Bixby_1999

HIGH FREQUENCY DIGITAL SURFACE COMMUNICATIONS

G. Edward Bixby, AK0X

In Memory of Raymond Petit, W7GHM
my long term friend and inventor of
CLOVER

Paraphrasing Robert Zubrin in his book THE CASE FOR MARS:

- Satellite communications will be very expensive and therefore not a viable option for some time to come.
- Line of sight (VHF & UHF) surface communications will be limited because of the smaller diameter of Mars.
- Mars has an ionosphere, enabling global surface to surface communication in the short-wave radio bands.
- SSB HF communication should be possible at about 4 MHz during the day and at 0.7 MHz at night.
- Mars' ionosphere extends upwards from an altitude of about 120 KM but is generally weaker than earth's.

PREMISE: Scientists exploring Mars will not be satisfied with simple voice communication. They will want to send error free digital communications permitting direct computer file transfer of technical data, including pictures, data files and spread sheets as well as fax.

Obviously, line of sight VHF and UHF communication distances can be considerably enhanced by the use of repeaters strategically placed on high points. Such repeaters, in common amateur radio use, receive on one frequency and automatically re-broadcast received communications on another frequency. The dual receive/transmit frequencies are necessary so that the repeater doesn't hear only itself in a severe feed back loop. Since they are placed at high altitudes, the line-of-sight distances are extremely enhanced. You can imagine the range of a repeater placed on top of Olympus Mons! But this talk is about HF communications employing ionospheric propagation.

First, I would like to make some general comments about single sideband (SSB) high frequency (HF) communications which utilize the reflective properties of earth's ionosphere for signal propagation instead of the more common line-of-sight transmissions. Higher frequencies pass more easily through ionized layers (there may be many) and suffer less scattering and in general less distortion. They also suffer less from static and multi-path, interference caused by a signal taking more than one path from transmitter to receiver. Since the path lengths are likely to be different, signals taking different paths in general arrive at the receiver at slightly different times resulting in distortion similar to ghosting on a TV when roof top antennas are used instead of cable. Because of these problems, simply increasing power won't necessarily result in the expected increased signal quality or, in the case of digital communications, increased data rates.

The earth's ionosphere actually consists of multiple layers at different heights which wax and wane with the day/night cycle, the eleven year solar cycle, solar storms, and some say even the weather as well as the geographic location of the communicating stations. These layers extend from a low of 48 KM to over 300 KM. The higher layers are primarily responsible for worldwide communications while the lower ones tend to scatter, and therefore weaken, the signal during the day. At night, however, the lower layers mostly disappear while the upper ones are significantly weakened. Therefore, during the day frequencies of 14 MHz or up are most useful while at night frequencies as low as 1.8 MHz or lower can be heard around the world. The exact nature of the propagation is strongly dependent on the solar cycle as well and sometimes varies considerably within a few minutes. Zubrin implies Mars has a much simpler single low lying but weak ionospheric layer that may mitigate these complications.

Another characteristic of ionospheric propagation is what is known as "skip zones". The signal bounces off the ionosphere and returns to earth at some distant point. At distances beyond the ground wave (essentially line-of-sight propagation) and less than the return point of the sky wave, no signal is heard. Signals skip over these intermediate distances. This process can continue, with the sky wave reflecting off the earth's surface and again off the ionosphere to return again at some even more distant point. This process can continue over and over again, sometimes completely circling the globe with, of course, significant attenuation of the signal at each hop.

Another phenomenon that occurs is known as "ducting". A transmission can be continuously reflected between two ionization layers, traveling a significant distance around the earth before returning to the surface where an uncharacteristically strong signal may be heard.

In general, if we want long distance communication, a low angle of radiation is desirable; one that essentially radiates strongly parallel to the earth so that it reaches into the ionosphere only as the surface of the earth curves away. This results in the fewest number of "hops" to a distant point. Now, if local communication (up to say 300 KM) is desired, lower frequencies are chosen which bounce off the lower layers. Antenna designs that produce nearly vertical radiation (straight up) and therefore effectively illuminate these lower overhead ionization layers are most effective. These antennas tend to be horizontal wires such as a dipole placed close to the ground (less than 1/8th of a wavelength) where the reflected waves from the ground combine with the original wave from the antenna to produce a strong vertical component.

