Radio CBTC for the Las Vegas Monorail

C. Bantin & J. Siu
Alcatel, Canada

Abstract

In July 2004, the Las Vegas Monorail entered revenue service with a revolutionary new radio communication-based train control system (CBTC). The radio system is the world's first application of wireless technology in a revenue-earning urban rail transit system using antennas and over-the-air transmission. The system has performed perfectly and is a testament to the application of sophisticated technology to a challenging environment. The entire data communications system (DCS) serving the CBTC is based on the Ethernet 802.3 open-standard. The radio segment uses the 802.11 extension of this standard to mobile networks. Specifically the 802.11 frequency hopping spread spectrum (FHSS) variant is used because it has the best resistance to interference while providing more than adequate data rates. The entire network was designed and built using commercial off-the-shelf equipment, where the 802.11 standard has been implemented for mobile networks capable of routine hand-overs between access points at up to 100 km/h. The system design, using the standard in the 2.4 GHz ISM band, involved several detailed stages to ensure faultless operation. First the nature of the CBTC data has to be understood and the network has to be designed to accommodate this data. Next the link budgets have to be determined as applied to the unique guideway environment used by the monorail. Signal strength surveys were carried out to determine where to place access points such that there was full and redundant coverage throughout the guideway. Furthermore severe co-user interference (802.11b or WiFi) is present at several locations along the guideway and had to be accommodated. Finally the operating parameters of the radios had to be fully understood to optimize the hand-over process. The complete system was extensively tested and many weeks of data was collected in order to verify that the radio system met the performance requirements.
Keywords: radio CBTC, ATC, driverless, monorail.

1 Introduction

In July 2004, a new monorail train entered revenue service in Las Vegas carrying passengers between hotels along the East side of the Strip. The route extends from the Sahara hotel at the North end to the MGM Grand hotel at the South end for a total length of approximately 7 km. This monorail train is unique in that it is both driverless and uses an automatic train control (ATC) system based on radio communications between the trains and control equipment distributed along the guideway. The recent adoption of radio for communications-based train control (CBTC)[1] has provided new opportunities for low cost, flexible, extensible ATC systems. In the past, radiating cable and waveguide systems have been used for CBTC communications, but these solutions are restrictive, difficult to deploy and expensive to install and maintain. Other free-space radio links have been set up in trial situations, but the Las Vegas Monorail system is the first to achieve revenue status. The Las Vegas monorail ATC uses a data communications system (DCS) based on the IEEE 802.3 Ethernet standard. The radio segment of the DCS uses the IEEE 802.11 extension of this standard in the 2.4 GHz ISM band [2,3]. More specifically, it uses the frequency hopping spread spectrum (FHSS) variant of this standard. The FHSS system has more than enough data capacity, and offers superior resistance to interference from both co-users of the ISM band and other sources. The performance of the DCS network, and particularly the radio portion, has been extensively measured, and the network does not impose any constraints on the performance of the ATC.

2 System design

The philosophy for designing and implementing the DCS network is to employ open standards and to use commercial off-the-shelf (COTS) equipment [4]. The advantages of this approach are cost, performance, availability of equipment from multiple sources, and the potential for interoperability. The choice of standard is based on the best type of network that meets the operational requirements of the ATC. The ATC system communicates with control telegrams, in the form of data packets, which are exchanged between the control equipment and the trains. These data packets contain vital commands to control the trains and to report on the status of each train. Although the commands themselves may be vital, meaning their correct interpretation and implementation is a matter of safety, the DCS is a non-vital network. Its principal characteristics are throughput, which is the achievable average data transfer rate, and latency, which is the delay between transmission and reception. Corruption of the telegram data is detected through extensive checks within the ATC equipment itself. Over the radio links, data encryption is used to protect the network against external intrusion. Except for emulation, the worst consequence of intrusion through the air interface is a denial of service (DNS) over the particular link involved. It can be though of as an extended period of missing data packets. While this could potentially affect operations, it is not, in itself, a safety concern.

The threat of emulation, on the other hand, is a safety concern, and the defense against this to use a secure authentication process for all communications [5].

