99% is good enough

By Graham Sharples, MD
Radiometrix Ltd

Radios are specified using a number of simple laboratory measurements which give an indication of the performance to be expected of them. Their eventual operating environment is usually far from benign and requires additional design considerations to achieve a reliable radio link. This paper examines the unpredictable nature of in-building propagation, looks at common sources of radio interference, and suggests diversity and redundancy techniques as a means of improving reliability.

The laboratory

The primary function of a transmitter is to generate RF power, usually as much as the regulations and cost constraints permit. The receiver is designed to detect as weak a signal as is possible i.e. have the greatest sensitivity. The path loss capability of the pair is the ratio of transmit power to receive sensitivity. A typical 433MHz transmitter of 10mW power output (+10dBm) and a matching receiver with a 2.2µV detection sensitivity (-100dBm) have a path loss capability of 110dB, i.e. they can overcome 110dB of attenuation.

The ideal world – free space

If we now connect this 433MHz transmitter and receiver to a pair of ideal isotropic antennas (0dB gain in all directions) and assume free space propagation (spreading losses only), we can calculate the free space range from:

$$\text{Range} = \frac{\lambda}{4\pi}\sqrt{\frac{P_{tx}}{P_{rx}}}$$

$$= \frac{23.87 \times 10^{\frac{L}{20}}}{f}$$

where R = range in metres
f = frequency in MHz
L = path loss in dB

This gives 17.4 km for our 433MHz example.

This figure is far higher than the 200 metres or so that can be expected as a working range in and around buildings, and serves to illustrate just how hostile the “real world” is.

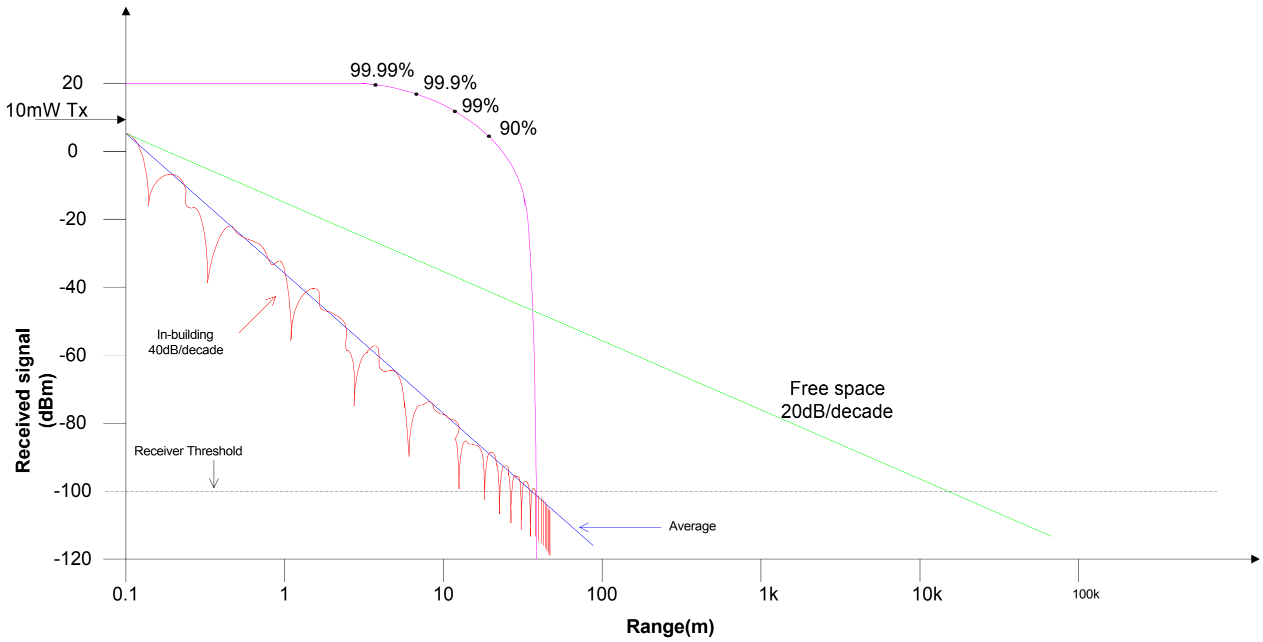
Propagation within a building

Signal propagation within a building is strongly dependent upon the topology, construction and content of the building and is influenced by the following:

1. Reflection from flat conducting surfaces such as metal cladding, galvanized roofing, foil backed plasterboard, metal coated anti-reflection glazing or any surfaces greater than a wavelength in size.

