convenient for the sake of analysis. The portion of transmitted power recovered at the receiving end, L_{ii}, is

$$L_{ii} = \left(\frac{\lambda}{4\pi D}\right)^2$$

Given this representation, the amount of signal received decreases as either frequency or distance is increased.

A receiving antenna serves to intercept and recover a portion of the transmitted power. If the entire surface of a sphere of radius D were surrounded by perfect receiving antennas and if the outputs of all these antennas were totaled the sum would be the total transmitted power. No power is actually lost along a free space path. The amount of power a particular receive antenna recovers

depends on how big an aperture or "bucket" it represents and the strength of the transmitted field at the antenna's position on the sphere.

To start to make this model more like a real-world situation we can substitute for the fictitious isotropic antenna a directional antenna which can actually be constructed. Except for excessive dissipative or matching loss, a directional antenna is by definition one which has gain. It gives gain because it causes power which the isotropic antenna would have spread evenly in all directions to be focused or concentrated in one or a few directions while reducing it in other directions. It essentially redirects power from some undesired directions to another desired direction. Antennas of higher gain have more of this focusing ability. The gain, G, of an ideal antenna may be stated in terms of its capture area or aperture, A,

$$G = \frac{4\pi A}{\lambda^2}$$

where A $=$ antenna aperture, expressed in the same units as λ^2

From the vantage point of the receiving antenna on the sphere it makes no difference whether the source is a 100W transmitter and an isotropic antenna or a 10W transmitter feeding a directional antenna with a gain of 10. The field strength at a receiving antenna located in the far field would be the same in either case.

Antenna effective aperture is a measure of the useful area of an antenna. Figure 1 shows the apertures of some familiar antennas.
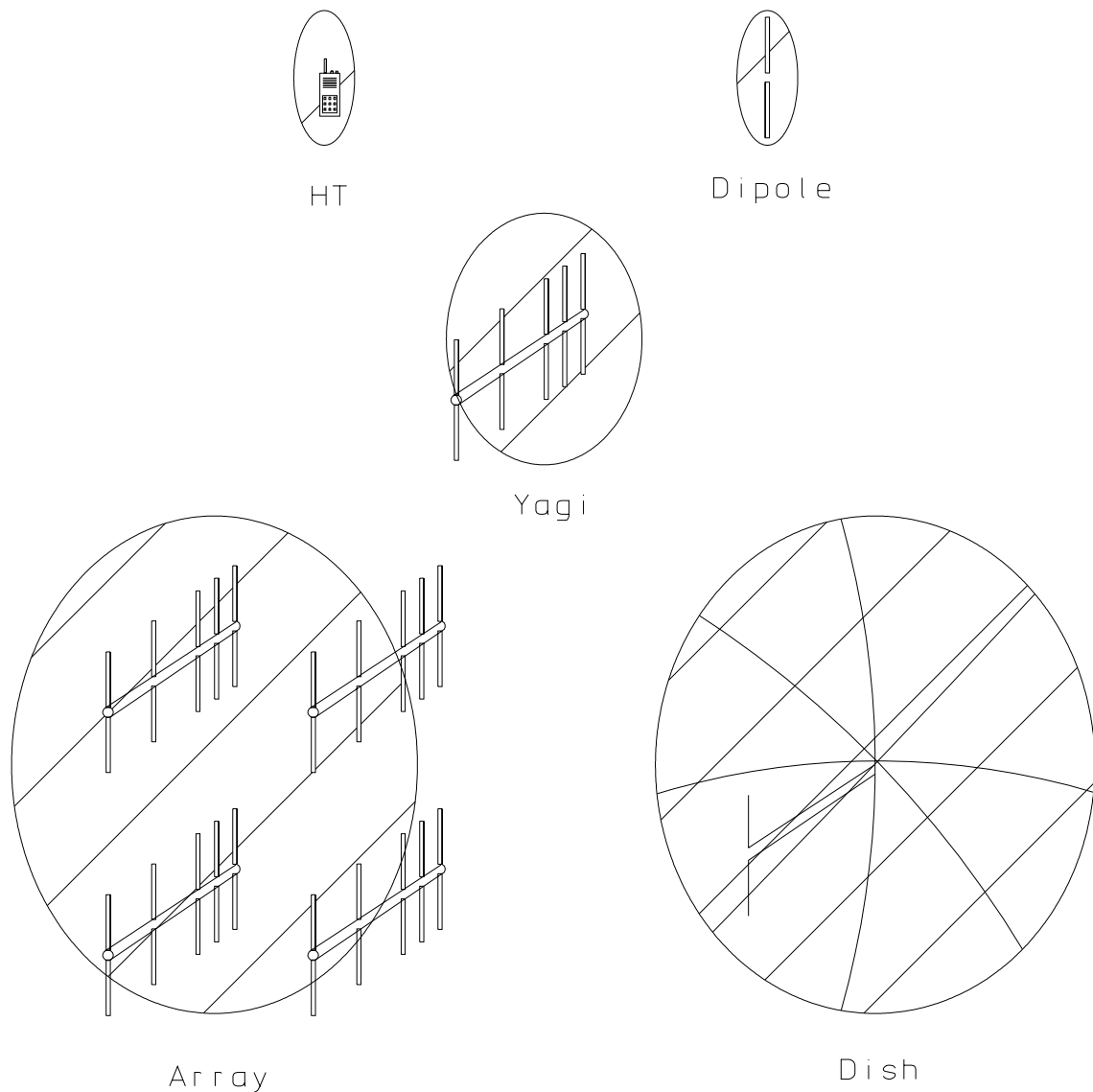
.