The ATC data packets contain the telegram information and data integrity codes (such as CRC checks). They also have an appended authentication code, which is a hash sequence, using the IPSec protocol [6], based on the data contents of the packet and a secret key. These data packets are exchanged between control equipment and the trains on a poll-response basis. Typically there are three polls per second from the control equipment and three responses from each train. The data packets are about 120 bytes long; therefore the overall data throughput required for this exchange is well under 10 kb/s per train. The 802.3 packet transmission protocol is well suited to this type of communications and satisfies the requirement for open standards and COTS equipment. An accompanying compatible standard for radio communications, 802.11, is for radio LANs and provides explicit support for mobility. An additional advantage of the 802.3 or 11 standard is that it uses contention-based access [2]. There are three main areas of where contention-based access is preferred:

1) **Data delivery**. With contention access packet delivery is guaranteed at the expense of added latency. The MAC protocol permits retransmission of the packet following collisions. This is a self-regulating mechanism and ensures that all contending parties eventually get their packets through. By contrast, with a deterministic access scheme circuits have to be assigned. When a request is received it must be processed by the network, and there is always a chance of not having a circuit available.
2) **Data integrity**. Guaranteed delivery also means error-free packet delivery. Errors in a transmitted packet are treated just as if there was a collision and the retransmission protocol takes over. With deterministic access there is no opportunity to resend data when there are errors in the transmission (at least at the physical or link layer) and an error control scheme, such as forward error correction (FEC), is needed to minimize packets containing errors. Residual errors have to be dealt with in the higher layers and this involves more complexity and a lot more overall latency.
3) **Interference Protection**. Retransmission protocols can be used very effectively to combat interference. The frequency-hopping scheme is particularly effective at this. With deterministic access there is no such mechanism unless it is dealt with in the higher layers with the attendant complexity and added latency.

A significant advantage to using open standards in general, and 802.11 in particular, is that the license-exempt frequency bands can be utilized [3]. Many regulatory authorities set aside common frequency bands for shared use, without requiring specific licenses. The equipment using these bands must obey certain restrictions, such as maximum transmit power levels, and must be certified to meet these restrictions. The 802.11 contention-based access methods allow sharing the band in an orderly manner, and the spread spectrum techniques utilize the band to full advantage. Licensed bands, on the other hand, may appear

to offer the advantage of having no co-users and is interference free. However, an intentional intruder will not respect the license conditions and the licensee is left with no protection. The radio-based DCS uses the 2.4 GHz ISM band. Furthermore, the FHSS variant utilizes this band to the best advantage because it occupies the full available bandwidth during its hopping sequences.

3 Safety and security

For security purposes the DCS is considered to be an open network because it is relatively easy for an intruder to attempt to gain access through the air interface [7]. This means that the entire network must be considered untrustworthy and is subject to malicious attacks. Of these attacks, the most straightforward is a DNS attempt launched by trying to associate with one or more radio access point or to simply produce enough interference to prevent any legitimate radio from associating with it. The DCS network provides several defenses against DNS attacks. In the link design, the high-gain narrow-beam antennas provide some protection against intrusion from external sources through antenna pattern discrimination. In addition, the wanted signal levels from the access point and mobile radio are kept as high as practical to minimize any impact of unwanted interfering signals. The most important defense, however, is redundancy. Each mobile radio can associate with two or more access points. Therefore individual access points that are subjected to deliberate interference, or even a successful DNS attack, can be avoided without affecting the network performance. Likewise, each end of a train has an independent mobile radio. If all the access points seen by one end are simultaneously blocked, or if the mobile radio itself is subjected to interference, the radio at the other end of the train can continue to provide network connections. Note that even in the instance of a successful DNS attack over the air the worst consequence is to cause one train to stop, and this in itself does not represent a safety hazard. The consequences of an emulation attack, however, are much more severe than for DNS. A safety hazard can exist with a successful emulation attack because legitimate train control commands can be sent and received. The DCS network has several layers of defense against emulation. First, it is difficult for an intruder to gain access to the network over the air. If necessary, the frequency hopping sequences can be made non-standard, making it more difficult to follow the signal. Second, authorization by password is required in order to associate with an access point. Third, and most powerful by far, is the use of end-to-end authentication over the DCS network through authentication gateway devices. These devices provide IPSec security with IKE dynamic key management [6]. This allows ATC components to communicate with each other by creating VPN (virtual private network) tunnels through the DCS network.

