34th Space Symposium, Technical Track, Colorado Springs, Colorado, United States of America
Presented on April 16, 2018

PROTECTED SATELLITE COMMAND AND CONTROL (C2) WAVEFORMS AND ENHANCED SATELLITE RESILIENCY

Bryan Butler
Kratos RT Logic, butler@rtlogic.com

ABSTRACT

Satellites in earth orbit were once considered relatively safe. However, these assets grow more vulnerable as adversaries are increasingly displaying both the intent and the means to compromise and cripple space capabilities in their efforts to challenge our strategic advantage. In addition, congested spectrum is leading to more instances of benign interference. These vulnerabilities are driving change in military space and countering them requires more resilience, agility, and speed in order to predict, pre-empt, and prevent the growing range of threats.

Among counter-measures under consideration is the concept of building protection into the command and control link for satellite programs. A case is made, based on the evolution and emergence of several competitor nation-states, that our satellite systems need superior protection and resilience in their command and control (C2) capability. Techniques for achieving that protection and resilience using modern waveforms and technology are presented. The security implications of protected C2 on both ground and space segments are described, followed by a discussion of how such a capability could be developed and fielded.

The presentation contains a general overview of the threats, including detection, interception, jamming, and spoofing. Protected waveform techniques, such as spread spectrum, forward error correction, cryptography, and filtering based on time and position are discussed.

An overview of the impacts and transition concepts for both the ground segment and space segment are reviewed, including re-use of existing assets, technology transition, and options for transition to a commercial model. Tradeoffs between re-use of existing spectrum vs allocation of new spectrum are discussed.

INTRODUCTION

The satellite command and control (C2) links used for telemetry, tracking, and commanding functions (TT&C) are a critical part of the satellite mission. There are a number of basic techniques used for TT&C. However, the technology used for satellite C2 has not kept pace with current telecommunications technology. It is reasonable to ask why it is that an average cell-phone has more anti-jam capability than a typical satellite C2 link, especially since the cell phone was not designed specifically as a protected communication device. The reality is that the cell phone technology, along with commercial broadcast standards (e.g. DVB) and internet protocols, have incorporated new technologies that have become available due to the ever-increasing capability of digital systems. These systems handle multiple users under very dynamic operating conditions, which leads to some degree of inherent resiliency.

TECHNOLOGY HISTORY

There are two basic approaches to building C2 links: direct ground access and relays. In the direct ground approach, a tracking station contacts the satellite directly. This mode of operation is the primary mode of operation for the Air Force Satellite Control Network (AFSCN). In the relay operation (e.g. TDRSS), an uplink signal is sent from a ground station to a relay satellite, which then transmits a forward link to either another relay, or to the destination satellite. The return link from the destination is received by the relay, which then transmits the signal to the ground station via the downlink.

In both of these systems, a mission control center is presumed to be the initiator of all commanding authority, and the consumer of all telemetry. In some cases, the control center may be segregated or distributed, but in either case, the end-to-end communication is presumed to be between the control center and the spacecraft.