HT

Dipole

Yagi

Array

Dish

Figure 1. Relative Apertures Of Some Common Antennas

For simple single element antennas like a monopole or dipole, the aperture can be approximated by the area of a rectangle which is a half wavelength long and a quarter wavelength wide. It is not affected by the physical size of the conductor used to make the dipole. Also, even a shortened dipole or monopole, like the "rubber duck" antenna used on handheld radios, has almost the same aperture as that of a full-sized dipole. Consequently, this is the minimum aperture a real antenna can have.

Adding more elements, or electrical size, makes an antenna more complex and increases the aperture relative to a dipole or isotropic antenna. Antennas with more elements or electrical size have a relatively larger aperture on receive at the same time they have a more directional beam.

Now let's look at what happens to antenna aperture as frequency is changed.
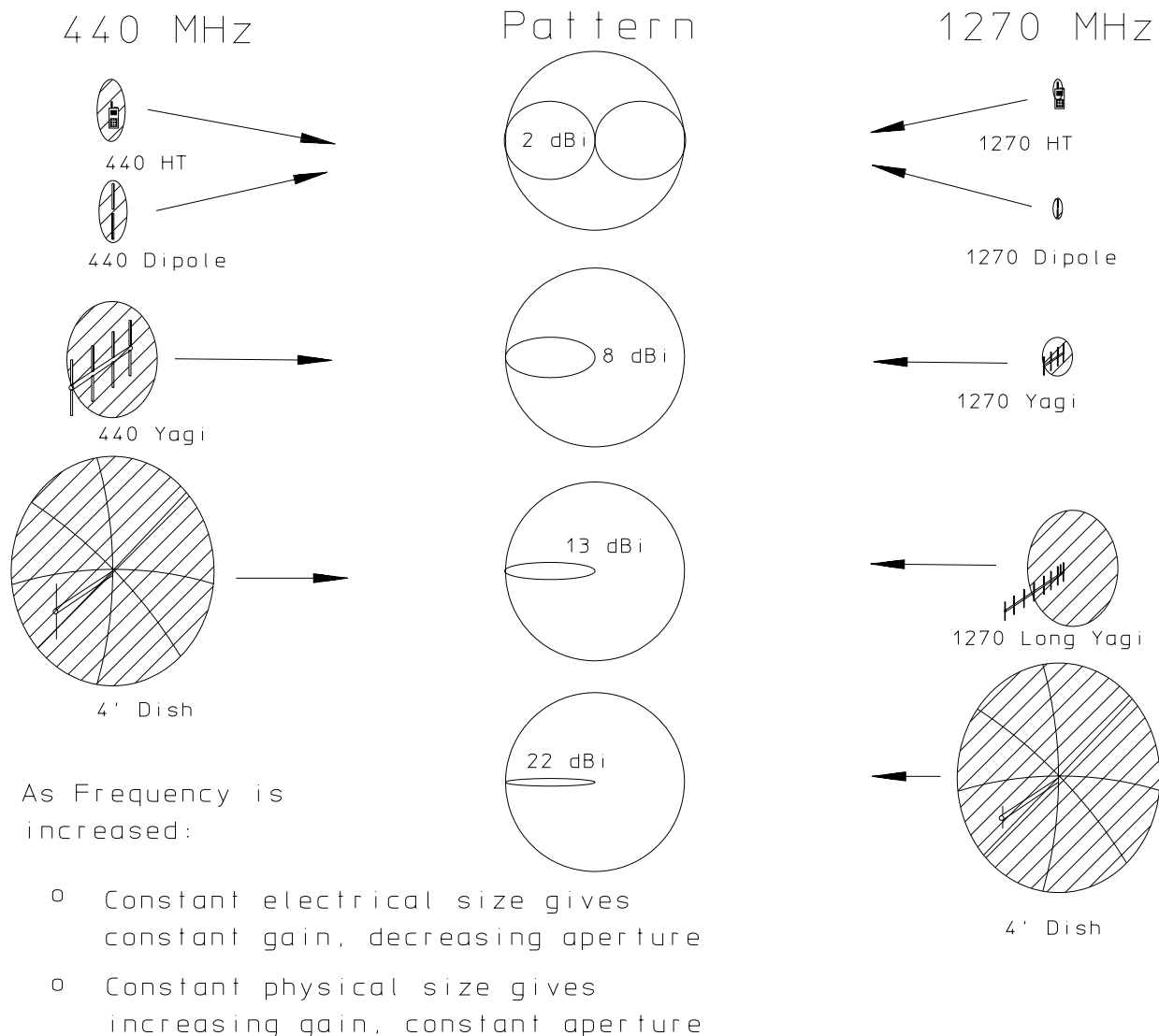

Figure 2. Antenna Apertures and Patterns Vs. Frequency

   Here again are some common antennas and apertures shown to scale for two amateur bands. As the antenna electrical size or number of elements is increased apertures increase too but for a given electrical structure the higher frequency antennas start out with a smaller physical aperture because their wavelength is smaller.