2. Re-radiation from thin conductors such as pipe work, electrical wiring, steel frame works and any conductor of greater than a half wave in length.

3. Absorption by lossy materials such as damp concrete, stonework and people.



Range curve for a 433MHz 10mW TX (unity gain antenna)

Multipath interference

Reflection and re-radiation of the signal causes a strong 3-dimensional standing wave pattern to be set up within the building. The signal strength at any particular point in space is determined by the sum in amplitude and phase of both the directly transmitted signal and all the passively re-radiated signals. It follows that in some positions signal cancellation will occur. These positions are known as “null spots” and appear as localised drops in signal strength when compared to the average strength in the surrounding space.

A receiver placed at random, has:

- a 10% probability of being in a >10dB null.
- a 1% probability of being in a > 20dB null.
- a 0.1% probability of being in a > 30dB null ……..etc.

This effect is bad enough, however it gets worse. The standing wave pattern will change - and with it, the position of the null spot - as the objects that contribute to it are moved. Some of these objects, such as metal furniture, filing cabinets, power cords etc, are moved infrequently. Others such as people, vehicles and ventilation fan blades, move rapidly and regularly. Perhaps the nastiest variable re-radiator is the fluorescent lighting tube – it behaves as a conductor which appears and disappears at twice the mains frequency and gives rise to "hum spots", which are null spots that have a 100Hz amplitude modulation. In many applications of in-building radio links either the transmitter or receiver or both are mobile, and may at any time be moved through a signal null.

Sometimes these effects are beneficial. For example, reflections between floor and ceiling in a reinforced concrete buildings act as a waveguide and will enhance propagation across a floor at the expense of vertical coverage. Re-radiation can often provide good coverage in areas which would otherwise lie in shadow from a direct signal. There may also be benefits in terms of antenna cross-polarisation losses - since the re-radiated signals tend to have indeterminate polarisation, there is no discernible need to orientate antennas in the same plane, nor does the antenna polar diagram have any significant importance since re-radiated signals are arriving from all directions.

From the foregoing it can be concluded that signal levels within a building cannot be determined with any degree of precision, but may only be expressed statistically in terms of averages and probabilities. There is always a finite possibility of exceeding the path loss capability of a radio link even at very short range.

Radio interference within a building

In many ways, local interference has the same effect upon a radio link as being in a propagation "shadow", i.e. loss of signal in a particular area. Depending upon the source the interference can vary from mild (e.g. 1 metre radius around a computer), to denial of the entire building where the interference is a strong on-frequency carrier. Unlike signal propagation nulls which are static or slow moving, interference is often intermittent. It may vary from occasional 'clicks' from light switches etc to a few minutes from a nearby cell-phone, or it may last the whole day whilst a computer is turned on.

Sources of interference to beware of:

Computers and other digital electronics can produce broadband noise and weak clock harmonics to 1GHz and above. It is worth noting that even EMC-approved equipment could still be legally radiating spurious signals that are 40-50dB above our example receiver's noise threshold.

An extremely common and particularly difficult variation on the above is interference from digital electronics within the product in which the receiver is used. Since the interfering source is usually within 5 to 20cm of the receive antenna and is always present, it masks all incoming signals below a certain level. The result is that the receiver is permanently "deaf".

Microwave ovens and industrial heaters - multiple unstable 2.4GHz carriers.

Switch mode power supplies - harmonics up to 100MHz and above.

Amateur radio transmissions on 433 MHz.

Other low power radio systems in the local area.