The height of the Martian ionosphere, according to Zubrin, extends upwards from about 120 KM roughly corresponds to one of earth's lower layers, the "E", layer which strongly reflects lower frequencies during the day. Now, if the communication we desire on Mars is from a local base to roving vehicles, standard vertical antennas used for mobile operations will probably not be as effective as horizontal antennas. Horizontal loops placed on the roof of the rover and low long wire dipoles at the base will probably provide optimum performance.

The Amateur Radio community has developed and utilized global HF SSB communication technology for decades. Amateur bands span frequencies from 1.8 MHz to 1300 MHz and multi-band 100 Watt transceivers suitable for mobile voice and digital communications are widely in use. Digital technology in the form of CW (Morse code) and RTTY (Radio Teletype) have been widely used for years. More recently, however, more modern techniques such as Amtor, Packet, Pactor, CLOVER, PSK31, BPSK, MT63 and MT-Hell have been developed. Of these, in the author's opinion, only two are suitable for error free high speed SSB HF digital communication of computer files and keyboard conversation: CLOVER and Pactor. There are two versions of both modes, CLOVER II, CLOVER-2000 and Pactor I and Pactor II. Pactor technology was developed in Germany while CLOVER is strictly a US product invented by my friend Raymond Petit, W7GHM. Since I am more familiar with (and closely associated with the development of) CLOVER, I will dedicate the remainder of this paper to CLOVER technology and the contribution it has for digital HF SSB communications.

The HF environment is inherently noisy with problems with static, multi-path, and interference with other stations a constant problem. Although I doubt that interference from other stations on Mars will be a problem for some time I suspect the other two limitations will apply. CLOVER has been especially developed to operate in this noisy environment by using a combination of modern technologies including a very narrow bandwidth wave form, Reed Solomon error control coding, an adaptive protocol based on the quality of the signal path, and digital signal processors. CLOVER hardware consists of a single long slot IBM compatible computer board and three wires to a standard SSB transceiver: receive audio, transmit audio, and a push-to-talk line to key the transmitter. Simple menu driven software makes use of CLOVER very easy to learn and other coding schemes come with the software permitting use of RTTY, Amtor, and Pactor I (but not Pactor II) as well.

Let's look at this technology more closely. A narrow bandwidth waveform consisting of Dolph-Chebychev pulses spaced 32 ms apart is used. These smooth pulses are the envelope of the audio frequency tones at about 2 kHz generated by the digital signal processor. The amplitude of the tone is slowly changed from zero to a maximum and back to zero, forming the individual pulses. The bandwidth of a single frequency train of pulses is about 100 Hz at -60 dB. This narrow bandwidth enhances the signal to noise ratio and therefore performance.

Data is encoded in the pulse string by a combination of changing adjacent pulse amplitudes or a phase change of the audio tone at the point in time when the amplitude is zero. By performing the phase change at zero amplitude, the signal is not broadened as it would be in standard phase shift modulation technology.

Let's look at this phase modulation a little more closely. The audio tone is essentially a sine wave which has both an amplitude (how big it is) and a phase (when it passes through zero amplitude or when it starts, however one wishes to think of it). By changing the phase between adjacent pulses, by 0 or 180 degrees for example, data can be passed. Zero phase change can represent a binary zero and the 180-degree phase change can represent a binary one. Thus, one bit of data is sent every 32 ms for a base data rate of about 31 bits/sec. This is called binary phase shift modulation (BPSM) since two phase angles are used. BPSM is the base data rate of CLOVER.

But, by using four phase angles (0, 90, 180 & 270 representing the binary 00, 01, 10 and 11 states) two data bits can be sent every 32 ms doubling the data rate to about 62 bits/sec. This is called QPSM, quaternary phase shift modulation. CLOVER got its name from this modulation when Ray displayed QPSM with noise on an oscilloscope and his wife Joyce exclaimed: "It looks like a clover leaf!".

CLOVER also uses 8PSM which adds another data bit every 32 ms as well as 8P2A in which two levels of pulse amplitude add another data bit. The highest data rate is achieved with a 16P4A modulation wave form which adds two more data bits every 32 ms. Thus, 16P4A provides a data rate of 6 bits every 32 ms or about 187 bits/sec.