   As antennas much more complex than dipoles are considered, the aperture size is increasingly related to electrical antenna complexity. A 160M dipole may have more aperture or "intercept area" than a 150 foot parabolic reflector even though it is electrically a much simpler antenna. However, at frequencies where the dish antenna is electrically large, where it is at least 10 wavelengths in diameter, the physical aperture is relatively constant.

   Notice the dish antenna near the bottom of Figure 2. It has the same physical size on each band, although its electrical size, measured in wavelengths, is about 3 times as large on 1270 MHz as it

is on 440 MHz. It has about the same physical aperture on both bands although its gain, directivity and its ability to focus a transmitted signal is about 10 times greater at the higher frequency.

If we now substitute a directional receiving antenna of constant physical size, A_r, and with gain, G_r, for the isotropic receiving antenna the portion of transmitter power transmitted between them, L_{id} is

$$L_{id} = G_r L_{ii} = \frac{K_{id}}{D^2}$$

$$\text{where } K_{id} = \frac{A_r}{4\pi}$$

$$A_r = \text{effective aperture of the receiving antenna}$$

K_{id} is a constant and the constant aperture antenna recovers the same amount of the transmitted signal irrespective of frequency. There is no longer any frequency dependence to the equation, only a distance dependence.

This arrangement of using a constant physical antenna size instead of a constant electrical antenna type makes a lot of sense in practice. Almost always the limitations to amateur antennas at the hamshack or at a high level site are in terms of antenna physical size rather than antenna electrical size. Antenna and tower wind loading, rotor capability and antenna size are constraints much more often than number of elements or antenna dimension measured in wavelengths.

We have now come to the point of realizing that there is nothing inherently wrong with increasing frequency when we are seeking to communicate between two different locations, that is, transmitted energy doesn't mysteriously evaporate in space. If we put up a given sized "collector" and transmit a given intensity in its direction the same amount of received signal may be recovered, no matter what frequency is used.

But we aren't quite done changing the antennas yet. In exactly the same way that the receive antenna size is more likely to be physically rather than electrically constrained, so is the transmit antenna size. We aren't limited to using a particular electrical size for transmitting. In fact we are likely to want to use the same antenna for both receiving and transmitting since most of our communications will need to be two-way.

In this third rendition of the equation we transmit as well as receive using an antenna of constant physical size. This result is also known as the Friis Transmission Formula[2]

$$L_{dd} = G_t L_{id} = \frac{A_r A_t}{(\lambda D)^2}$$

$$G_t = \text{transmitting antenna gain}$$

$$A_t = \text{transmitting antenna effective aperture}$$

Because the directivity and gain are now increasing with frequency, the effective radiated power, ERP, is also increasing. The transmitter power is better focused to go only toward the receive antenna and not elsewhere, as frequency is increased. Once again there is a frequency dependence in the equation but this time instead of things getting worse as frequency is increased, as was the case with constant electrical antenna size, with constant physical antenna size the amount of