4 Radio network

The radio segment of the DCS network consists of a series of access points deployed along the monorail guideway. The train mobile radio associates with

these access points as the train travels along the guideway, forming radio links and therefore data paths over the network to the wayside. There are two possible links from each train at all times, one from each end. According to the 802.11 MAC protocol, the mobile radio makes the determination to hand-over communications from one access point to another based on the strength of beacon signals that it sees from the various access points in its neighbourhood. The design challenge is to provide the appropriate RF signal environment for consistent and smooth hand-over along the route. Therefore, the locations of the access points were carefully selected so as to provide continuous and redundant RF coverage to each end of a train at all points along the guideway. This coverage ensures a minimum signal level consistent with the threshold settings of the radios. The access point antennas along open sections of the guideway are mounted on poles between the guideway beams. Two high-gain panel antennas are mounted at the top of the pole, one pointing in each direction along the guideway and serving both guideway beams. The height of the poles above the guideway beams is approximately 3 m, enough to allow line-of-site propagation over the top of the trains. The access point electronics are mounted in an environmentally controlled cabinet near the base of the pole beside the guideway beam. The mobile radio antennas are medium-gain sector antennas mounted in pairs at each end of the train. Each pair provides diversity reception for the mobile radio. The main beam faces directly forward away from the train, therefore the antenna pattern is essentially unaffected by the train itself or its occupants because of the very low back and side-lobes. With the sector antennas mounted in a vertical (physical) orientation, they have a narrow 25-degree beamwidth in elevation, and a wide 120-degree beamwidth in azimuth. Therefore they can see access points even while the train is negotiating curves. They also provide a gain of 12 dBi while maintaining good discrimination against interference.

The radio equipment consists of 802.11 frequency-hopping radios connected to bi-directional amplifiers and high-gain antennas. For the access point, a single amplifier is connected to two 16 dBi panel antennas through a splitter/combiner. An effective isotropic radiated power (EIRP) of +34 dBm (the maximum allowed is +36 dBm [3]) is achieved from both antennas with a 250 mW amplifier power output. For the mobile radio, two independent antennas and their associated bi-directional amplifiers are connected to the radio to provide the propagation path diversity. The radio automatically selects the strongest signal from the currently associated access point via the separate antennas. An EIRP of +34 dBm is achieved from either sector panel antenna with a 250 mW amplifier and a 12 dB antenna gain. The receive path net gain, taking into account the antennas, splitters and cable losses, is designed to be 20 dB, at each end. Therefore, because of reciprocity, both radios receive the same signal power when the corresponding transmitter uses the maximum EIRP. It is a necessary condition that the links be balanced in this fashion because the hand-over functions could not operate properly if the mobile radio could hear the access point while the access point could not yet hear the mobile radio, or vice versa. The received signal power is set to be no less than –50 dBm at any point on the

guideway by carefully placing the access points. Although the sensitivity of the radio is a lot lower than this (at –80 dBm [2]), a large margin is used to ensure a signal coverage that was above most of the expected interference levels as well as to allow for signal fluctuations. In other applications, where interference is not as severe, a signal threshold of –65 dBm can be used. This provides 15 dB of margin, and is sufficient to accommodate temporal and spatial signal fluctuations.

Antenna deployment along straight sections of guideway is straightforward. The main propagation considerations are reflections from the ground (guideway) and any nearby buildings or objects such as billboards. An example of the signal from one access point as seen by the train receiver (before the receive amplifier) is shown in figure 1.