However, CLOVER doesn't stop there. Four audio tones with a separation of 100 Hz are generated, separated in time by 8 ms for a repetition period of 32 ms, the basic pulse period. Each tone carries independent information so the data rates quoted above are multiplied by four. The resultant signal is now about 500 Hz wide still at -60 dB and exhibits four distinct peaks at the four audio frequencies located in a central plateau in frequency space. This distinctive waveform is independent of which modulation (BPSM, 16P4A etc.) is being used although the time trace of the various modulation techniques (and the sound heard by the ear) varies considerably. Thus, CLOVER has a very unusual characteristic in that its bandwidth is independent of data rate! This 500 Hz signal is easily handled by modern amateur HF SSB transceivers and is narrower than normal voice SSB communications and a relatively low-level duty cycle mode. Therefore, the transceivers' built-in filters can be effectively used to further improve the signal to noise ratio.

Actually, there are two distinctly different versions of CLOVER. The one described above is used by the amateur community and has the 500 Hz bandwidth stated above. But a commercial version has also been developed which takes full advantage of the 2 kHz bandwidth assigned to commercial HF users. The pulse interval is shortened to 16 ms and eight audio tones are used in place of the four, resulting in a full 2 kHz CLOVER bandwidth. Thus, the data rate is four times faster in this commercial version over the amateur version. This is called the CLOVER-2000 (for the 2 kHz bandwidth) protocol.

But which modulation technique does CLOVER actually use? As mentioned previously, BPSM is the base or default modulation used. Although the slowest, it is the most robust mode in overcoming a noisy environment. Multi-path, for example, can smear out the measured phase changes and noise can distort the amplitude modes. But the measured phase change has to be off by more than plus/minus 90 degrees before a data error is made and BPSM doesn't use amplitude modulation. Obviously, smaller errors in measured phase changes will result in corrupted data for the higher data rate modes. However, BPSM is fast enough for normal keyboard conversation under normal conditions.

But what if higher data rates are needed, like when one station wants to send a data file such as a JPEG picture? CLOVER continuously monitors the quality of the communication path between two conversing stations in both directions and keeps both stations informed of that quality by automatically encoding the information in what are called Clover Control Blocks (CCBs). Measures of the signal to noise ratio (SNR), phase distortion (PHS), modulation type, and transmitter power are continuously transmitted to both stations for both directions of communication. Thus, when a large amount of data needs to be transmitted, CLOVER automatically makes a decision of what data rate is to be used based on the quality of the communication path. If the path is very poor, communication may proceed in the base BPSM because attempts at higher data rates would result in nothing but errors with no transmission of data. But if the path is good, CLOVER shifts into high gear and data can really fly when the higher data rates are possible. Data rates in the two directions are independent. One station can be sending data in BPSM just acknowledging correct receipt of the data while the other is sending in 16P4A! The CCBs are always sent in BPSM to minimize errors.

This technique is called an **adaptive protocol** and no other modulation known to this author uses it, certainly none suitable for the HF environment. A word of caution, however. Because of this very difficult environment, attainable data rates are much below those of hard-line techniques used on the Internet or UHF satellite based communication. A data rate table for the CLOVER-2000 protocol is included later in this paper.

I have talked a lot about controlling errors but how is this actually accomplished? Many people are familiar with parity checking, the technique of using an 8-bit byte consisting of 7 data bits and the 8th as the parity bit. If there are an even number of 1s in the 7 data bits, for example, the 8th parity bit is left a 0 so the number of 1s remains an even number (thus the name even parity). If there are an uneven number of 1s in the 7 data bits, the parity bit is set to 1, again resulting in an even number of 1s. This method can detect an error if one bit, for example, is changed from a 0 to a 1. It can detect errors which change an uneven number of bits but not if an even number of bits are in error nor can it correct the error, but only claim that an error has occurred.

But more sophisticated mathematical techniques can be applied. Additional error control bits can be added such that a rigorous methodology can be applied to the received data so that, if not too many errors are made, the exact bits in error can be determined and reversed and the data therefore corrected. CLOVER uses one of these techniques called Reed Solomon Error Control Coding. Three different levels are available: Robust in which 60% of the bits are data with the remaining 40% are error control bits, Normal 75% coding, and Fast in which 90% of the bits are data. If all errors can be corrected, re-transmission of the data block is not required. If the data cannot be corrected, the receiving station automatically requests re-transmission of the data block. A Check Sum Comparison is also made at the end of each data block to ensure absolute data accuracy.