transmitted signal reaching the receiver increases with increasing frequency. In fact, an increase in distance incurs no additional reduction in recovered power if frequency is increased by the same amount.
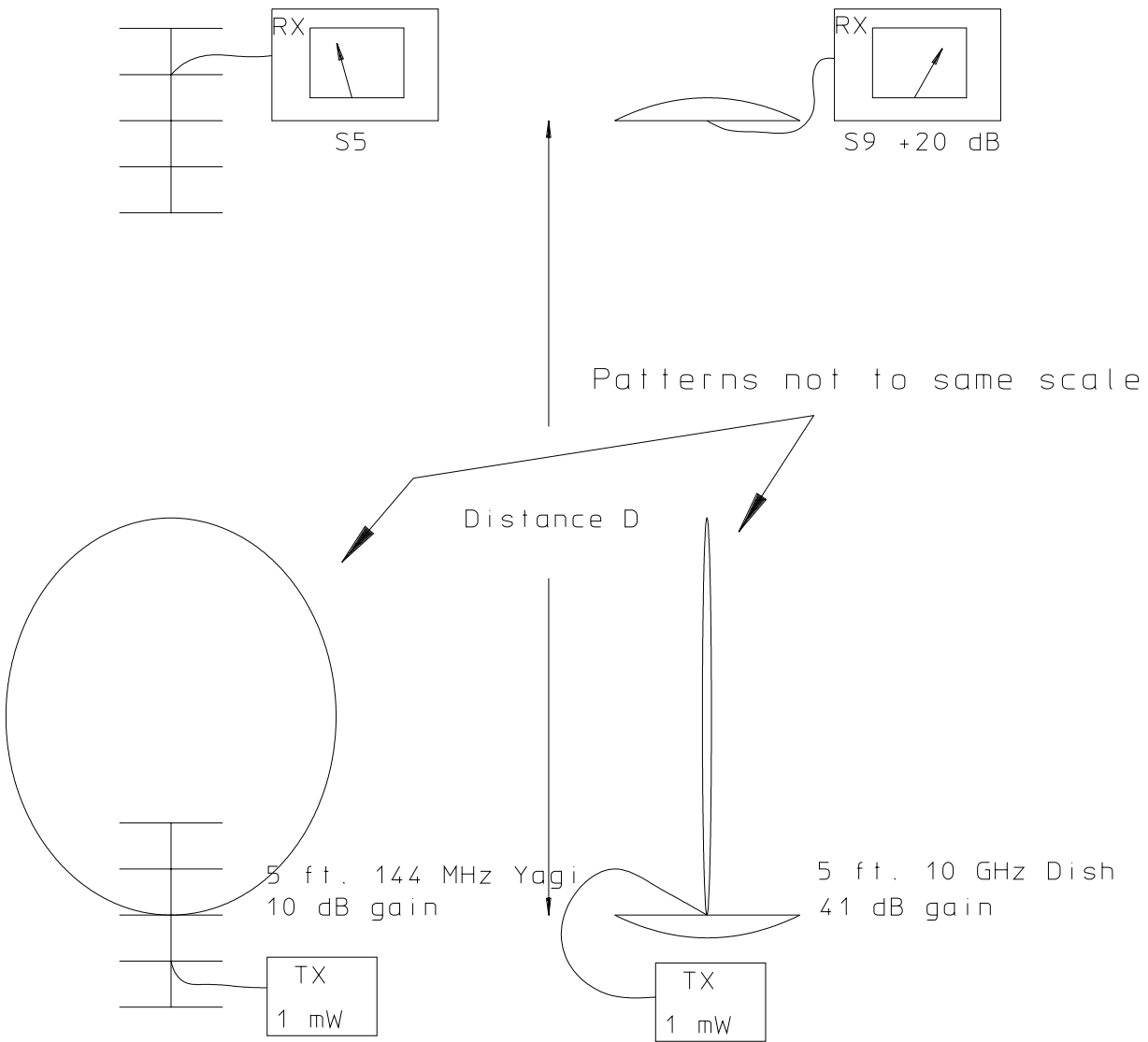

Figure 3. Practical Example Of Frequency/Directivity Improvement

   As a practical example, Figure 3 shows that a transmitter using 5' antennas at 10 GHz can provide more than 1000 times as much signal to the receiver as the same transmit power and antenna size on 144 MHz. Or, as an alternative, 1000 times as much information could be transmitted over the 10GHz system as over the 2M one in the same time interval.
   In summary, as frequency is increased when using constant physical antenna size for both receiving and transmitting the receive antenna has constant aperture and intercepts power within the same area but the transmitting antenna focuses the transmit power more resulting in more power recovered by the distant receiver.

Returning to the original goal of amateur communication of maximum "information carrying energy", we discover that to maximize "high-information-rate energy transfer", we must choose the highest available frequency where:

To maximize information/energy transfer, we must choose the highest available frequency where:
Physical antenna size can be maintained
Maximum ERP can be achieved. This is a trade-off of transmitter power expense for antenna gain
There is sufficient spectrum to support the signaling method and information rate
Propagation over the desired path doesn't attenuate prohibitively more than the free-space line-of-sight (LOS) case

Higher frequency, shorter wavelength, point-to-point links are the most effective way to transfer information by radio and therefore offer the most attractive solution to the problem of communicating high volumes of information rapidly.

Within the context of current radio amateur resources, this is indeed fortunate. It is presently only the amateur microwave and millimeter bands which offer the possibility of highly directional antennas and have sufficient bandwidth to support high volumes of information transfer which will be necessary for a successful high speed digital amateur network.

RESOURCE SHARING

The goal of an amateur network is to provide information exchange among a large number of users; not just an optimimized, high rate information transfer across a single link. Such exchange needs to be performed within the bounds of amateur hardware and spectrum resources. Also these available resources and the resulting supported services must be equitably shared.