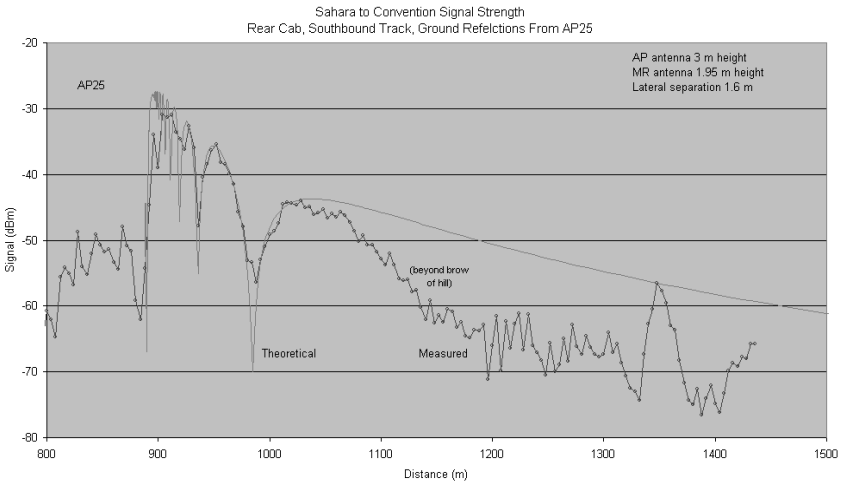

Figure 1: Signal strength example.

This figure illustrates the effects of a single ground reflection and diffraction over the brow of a shallow hill. In these straight guideway situations, the 16 dBi panel antenna gain provides a relatively narrow 32 degree beamwidth for discrimination against external interference. For curved sections of guideway a wider azimuth beam width is needed and the antenna pointing is offset around the bend in the guideway. Antenna deployment in stations is more complex because of the station design. The platform is between the guideway beams and there is a wall separating the two directions of travel. In these cases there is an access point located at each end of the station with a special antenna arrangement. One antenna is located between the beams and pointed away from the station. In the direction into the station there are two antennas, one pointing down the beam on each side of the station. These antennas are fed by a second splitter, therefore each one has 3.5 dB less EIRP than the single antennas pointing down the guideway, and similarly 3.5 dB less gain on the receive path.

5 Network and radio performance

The on-board mobile radio (from either train end) associates with the access points roughly in sequence as the train moves along the guideway. However, a mobile radio is in no way obliged to associate with each access point it passes. This is an essential feature of the protocol. The decision to re-associate, or hand-over between access points, is made by the mobile radio based on the signal level it sees and not by the access point or by any other element of the network. In this way the network is not burdened with the need to keep track of where each train is located and where it is going. The hand-over algorithms in the mobile radio are one of the most critical aspects of the radio design. The 802.11 standard provides a number of MAC layer services that a radio manufacturer may use in order to implement the hand-over process, although it does not explicitly define this process (802.11e is attempting to do this). The mobile radios make full use of these services and can manage seamless hand-over at speeds over 100 km/h. The most consistent and repeatable hand-over sequences are obtained when the decision to leave an existing access point and the decision to join a new access point are the same, and equal to the value of minimum signal level coverage obtained along the guideway by the choice of access point locations. At any time each mobile radio can see at least two access points that meet this criterion.

Interference in the 2.4 GHz ISM band can be very severe in Las Vegas. A sweep of the route reveals numerous 802.11b (WiFi) hot spots and several fixed links. A lot of these are located in and around the Convention Center and at least one is located on the mezzanine roof of the Convention Center station. The FHSS implementation used for the DCS radios is particularly well suited for dealing with this interference and indeed other kinds of interference, intentional or otherwise. There are three main mechanisms used for dealing with interference:

1) **Ignoring**. If the level of the interfering signal is sufficiently low the effect on the performance is small. If errors in the data packet do occur as a result of this low-level interference, then the packet is resent (initially in the same dwell period, or same hop frequency). The level below which the interference is ignored is set within a radio by the same parameter that is used to determine when the channel is busy.
2) **Avoiding**. The most powerful mechanism that is available to combat interference is to avoid it. The FHSS scheme is well suited to do this. If the channel is deemed busy or if a data packet is consistently received in error then it is resent on the next frequency in the hop sequence. The minimum spacing between frequencies in the hop sequence is 6 MHz. It can also be much larger than this. This means there is a good chance of moving out of the band containing the interference for the next hop frequency, thus avoiding the interference.
3) **Contending**. The FHSS scheme shares the same MAC protocol as other 802.11 users of the band. Therefore, even if there is co-user interference, the contention access process will ensure the data packet gets through, albeit with some added latency. It is also possible,

however, that the frequency will have hopped to a new value before the retransmission occurs (i.e. mechanism 2).