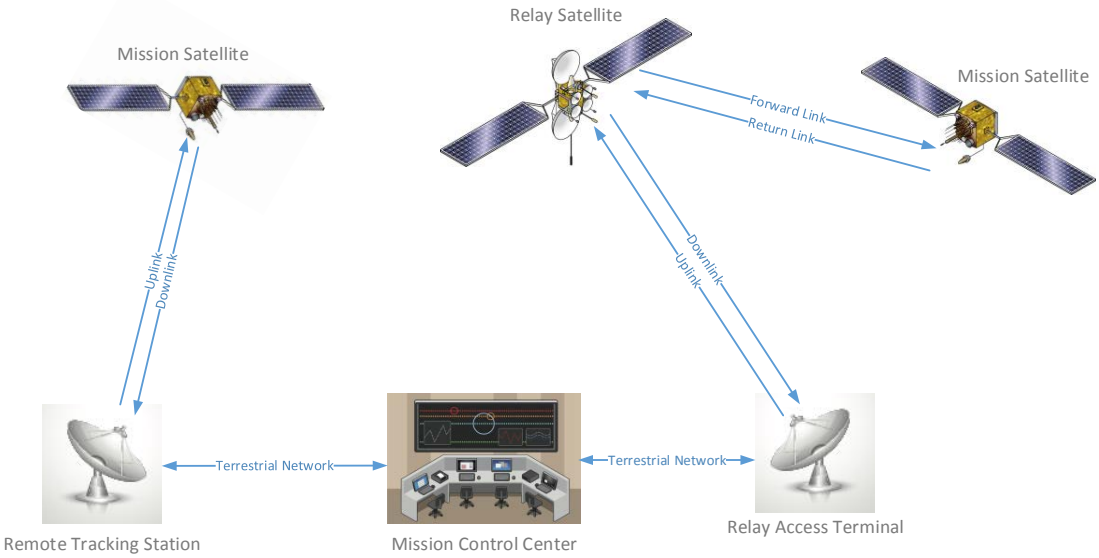


Figure N: Command and Control Architecture. *Satellites may be controlled by a direct link from a remote tracking station (left) or through a relay system (right). Relay may be one or more hops. Each node in the network may operate in either a "bent-pipe" mode, or in a "store-and-forward" mode.*

Each intermediate node in the network may operate in either a "bent-pipe" mode, or in a "store-and-forward" mode. A "bent pipe" node does not participate in the link security, but also does not require embedded cryptography. A "store-and-forward" node can provide increased security by eliminating cross-correlation between incoming and outgoing signals, at the cost of node complexity and cryptography.

For any future secure C2 system, support for both direct contact and relay modes of operation leads to improved resiliency by ensuring that there are multiple access methods to reach each mission satellite.

Current Satellite Command and Control Standards

This section provides a very brief overview of the major satellite command and control waveforms currently in use by the United States government defense and civil space programs.

Space-Ground Link System (SGLS)

The primary waveform for US military satellites is defined by the Space-Ground Link System (SGLS) standard. This waveform suite defines a commanding uplink using FSK modulation with an AM modulated embedded clock, a telemetry downlink using subcarriers, and a pseudorandom noise (PN) ranging signal. There are some provisions for direct modulated telemetry, but in practice, most users implement subcarrier telemetry.

Unified S-Band (USB)

The so-called Unified S-Band quasi-standard borrows some technology from the NASA STDN standard, such as subcarrier commanding techniques and standard subcarrier telemetry. USB sometimes uses tone ranging, but PN

ranging techniques are also used. Interestingly, USB has been used on other frequencies, so it is neither "unified" nor exclusively "S-band".

Tracking and Data Relay Satellite System (TDRSS)

The NASA-operated TDRSS defines several command and control links for use through its network of relay satellites. The multi-access system, which we will focus on here, is implemented on S-band, and provides both forward (commanding) and return links (telemetry). The multi-access waveforms are implemented using a PN-spreading technique, and the PN sequence is also used for ranging.

These standards represent a survey of basic satellite C2 capability. There are a multitude of other waveforms, including commercial standards and one-off designs in use by small-satellites that are very similar in concept. There are a few specific implementations of hardened C2 links, but they are the exception, not the norm. These hardened links are specific to the particular programs; there is no standard for hardened C2 links.

CURRENT TECHNOLOGY VS THE THREAT

The technologies described above are quite old. The SGLS standard was established in the 1960s, as was the STDN standard, on which USB is based. The TDRSS standard is somewhat newer, with TDRS-1 being first deployed in 1983. Meanwhile, our cell phone industry has had 5 major waveform technology refreshes (AMPS, GSM, CMA2000, UMTS, and LTE), and audio media with 7 tech refreshes (78s, LPs, 8-tracks, cassettes, CDs, digital downloads, and now streaming). Understandably, technology transitions on space programs can be difficult, given the long procurement times and even longer operational lifecycles. However, an examination of the capabilities (or lack thereof) of current C2 waveforms indicates a need for update.