Having explored the requirement for highly focused beams for optimum information transfer over a single link, let's turn to the question of how point-to-point (PTP) physical and link layer implementations affect the sharing of resources compared with omnidirectional-to-omnidirectional (OTO) implementations.
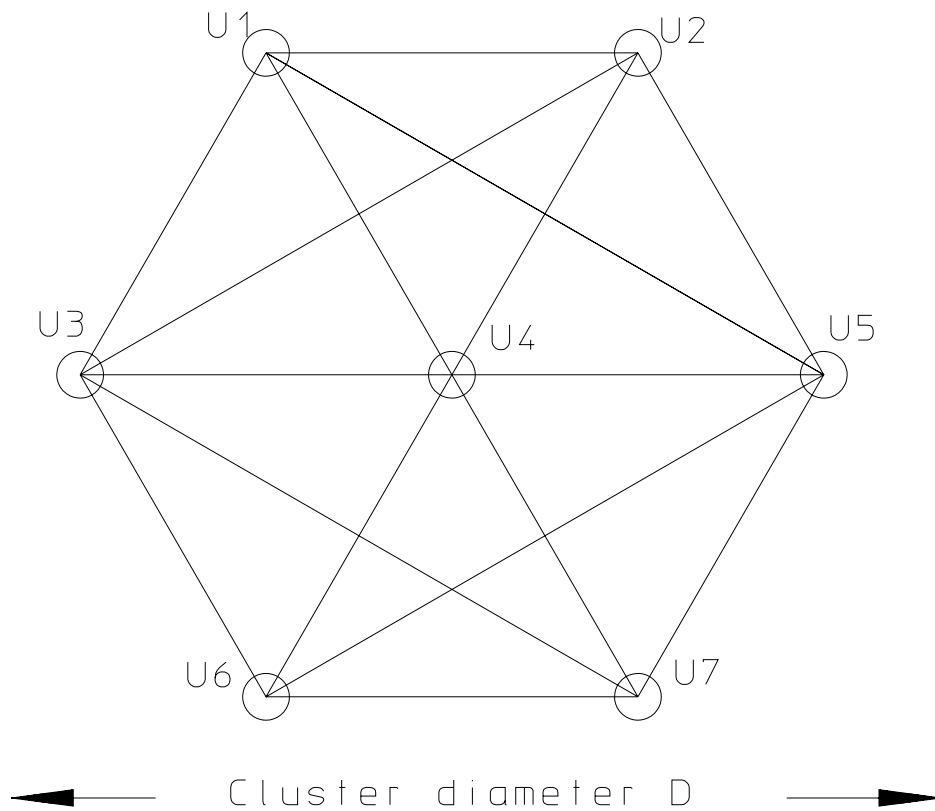
"

Figure 4.   CSMA/OTO Cluster

   Figure 4 shows an idealized representation of current amateur carrier sense multiple access (CSMA) packet radio. A number, U, of local users within radio range of each other are using omnidirectional antennas to receive one another's transmissions on a shared channel. If a "perfect" link layer protocol is in use, the channel capacity, C, may be thought of as being perfectly divided among these OTO users. If the channel is time shared each user may obtain the full channel data rate but only a portion of its information handling capacity. Each user can expect a share, R, of this capacity

$$R = \frac{C}{U}$$

R = user throughput rate
C = channel capacity
U = number of users sharing the channel

   As anyone experienced with amateur packet on congested channels located in typical terrain knows, this is very definitely an ideal rate. The value of U is indeed variable at any particular time.

The realities of hidden transmitters and packet collisions also degrade the ideal to an ALOHA case and rapidly reduce the throughput rate a great deal more.

When a station transmits in this environment, the channel effectively becomes unavailable to all other potential users. Transmitted power which is only being used to communicate with one other station (in a non-broadcast protocol) is causing the channel to be unusable by all other stations within radio interference range for the duration of the transmission. In addition, most of the transmitted power is going in directions other than that desired for that particular transmission. Consequently the data rate must be reduced so that the noise power in the occupied bandwidth is enough smaller than the recovered signal power to allow successful demodulation of the data. In this example, omnidirectionally transmitting the power is causing a double problem:

Less than optimum use of both the channel and the stations' hardware

Removal of the channel from use by other stations

Neither the information rate nor the resource sharing aspect of this implementation is optimum.

Fortunately, at the same time a PTP link more effectively transfers information it can also improve the resource sharing attributes of a network and improve the effective throughput rate available to end users. The directive nature of the antennas on a PTP link may reduce interference to other users at the same time it improves the performance of its own link relative to an OTO implementation.

An example may be useful to compare PTP link architecture with OTO

Example Benefits of PTP Links

Figure 5 contrasts OTO communication with PTP communication. In the two cases stations are shown at the vertices and center of a hexagon. The users are assumed to be members of a cluster located in an environment with many other users present at constant population density.

Directional antennas with a gain of 20 dB, which corresponds to a 3 dB beamwidth of about 20 degrees, are chosen for the sake of the example. An operating frequency in the PTP case of 10 times that in the OTO case could provide this directivity while maintaining similar physical antenna sizes. In the PTP case there are six users and a server, added to properly route the information. The same shared channel width and antenna size is assumed. This might not be the case in practice since PTP dedicated hardware could easily have a dedicated channel too but the assumption allows fairer comparison with the OTO architecture. The hatched areas show the regions which are effectively "deallocated" during transmission. Any other potential network users within these regions can not expect to have the same channel fully available since transmissions by cluster members can make the channel unusable by another station using an omnidirectional receive antenna.