Strong near-frequency transmitters. Unlike all of the above, which occur on the frequency which the receiver has been designed to respond to, response to this type of interference is a common receiver weakness and depends heavily on its selectivity and strong signal handling abilities. It is becoming increasingly important - particularly at 868MHz with the adjacent cell phone band, and the introduction of TETRA at 410-430 MHz and latterly at 870MHz.

Designing for uncertainty

From the foregoing it can be seen that operating range within a building is both unpredictable and extremely variable. Since our aim is to design a reliable radio link with a reproducible working range, it is worthwhile to examine the various techniques available to improve reliability.

The simplest and by far the most common approach is to use excess signal levels (transmit power) to ensure that at the maximum working range the average signal level is at least 30dB above the receiver's detection threshold. This is simply checked by attenuating the transmitter output by 20dB and verifying at least 90% signal reliability at the desired range.

The figure of 30dB is chosen for a null probability of 0.1%, or conversely a 99.9% link reliability. A lower figure may be acceptable for an uncritical application such as a wireless door chime, or a manufacturer may require a higher safety margin for critical applications such as fire alarms or help call devices.

Excess signal above detection	Signal null probability	Link reliability	Range de-rating	Applications
0dB	>50%	<50%	1.0	car locking, toys
10dB	<10%	>90%	0.5	door chimes, DIY alarms
20dB	<1%	>99%	0.3	monitoring systems
30dB	<0.1%	>99.9%	0.2	professional telemetry
40dB	<0.01%	>99.99%	0.1	critical radio links

This method of de-rating the range or increasing TX power to gain reliability is both wasteful and anti-social, but it is simple. From the above it can be seen that methods to gain higher reliability without excessive signal level are of interest, particularly for more professional / critical radio links.

Redundancy and Diversity

From the simple null spot probabilities stated earlier it follows that if one receive antenna has a 1% probability of being in a >20dB null, then the probability of two receive antennas both being in nulls is 0.01%, or the same link reliability for 20dB less excess signal. Put another way, a threefold improvement in range (10 times coverage area) can be achieved.

The use of two antennas (and usually two receivers) in an "OR" configuration is known as Spatial Diversity. The antenna spacing and orientation is uncritical - provided it is sufficient to prevent significant mutual coupling and ensure that that both are not in the same null, any spacing from a quarter wave to many wavelengths works well. The technique may be extended to 3 or even more antennas / receivers "OR"ed together, however the law of diminishing returns applies.

Spatial diversity is economically most viable when used at the master or hub of a star network. Transmitter spatial diversity, where a message is sent using a combination of two or more separate transmitters / antennas, is also possible and provides similar benefits. In this case the message must be sent twice, first on one antenna then repeated on a second antenna sited in a different position to the first. Since the message is sent twice there is also some immunity to impulse interference.

Finally, transmit and receive diversity may be employed together in bi-directional links to achieve 99.99% reliability for only 10dB excess signal.

Time diversity is a commonly employed and very effective technique. Simply repeating a message several times with random off periods, or using bi-directional links with intelligent hand shaking, gives a high degree of immunity to impulse interference. Clearly, if the receiver is permanently in a null this method cannot improve signal reliability in a static environment - but if either end of the link is moving, time diversity has the same effect as spatial diversity in improving link reliability.

Frequency diversity is an excellent method of improving interference immunity. It can also provide a degree of spatial diversity, since the position of null spots is frequency dependent and with sufficient frequency shift a signal can be moved out of a null spot. Calculating the required shift can become quite complex. As a rough guide, for a quarter wave shift in null spot position (minimum effective) at a range of R metres:

$$\text{Required frequency shift} = \frac{300}{4 \times \text{R}}\ \text{MHz}\ .$$

This gives a figure of 1.875MHz shift at 40 metres range – just achievable in the 868MHz band and easily achieved at 2.4 GHz.

Finally

There is no such thing as a 100% reliable radio link. However, redundancy and diversity techniques can make considerable improvements to in-building link reliability and ensure a good reputation for your company's products.