A primary concern is the security of the C2 waveform. Although the data stream is usually encrypted, providing secrecy and some degree of authentication, the waveforms themselves do not in any way hide the traffic flow. It is readily apparent when commands are being transmitted, and the telemetry often has different modes depending on the operational state of the satellite (e.g. different data rates or modulation types) that are easily identified when examining the signal externals. The implication is that an external observer can infer things (for example, traffic patterns) about what is happening on our systems, with the possibility of either passive or active exploitation.

The waveforms are also not generally robust against receipt of bogus signals. Protections are built in, using cryptography, to prevent the most obvious and egregious intrusions (such as unauthorized commanding of a satellite), but this is not the only concern. Consider what would happen if an intruder attempted to transmit an uplink to a typical SGLS satellite. The frequencies and modulation to do so may not be public knowledge, but they are not hard to reverse-engineer from 50+ years of observation. Even if the intruder can't directly command the satellite, they could tie it up dealing with a fake signal, possibly draining batteries or keeping a legitimate commanding node from accessing the satellite.

The waveforms themselves are not very bandwidth or power efficient. The SGLS commanding waveform is over 10 dB worse than uncoded BPSK for power efficiency, meaning that a 1 kW uplink station is required, where a 100 W station should be sufficient. Likewise, the subcarrier downlinks leave considerable power in the residual carrier, which carries no data, and thus is effectively a waste of power on the satellite.

The TDRSS waveforms are somewhat better than SGLS or USB. They use spread-spectrum techniques with power efficient BPSK modulation. The spreading code affords a degree of code-division multiple access (CDMA) capability to the TDRSS system, and also provides an embedded ranging signal. However, the signal as currently defined does not provide any covertness to the data flow, and the code sequences are public and short, and are thus relatively easy to observe.

The waveforms are often implemented without forward-error-correction (though this is becoming less common with newer programs). Without spread spectrum or FEC (or, preferably, both), the signals are susceptible to

inadvertent interference. As the civilian use of the non-renewable resource called RF spectrum continues to increase, the likelihood of interference will increase accordingly.

PROTECTED WAVEFORM TECHNOLOGIES

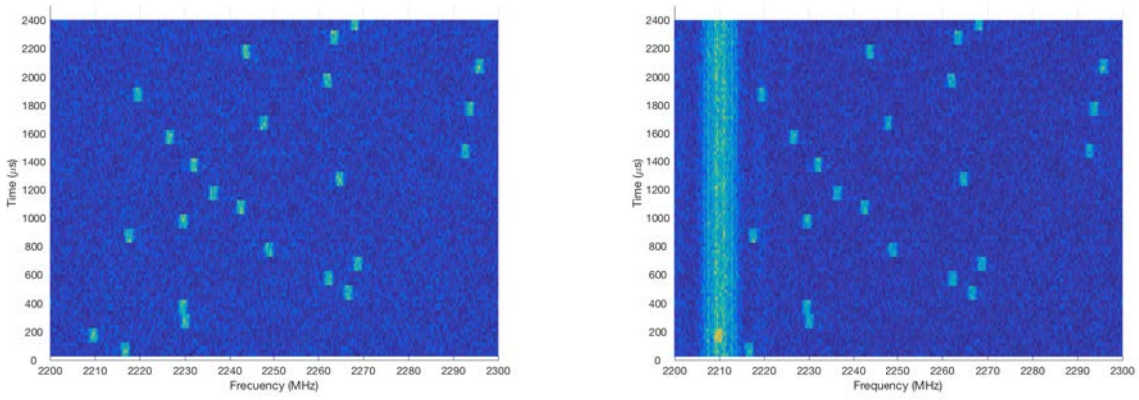
Spread Spectrum

This section describes spread-spectrum techniques that are used to provide a number of useful features:

- Multi-user access. Although technically a spread-spectrum waveform is not "bandwidth efficient", it does allow multiple users to share a portion of spectrum. The total capacity of the channel, when the number of user bits per Hz is considered, is often nearly the same when compared to a conventional FDMA channel.
- Anti-interference. Spread-spectrum is sometimes touted as "jam-proof", which is an unfortunate (and unrealizable) characterization of an actual feature, in that narrowband interference occurring on the spread waveform becomes wideband interference (at the same power level) when the received waveform is de-spread, thus reducing the net effect of the jamming or interference. Nothing is ever "jam-proof", but the advantage of a spread waveform can be easily characterized by the processing gain, which is roughly the ratio of the bandwidth of the spread waveform to the net bit rate.
- Covertness. If the waveform is spread with a secure spreading function, the signal will be hard to detect. Signal features, such as symbol rate or frame markers, can often be obscured by the spreading function. Furthermore, a non-repeating secure spreading function will be robust against problems such as replay attacks and cyclostationary detection.

Frequency Hopping

One simple approach to spreading is to randomly vary the carrier frequency over time. Traditionally, this is done in discrete time intervals, with the carrier frequency defined by a TRANSEC function over an interval of time, known as the hop period. FH waveforms achieve their robustness by avoiding interference for all but a short segment in time. Redundancy (including FEC) recovers any part of the signal that is lost. TRANSEC-driven FH prevents "following" jammers since the sequence of hop frequencies is only known to the transmitter and receiver. Frequency hopping is well-known as a robust means of achieving anti-jam performance with very high processing gain, but this requires fairly high bandwidths (on the order of 1 GHz).

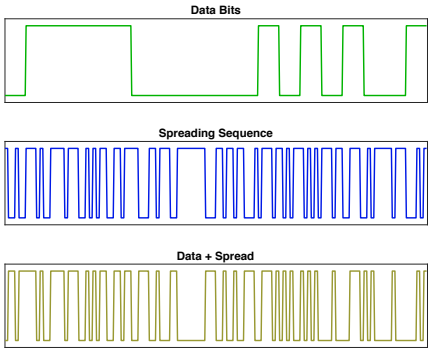


a. Frequency Hopping b. Frequency Hopping with Interference

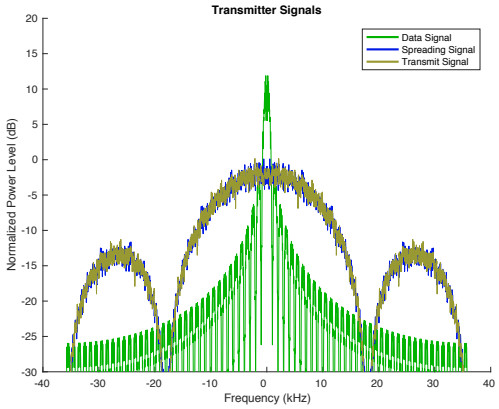
Figure N: Frequency Hopping Spread Spectrum. *(a) Frequency hopping changes the center frequency dynamically and pseudo-randomly. (b) Each individual hop is usually processed as an independent burst. Interference may collide with some hops, as shown in the lower-left corner, but the combination of hopping and forward error correction recover the data from the lost hops.*

Direct Sequence

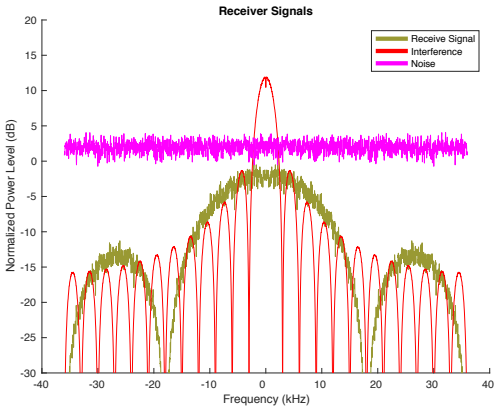
Direct sequence is another basic approach to achieving spread spectrum characteristics. The transmit symbols are mixed with a pseudo-random "chipping" sequence to achieve what appears to be a modulation of much higher rate. Without an increase in transmit power, the signal would not be received without significantly increasing the G/T of the receive station. However, the processing gain is achieved by a cross-correlation process in which the spreading sequence is removed by the receiver, yielding the received symbols as they were prior to spreading (plus noise-induced errors). The processing gain is the ratio of the chipping rate to the symbol rate, and allows for a corresponding reduction in the power spectral density of the transmit signal. This has the effect of disguising the signal so that it looks like noise to non-authorized receivers. The de-spreading process also has the benefit of performing spreading on any narrowband interference, yielding a significant gain in overall signal-to-interference (S/I) ratio.