OTO (dipole)
Antennas
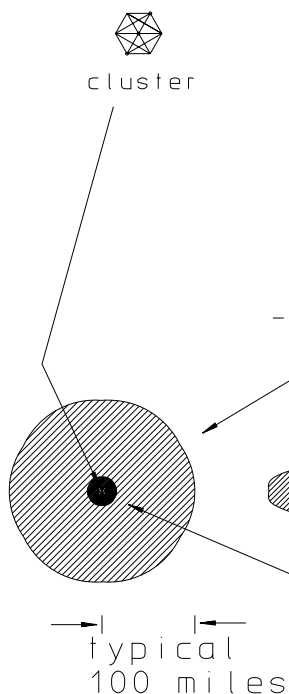
20 dB gain PTP
antennas
(reduced PTP Tx power)

cluster

cluster

-20 dB

-10 dB

environment
limits
QRM

typical
100 miles
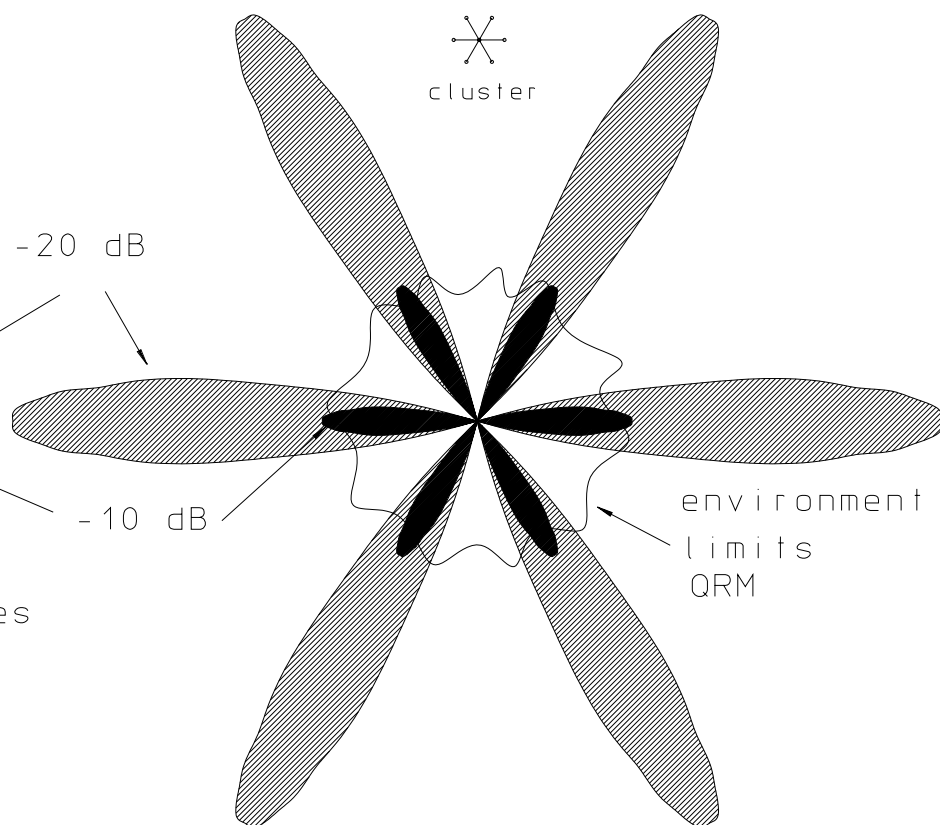
Figure 5. Comparison of OTO with PTP

The ends of these links are symmetric, the receive and the transmit antennas are the same. On receive the gain is a result of holding aperture constant as frequency is increased, on transmit this gain produces 100 times the ERP of the OTO case. This means that if transmitter power is held constant, 20 dB more signal power will be recovered. If a wider channel were used this could allow for an increase in the data rate of 100 times.

The PTP case does require the existence of a server, and a potential reduction of the throughput since the data must be sent twice, once to the server and again to the receiving user (our example uses only one shared channel and therefore no "cut-through"). But this is offset by the factor of two reduction in link path lengths; each user need only communicate with the server at the center instead of across the entire cluster diameter. This results in a net additional benefit. The potential improvement to a pair of users using PTP instead of OTO equipment is

100 (Tx antenna gain) X 4 (half pathlength compared to OTO) X 1/2 (data must be sent twice) =

200 times

$$C_p = 200C_o \text{ is available or } C_p = C_o \text{ with transmitters running 1/200th the output power}$$

where

$$C_o = \text{OTO Channel Capacity}$$

and

$$C_p = \text{PTP Channel Capacity}$$

This improvement is at the cost of an additional radio and data processing hardware for the server. However lower link budget margins are required to assure similar link up-time on the shorter path. In terms of energy and information transfer, the PTP case provides

$$G_t + 6 \text{ dB} - 3 \text{ dB} = 23 \text{ dB of improvement}$$

.

$$G_t = \text{transmit antenna gain}$$

In terms of channel reuse, if the power is reduced the same user data rate may be supported with a fraction of the interference.

Antenna directivity and gain may be considered in terms of beamwidths. The gain of the directional transmit antenna increases because its beamwidths are decreasing. Use of a "square" antenna, one with similar dimensions in both vertical (elevation) and horizontal (azimuth) planes will produce similar beamwidths in these two planes. Thus a consequence of increased transmit antenna gain is a focusing of power in both planes. However if all the channel users in our example are located on the same horizontal plane only the narrowing of the beamwidth in azimuth serves to reduce interference. Increasing transmit antenna gain by G increases the ERP by this same amount but reduces the azimuthal beamwidth by only sqrt(G). The G times increase in ERP increases the range by sqrt(G) but at the same time the narrowing of the beamwidth reduces the subtended angle by sqrt (G). The net result is that the same total geographic area is affected in PTP as for the OTO case.