This is the Automatic Repeat Request (ARQ) mode in which two stations are talking exclusively to each other. One station transmits for two seconds, and then the other station transmits for two seconds in the basic BPSM keyboarding mode so data flows almost simultaneously in both directions making it very similar to a full duplex link although only one radio frequency is used. This bi-directional link is also unique to CLOVER. If a station has data to pass in which the higher data rates are to be used, its transmission block is automatically extended to about 20 seconds with the receive station acknowledging in 2 second blocks. If both stations have data to send, both will go to the 20-second blocks and may even use different modulation techniques if the path quality is different in the two directions! In other methods, communication is uni-directional with the receiving station only acknowledging proper receipt of the data. A “hand-over” code is transmitted when it is the other station’s turn to talk.

A broadcast mode called Forward Error Control (FEC) is also available in which one station can broadcast data to several other stations simultaneously. Only one station transmits with the other stations passively listening. The error control protocol is fully active but repeat requests for errors which cannot be corrected is not available in this mode. The operator of the transmitting station chooses the modulation method and amount of coding to use at the beginning of the transmission. A “dual diverse” mode is available in FEC which provides even more robust communication. In this mode, tones 1 and 3 (in the amateur version) send the same data, and tones 2 and 4 send identical (but different from tones 1 and 3) data in BPSM format. Thus, two chances of decoding correct data are available, the best one being chosen of course.

As can be imagined by now, the CLOVER signal is very complex and the encode/decode software logic very complicated. But it works! And this complexity yields one very nice benefit: an ARQ communication between two stations is quite secure from outside listeners. While someone with a CLOVER board can copy bits and pieces of the communication, data from both stations will be intermingled and of course non-correctable data errors cannot be corrected by a repeat request from this third party. Without the CLOVER hardware and software, correct encoding of the communication by any third party station is virtually impossible!

All this data encoding and decoding is accomplished by the digital signal processor which generates the transmitted audio wave form as well as employs very strong digital audio filtering of the received signal in addition to that filtering provided by the receiver. The PCI-4000 digital signal processors are the DSP 56001 24-bit processor with an 8Kbyte x 24 bit RAM and a 68EC000 16-bit processor with 32 Kbytes x 16 bit RAM and 16Kbyte x 16 bit ROM. At least I hope I got all that right. Software is downloaded to the PCI-4000 board from the host CPU disk file programs. The software is menu driven with multiple screens and “magic keys” are available for ease of use. Computer files can be selected and transmitted easily by choosing a file handling facility and identifying the appropriate path. A data compression technique is automatically enabled for transmission of binary files.

I have personally used CLOVER to talk to other amateur operators located throughout the US and other operators have communicated around the world with it. I have received data files of a few hundred kilobytes absolutely error free although it might take an hour or two if the path is poor. But this is the reality of HF SSB communications: not nearly as good as modern internet data rates but they can be error free and I believe CLOVER is the optimum solution available today.

I am convinced that no better digital communication method for the HF environment is available anywhere.

In his book 2001 A SPACE ODYSSEY, Arthur C. Clark has the space ship on-board computer, HAL, identify himself with the words:

“ I am a HAL Nine Thousand computer Production Number 3. I became operational at the Hal Plant in Urbana, Illinois, on January 12, 1997.”

CLOVER is a proprietary product of: HAL COMMUNICATIONS CORPORATION
Located in: URBANA, ILLINOIS

CLOVER was introduced to the amateur radio world at the Dayton, Ohio Hamvention in 1993.

With karma like that,

CLOVER BELONGS ON MARS

CLOVER-2000 PROTOCOL

ROBUST BIAS (60 % CODING)

MODULATION TYPE	BYTES PER FRAME	MAXIMUM ERRORS	BLOCKS PER FRAME	DATA RATE BYTES/SECOND
16P4A	900	300	6	165.4
8P2A	600	200	4	110.3
8PSM	450	150	3	82.7
QPSM	300	100	2	55.1
BPSM	150	50	1	27.6

NORMAL BIAS (75 % CODING)

MODULATION TYPE	BYTES PER FRAME	MAXIMUM ERRORS	BLOCKS PER FRAME	DATA RATE BYTES/SECOND
16P4A	1128	186	6	207.4
8P2A	752	124	4	138.2
8PSM	564	93	3	103.7
QPSM	376	62	2	69.1
BPSM	188	31	1	34.6

FAST BIAS (90 % CODING)

MODULATION TYPE	BYTES PER FRAME	MAXIMUM ERRORS	BLOCKS PER FRAME	DATA RATE BYTES/SECOND
16P4A	1356	72	6	249.3
8P2A	904	48	4	166.2
8PSM	678	36	3	124.6
QPSM	452	24	2	83.1
BPSM	226	12	1	41.5

ARQ FRAME TIME: 5.44 SECONDS