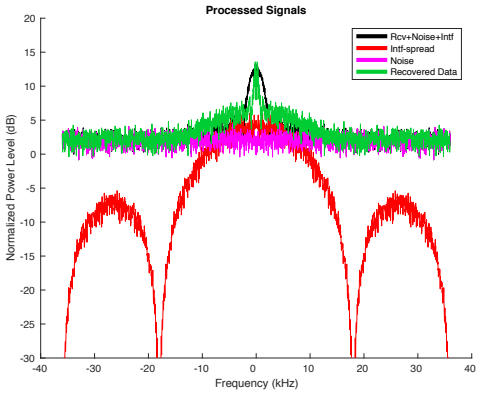
a. Time Domain



b. Transmitter Output



c. Receiver Input



d. Receiver Output

Figure N. Direct Sequence Spread Spectrum. *(a) Combining a data stream (green) with a pseudo-random spreading sequence (blue) yields another pseudo-random sequence with "hidden" data (gold). (b) In the frequency domain, the spread+data sequence is indistinguishable from the spread-only sequence, with no data rate information. (c) At the receiver, the spread signal is often received below the noise floor (pink), and may be subject to interference (red). (d) After receiver processing of the composite input (black), the interference is spread (red), and may impart a small increase in the noise floor (green). The data signal is now visible and recoverable in the center.*

Forward Error Correction

The use of forward error correction (FEC) is an excellent way of improving the overall performance of a data link with very little cost. Modern FEC using low-density parity codes (LDPC) can achieve very close to Shannon-bounded performance. FEC generally works best on channels with Gaussian noise, because the individual symbol errors are uncorrelated and uniformly distributed. By combining FEC with other techniques, such as interleaving and spreading, non-Gaussian channels (including interference) can often be transformed to have a Gaussian-like effect on the end result. Thus, the FEC performance is a key ingredient in achieving the highest level of robustness in a protected C2 link.

Cryptography

Encryption

Encryption is used to obscure the contents of the data being transferred, with the purpose of preventing disclosure of the data contents. Encryption should be implemented using a means that does not amplify normal system errors. For example, the cipher-block-chaining (CBC) mode will convert one bit error on the ciphertext side to (on average) 64 bit errors (assuming a 128-bit block cipher) on the plaintext side. Other modes, such at counter (CTR) or output feedback (OFB) do not suffer this problem, and are more suitable for communications on a Gaussian white noise channel. Advanced cipher modes, like Galois counter mode (GCM), provide high performance cryptography combined with authentication.

Authentication

Authentication is the verification of the authenticity of the data, that the data originated from the appropriate source, and has not been tampered with during transit. Authentication is subtly distinct from encryption, in that authentication is not concerned with preventing disclosure, but ensuring the integrity of the data at the receiving end. It is important that if data integrity is of utmost importance (for example, spacecraft commanding) that a true authentication mode be employed, and not simply rely on the failure of the decryption process to imply authentication.

Authentication protocols, in the form of message authentication codes and digital signatures, is well-established in standard networking protocols such as Transport Layer Security (TLS) and IPSec. Such techniques are easily adapted to satellite C2.

Ephemeral Keys

One general principle of any cryptographic system is that all of the security of the system should be contained in the key. The algorithms added to the system (such as COMSEC encryption, authentication, or TRANSEC) do not require protection if they are based on robust algorithms in which the key provides all of the security. Protection of the key becomes the fundamental task of the cryptographic functions. Furthermore, if the system contains other parameters or algorithms that require protection, the entire system becomes very expensive to develop and maintain, as these aspects must be protected by classification.