If user density, users/unit-area, is constant over the entire region, the number of users, U, affected by a transmission (to the -20 dB level) is

$$\text{Number of users impacted} = \text{Area} \times \text{Density}$$

$$U = \pi (10D)^2 \left(\frac{1 \text{ user}}{\left(\frac{D}{2}\right)^2} \right) = 1250 \text{ users (approximately)}$$

In practice, solutions depend upon on the terrain and the relative elevations of the cluster members and the server. Also, in a real-world PTP situation, changes in the number and locations of cluster members may effectively deallocate a circular region. Coordination sufficient to maintain effective use of the "quiet" regions in between the petals of interference is probably not realistic.

On a flat earth with all users in the same plane and only LOS paths the increased range of PTP could cause interference with any of G times as many users, effectively counteracting most of the PTP link performance gains with similar spectrum deallocation. However, propagation in real surroundings can help to limit this problem.

A realistic case might have a cluster diameter, $D = 10$ miles. If the server is located sufficiently elevated to be within LOS of all members it will probably be primarily the servers' transmitter which contributes to interference much beyond the cluster. This is because the cluster members locations are in general going to have absorptive ground clutter. Also for cases with narrow antenna beamwidths, perhaps gains above 35 dB, user antennas will be pointed upwards toward the server and the server's antennas will similarly be pointed downward. In one case transmitter power will radiate into space, in the other it will largely be absorbed by the earth. As long as the server isn't over-elevated, it is likely that at terrestrial distances greater than $10D = 100$ miles that attenuation will usually have increased by much more than the additional 20 dB LOS space loss because of reduced antenna gain at this vertical angle, scattering and absorption from distant clutter, and the earth's curvature. Energy which would have gone into interference on a flat earth winds up being absorbed or radiated into space where it ceases to cause a problem.

The likely result is that similar maximum areas will be affected in the two cases. If the PTP case causes interference out to, say, twice the range of the OTO case the number of affected users (to the -20 dB level again) is

$$U_p = \left(\frac{20D}{10D}\right)^2 U_o = 4U_o$$

$$U_p = \text{number of users affected in PTP}$$

However, if these potentially affected users are also using directional receiving antennas, perhaps as part of other similar small PTP clusters, virtual elimination of interference may be obtained resulting in potential for much better channel reuse and considerably greater improvement.

In this example then, the PTP architecture would provide user throughput rates, R_p of between

$$R_p = \frac{C_p}{4U_o} = 50\frac{C_o}{U_o} = 50R_o$$

for the case where the same channel must also be shared with other nearby users utilizing omnidirectional antennas to

$$R_p = C_p = 200C_o$$

if spectrum is reused with directional receive antennas and there is no competition for the channel.

Final user throughputs might then range from

$$R_o = \frac{C_o}{1250}, \text{ OTO user throughput rate for a congested OTO case to}$$

$$R_p = C_p = 200C_o, \text{ rate for a non-contested PTP channel.}$$

Thus the user throughput rate might be from one to two orders of magnitude improved in a PTP

spectrum limited case when the channel is shared with OTO users. But if spectrum resources are not limited or if antenna directivity is great enough to allow complete frequency reuse, PTP users obtain the full G_t + 6dB - 3 dB channel improvement and aren't required to share the channel as in the OTO/CSMA case. This could amount to a relative improvement of

$$R_p \ = \ C_p \ = \ 200 C_o \ = \ 200 \times 1250 R_o \ = \ 250,000 R_o$$

more than five orders of magnitude while running the same transmitter power.

SOME PRACTICAL IMPLICATIONS OF A PTP NETWORK

As already indicated, in order to take best advantage of PTP hardware it is necessary to choose the highest available frequency where antenna size, ERP, bandwidth and propagation permit effective operation.

Antenna Size
To maintain the full 20 log(Frequency) advantage, the comparisons were made between antennas of constant physical aperture. The maximum acceptable size for an antenna will depend upon many things but a 145 MHz dipole or groundplane has an aperture of about .25 M^2. This means that a properly constructed antenna with at least this much capture area will provide the expected improvement. Depending upon the antenna efficiency, the ratio of gain and directivity, the physical antenna size may have to be somewhat greater. An antenna of up to one meter on a side, seems acceptable for many uses.
A 1 meter dish can provide about 20 dBd (dB relative to a dipole) gain at 1250 MHz or 40 dBd at 10 GHz. To get this gain at 145 MHz would require a yagi nearly 10 meters long. At 1200 MHz and below yagis, arrays of yagis and collinear antennas are convenient. A very inexpensive conical approximation of a parabola can be made for 1200 MHz from "rabbit wire" having nearly 20 dB gain. At the high end of the amateur microwave range solid reflectors are probably most suitable. Surface accuracy which generally needs to be better than lambda/8, about 1/16 inch at 24 GHz, is commercially available in inexpensive dishes.
An upper limit on antenna size relates to propagation as described below.

Transmitter Power and ERP
To about the 100 milliwatt level, the cost of transmitter power is roughly constant through 10 GHz. To the 10 watt level it is approximately constant through 1300 MHz. Since using the minimum necessary power improves the network resource sharing, (not to mention being mandated by FCC regulation), it is desirable to use antenna gain, and higher frequencies rather than transmitter power to obtain the channel quality necessary for a given throughput rate. As other articles have already shown, transmit power of a few milliwatts is sufficient to communicate several Mbps. over a distance of 40 miles with 2 foot antennas[3].