The overall performance of the DCS network can be characterized by the data throughput and latency parameters. These parameters are related since throughput is usually specified as the effective data rate at a maximum acceptable latency. Therefore, in order to determine the performance of the DCS, extensive measurements were made to characterize the latency. These measurements were done by measuring the round trip delay incurred by sending 200 byte packets at 350 ms intervals from the trusted network interface on a train (from both ends and traveling in both directions) to a trusted network port on one of the switches at the central control site. The link thus includes the authentication gateway devices. The test data was sent simultaneously with the train control data. Data was taken during live runs, over many hours, along the entire guideway route, and from each end of a train traveling in both directions. The results are illustrated in figure 2. They show that more than 95% of the data packets made the round trip in less than the 350 ms transmit interval and 97.5% in less than 700 ms.

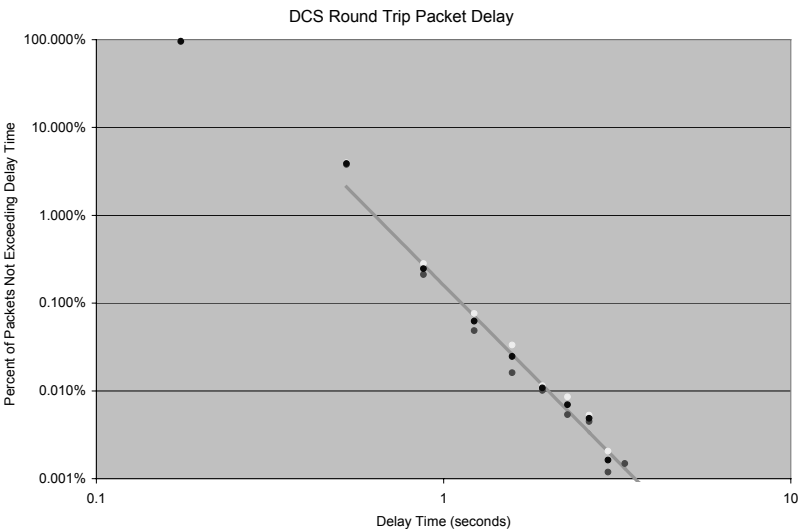

Figure 2: Round trip packet delay.

At the tail-end of the distribution there are instances of delays up to 3 seconds or more. These are low probability events involving less than .001% of the transmitted packets, and are due to occasional hand-over difficulties or temporal signal fluctuations due to propagation effects. As an example of the effect of these delays, consider the case for trains running 18 hours a day. Only one end of train is actively communicating at any time, and if there are 3 telegrams per

second to this train-end then, on average, there are 2 instances per day when a telegram experiences a delay of more than 3 seconds. The performance of the ATC is not adversely affected by these occasional delays because there are built-in algorithms to handle gaps in communications. In the extreme case a loss of communications for greater than 3 seconds will result in a train slowing to a stop, but automatically recovers when communications is restored, even before the train comes to a halt.

6 Conclusions

The DCS network in Las Vegas is part of the first radio-based CBTC system in the world to achieve revenue status. The network fully meets the availability and performance requirements to support driverless ATC operation. The use of 802.3 standards-based network solutions has resulted in the best choice for the network and particularly the radio standard. Further, the choice of 802.11 frequency-hopping for the radio segment has resulted in the best solution for dealing with interference. The use of data encryption adds privacy over the network, and the use of IPSec authentication adds the essential defense against emulation.

References

[1] Sullivan, T., *Radios for CBTC,* Railway Age Magazine C&S Buyers Guide2005.

[2] IEEE Standards, *Part 11:Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications,* ANS/IEEE Std 802.11, 1999 Edition (R2003).

[3] The FCC Rules and Regulations, *Operation within the bands 902-928 MHz, 2400-2483.5 MHz, and 5725-5850 MHz,* Part 15, Sec. 15.247.

[4] Alcatel TAS, *Open Standard CBTC Wireless Technology*, Alcatel Transport Solutions Division newsletter, Vol. 1, Issue 2, pp 1-2, April 2003.

[5] European Standard, Railway Applications: *Part 2: Requirements for Safety-Related Communications in Open Transmission Systems,* EN 50159-2.

[6] Internet Engineering Task Force (IETF), *Internet Security Protocol (IPSec).*

[7] Alcatel TAS, *Radio Communications – Securing 802.11 Safely,* Alcatel Transport Solutions Division newsletter, Issue 4, pp 1-3, January 2004.