Ephemeral keys provide a unique key for each. Ephemeral keys are established through a key-establishment protocol, such as RSA or Diffie-Hellman, in such a way that the two endpoints (the MCC and mission satellite) both arrive at the same shared secret key, but without revealing anything to external observers that can be used to compromise the communications. Ephemeral keys enhance system security since the compromise of one session key has no impact on the security of any other session, since each session key is unique and independent.

Electronic Protection Measures for Satellite C2

This section describes some of the protection measures that could be applied to satellite command and control operations.

Spectrum Monitoring

Spectrum monitoring provides situational awareness on signals within the operating service volume that may or may not be of concern. Some of these signals might be known emitters that share the spectrum, and are not of concern, except under unusual circumstances (such as proximity of a satellite to the emitter). Others might be new or unexpected emissions that should be tracked to determine if, and when, they are a threat.

Captures of real-time spectrum data can be analyzed for signal characteristics, cataloged into a database, and referenced in the future. Such a catalog helps sort out the legitimate threats from the known, detectable but non-threatening emitters.

Cancellation

The processes and technologies used for spectrum monitoring can also be applied to the problem of cancellation. If an interfering signal is completely characterized in the time-domain, it is possible to subtract it out, removing it as a source of interference. Cancellation is not always possible, since it is conditional on the time-domain characterization. However, it can be a valuable addition to the electronic protection toolkit.

Filtering by Position or Orientation

The concept of filtering signals by incident angle is well-known in the form of directional antennas (either steerable dishes or phased arrays). These antenna structures have the dual advantage of reducing the overall power required from the transmitter, at the same time reducing the observed signal power from unwanted transmitters.

One difficulty with the use of directional antennas for satellite C2 is the need to communicate even when the satellite is not oriented properly. This will occur after orbit insertion, prior to completion of the satellite initialization process, and may also occur following certain on-orbit anomalies. In either case, it is necessary to establish 2-way communications with the satellite using one or more antennas with a large beam size (with corresponding low gain).

A phased array antenna employing either beam-steering, null-steering, or both, can provide significant electronic protection. Beam-steering permits increased gain in the direction of the trusted emitter, and null-steering allows for nulling of known interference sources. Furthermore, a phased array can be switched to a near-omnidirectional pattern when required for satellite maneuvers.

In addition to directional antennas, filtering based on satellite position relative to the service volume of known C2 nodes (either ground-based or space-based) provides increased protection. If the satellite is aware of its position in orbit (even if attitude control has not been established), it can reject incoming signals that appear when the satellite is outside of a “blue” service volume. Furthermore, if “red” service volumes associated with known threats are established, the satellite can be aware of, and monitor for, attempted intrusions when it is inside of these spaces.

Filtering by Time

Filtering by time is another technique for inhibiting the inadvertent entry of erroneous or harmful signals into the satellite C2 processor. This method is particularly helpful for eliminating the possibility of a replay attack, mounted by an adversary recording a valid C2 signal, and retransmitting the signal at a later point in time.

To implement time-based filtering, the satellite must have a reasonably accurate representation of current time, maintained by a stable on-board reference. Long-term stability is maintained by regular time transfer from another time reference, such as the C2 commanding node or GPS. The window of acceptance for determining if a received signal is valid is defined by the total amount of uncertainty that is built up in the on-board timekeeper. This uncertainty is a function of the stability of the on-board reference, the time since the last time-transfer event, and the uncertainty in the satellite position.

OPERATIONAL IMPLICATIONS

Spacecraft Initialization and Anomaly Recovery

One of the more difficult problems to consider for a robust TT&C link is to handle the problem of spacecraft initialization and anomaly recovery. Ideally, the initial acquisition of a "rebooted" satellite will only occur with an authorized control authority, and no external party will be able to negatively influence the system. A satellite in safe hold has minimal to no state information (including time), so an ideal robust protocol would be one in which no a-priori state information is required. Once the initial acquisition is completed, the synchronization of state between the satellite and the controlling authority may be established and maintained.