Available Spectrum
In general, available bandwidth increases with increasing frequency. In most areas, the amateur bands above 450 MHz are underutilized. The largest contiguous amateur microwave band is presently at 10 GHz where 500 MHz is available. This is an extremely valuable resource which

should be used.   Present FCC regulations also promote use of higher frequencies for higher baud amateur communication.

Propagation

For effective use of resources it is necessary to have high quality, LOS paths. When none is available it will be desirable to break the path into multiple shorter paths. As shown in the PTP example, shorter links can produce a net improvement in performance. Presently many packet paths are not LOS. At vhf and lower frequencies, diffraction allows for some "bending" over obstacles and terrain. This is very expensive from a link budget point of view and in part accounts for much of the mediocre operation of hardware which should otherwise be capable of much better performance (typical 10-100 watt ERP and $>.25$ M^2 apertures). Even through the 24 GHz band, atmospheric attenuation is probably not a significant problem over the shorter path lengths which are desirable.

In situations where longer paths are unavoidable, perhaps on some backbone link connections, data rate may have to be decreased to ensure high probability of link function. The chief problem with longer paths is not the additional $20 \log(D)$ path loss so much as the variability in propagation through the troposphere. In well stirred air (as in parts of the Rocky Mountains) this is less of a problem but some areas, particularly near large bodies of water, exhibit great variability due to moist air masses. These variations on longer paths may make vhf-microwave DX interesting but they can require a lot of excess system margin to guarantee link availability.

These variations and the beam "de-steering" that tropospheric refraction can produce sets an upper limit on antenna gain. When the gain gets too high and the corresponding beamwidth gets too narrow the beam can be deflected from the intended direction. This sets an upper limit to available performance on a path. To achieve higher data rates the longer path must be broken up into multiple shorter paths which exhibit a smaller amount of deflection.

Other Considerations

In addition to the above constraints there are many other higher level issues which must be considered when designing and implementing physical and link layer network components to make optimum use of amateur resources.

From an efficiency point of view both mean square path length and total network path length need to be minimized. To do this, clusters need to be limited to a small number of members. However, operation of a local cluster may require the resources and cooperation of several individuals in the same way that local NBFM repeaters are constructed and maintained. A network of small high performance clusters requires more servers and more expensive higher performance hardware. Larger clusters require faster servers capable of processing many simultaneous higher layer communications.

Radio amateurs may be uniquely positioned to implement and maintain a wide area network. The physically distributed nature of the hobby's membership combined with radio communications skills and access to higher level sites in many locales may prove invaluable in creating such a network. To encourage amateurs with these resources to help construct a network may require some network applications which are particularly attractive to their interests. Applications supporting digital voice QSO, linking with other remote users and control and monitoring of existing analog hardware at high level sites may be attractive and help enlist their efforts in supporting a backbone of high speed hardware at high level sites.

For some amateurs, location may be such that no cluster server is available or easily constructed with resources at hand. In these instances it may be necessary that another amateur within LOS provide equipment full time to allow access to the network.

In general, higher user throughput data rates will require higher performance links and more coordination. Hardware, associated protocols and available user speeds may have to gradually migrate toward the ideal of small very high performance clusters. There will need to be ongoing access to such a network by entry level, lower speed users.

Perhaps above all else, the distributed and shared nature of a network demands organization and cooperation among amateurs. This may in fact be the most difficult task.

Local organization is required to install and maintain clusters and backbone connections. Between local clusters frequency coordination must be sufficient to allow good spectrum reuse, particularly as performance and channel bandwidths increase. Such coordination deals with a problem similar to the mapmaker's multi-colored map problem; how to assign a limited number of colors (channels) to a large number of areas (countries or states) so that adjacent domains do not have the same color.

Protocols which provide dynamic network routing without user involvement or knowledge need to be developed to allow for regional communication. Similarly national and international organization needs to oversee long distance connectivity. Satellites and other long haul links need to be developed and efficiently administered. It will likely take the combined resources of amateur radio to do this.

SUMMARY

The hobby of amateur radio nearly died at its inception. It was relegated to shorter "useless" wavelengths and left to die. Fortunately for amateurs the value of this spectrum combined with group cooperation and networking established the hobby in the face of this opposition.

If we are to continue as radio amateurs, the hobby must again achieve relevance within our culture through efficient use of our resources. In many ways we are uniquely positioned to do this but successful execution will demand attention to methods and a great deal of cooperation. The potential result is the birth of truly exciting new aspects to ham radio and the continuation of the king of hobbies.

References

1
 W. Lindsy and M. Simon, "Telecommunication Systems Engineering", Prentice-Hall 1973
2
 J. Kraus, "Antennas", McGraw-Hill Electrical and Electronic Engineering Series, 1950
3
 G. Elmore and K. Rowett, "Inexpensive Multi-Megabuad Microwave Data Link", Ham Radio Magazine, December 1989, pp 9-29