A satellite in safe hold is potentially quite vulnerable, as it, by definition, is not in an operational state, and may not have all of its built-in security mechanisms available. An analogy can be made to a computer which has been forced to reboot. If an intruder gains access to the BIOS of a system, the intruder owns that system, and can make whatever changes they desire, independent of the degree of hardening that has been applied to the operating system and application software on the system. The answer for the computer is to protect the BIOS, both physically (minimize the number of people who can access the hardware), and security (for example, BIOS passwords and locking down boot drives).

Satellites, by their nature, can't be limited by physical access. The satellite cannot be secured by a simple password (as with the BIOS example), as the password would be subject to a replay attack. However, there is some promise in the use of a bidirectional zero knowledge proof algorithm. In a zero knowledge proof, a challenger uses a secure protocol to verify whether or not a responder knows a particular secret. It is not necessary for the challenger to know the secret, only to know the proper response given a particular challenge. By implementing a bidirectional protocol, both the MCC and satellite are confident in the others identity, and can use that fact to derive a shared secret, which is then used to build an efficient, secure data link.

Initial acquisition of a satellite using zero knowledge proofs might work something like the protocol shown below.

1. The controller issues a challenge to the spacecraft. The challenge/response protocol has to be agreed-upon in advance between the controller and the spacecraft, and must be variable to prevent replay attacks.
2. The spacecraft responds to the challenge, proving to the controller that the spacecraft is holds the secret password, and is thus a trusted entity. However, the spacecraft can not yet verify the authenticity of the controller. Note that the controller is the only entity that knows the appropriate response to the challenge.
3. The spacecraft then issues its own challenge to the controller, with the result providing mutual authentication. At this point, the spacecraft and controller may synchronize their internal state (including time), and proceed with the spacecraft initialization process.

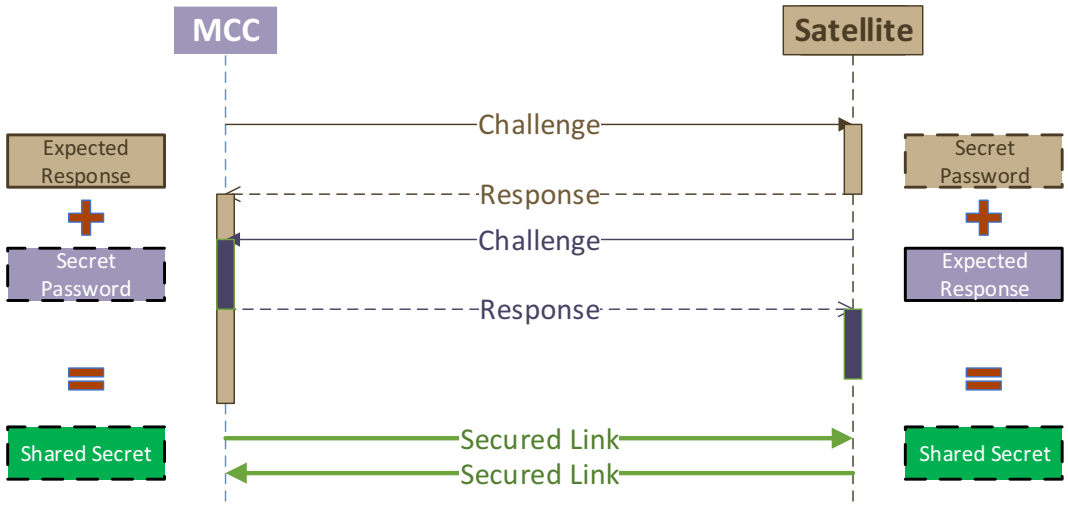


Figure N: Zero Knowledge Proof Login Protocol. *The challenge/response protocol is used by the challenger to verify that the responder knows a particular secret. The challenger does not have to know the secret, only the proper response for a given challenge. Following the bidirectional protocol, the two endpoints combine the secret knowledge and response data to form a shared secret, which is the basis for a permanent secured communication link.*

This protocol could also be inverted, with the initiator being the spacecraft, rather than the controller. There are operational considerations in either approach. If the spacecraft initiates the exchange, then it will have to start transmitting immediately upon restart. Depending on the condition of the power system (solar panels and batteries), this might be a concern. If the controller initiates the exchange, then it is possible for the system to be subject to potential denial-of-service type attacks. An intruder could issue continuous, bogus challenges to the spacecraft. With robust cryptography behind the protocol, these challenges would not succeed, but the spacecraft would be occupied answering these challenges. There are ways to limit the impact of such an attack (like limiting the number or frequency of incoming challenges), so it seems that this proposed approach is generally preferred. However, in certain circumstances, the use of satellite as the initiator may be warranted.

Spectrum Allocation

Most current command and control systems are located in L-band or S-band, which provides a great deal of robustness against weather effect (e.g. rain fade) and provides for relatively low impact due to dynamics. Both of these are conducive to use for satellite C2, given the criticality of being able to contact a satellite at any time, regardless of weather. Dynamics effects can be mitigated somewhat with knowledge of the system position, navigation, and timing (PNT), but this is not necessarily the case during satellite initialization.

Both of these bands are in high demand for terrestrial applications, and the relatively low frequencies (compared to C, X, Ku, Ka, etc) somewhat limit the overall bandwidth available for spread spectrum. Since the processing gain, and thus the interference rejection, of spread spectrum is dependent on the spreading bandwidth, it may be desirable to use a higher frequency band, at least for some types of C2 links.

Embedded Ranging and Time Transfer

Most existing TT&C waveforms include ranging capability, either in the form of a PN spread-like waveform (SGLS, TDRSS), or tone ranging (USB). With the advent of on-board GPS space receivers, many satellite systems are transitioning to the use of GPS for direct on-orbit PNT. As such, the use of ranging waveforms in the TT&C link may, on first glance, seem obsolete.

Reliance on GPS is becoming more ubiquitous, not just in satellite programs, but all through the US (both DoD and civilian users), and throughout the world. GPS is a very low user-cost, reliable, and high-performance system, so it is not surprising that it is so popular. However, this popularity comes at a cost. As we become more and more dependent on GPS, it starts to become a very vulnerable single-point of failure. Consequently, it is wise to consider alternate methods of achieving the same results. This doesn't mean that GPS use should be discontinued. On the contrary, the point is to have multiple systems with complementary redundancy to increase overall system resilience.

Fortunately, a spread-spectrum approach provides a nearly cost-free mechanism for implementing ranging and time transfer to augment any tracking data derived from other sources, such as GPS. The performance of a ranging signal is directly related to the bandwidth of the signal, thus by spreading the signal, the performance is improved, without having to increase the transmit power (GPS already exploits this feature). The only additional steps necessary to turn the signal into a ranging signal is to synchronize the spreading functions to a common time reference, and define a protocol for executing the time transfer.

Military+Commercial+Civil Multi-Use Technology

Two basic models for building secure communication systems have been employed in the past. The first approach is a purpose-built system, using a unique system specification, and may include multiple sensitive or classified elements. This approach is typically very expensive, as the supplier participation is minimal, the non-recurring engineering (NRE) is large, and the marketspace for the end-product is small. The second approach, exemplified by systems such as PKI and TLS, is to define open standards using robust, but publicly-available, security protocols (most notably from NIST). This approach allows for a broader range of suppliers, and a broader market for those suppliers. Even if the NRE is not significantly different, the overall cost to any individual user is much less.

This latter approach is recommended for a protected C2 waveform design.

Another aspect of an open approach (which can also be seen in the TLS design example) is the ability to upgrade cryptography without major changes to the overall system structure. TLS uses a concept called ciphersuites, which together define standards for encryption, authentication, and key establishment. It is flexible enough that large parts of the cryptography can be switched out (including wholesale replacement of the underlying ciphers), but the basic design is maintained.

SUMMARY

A highly robust communication mechanism for satellite command and control is required to prevent attackers from adversely affecting our ability to control our assets in space. This paper presents a number of concepts and principles for building in protections into satellite command and control links that combat both passive attacks (such as external observations and traffic analysis) and active means (including jamming or spoofing).

These techniques include the use of TRANSEC to build protected waveforms, in addition to traditional COMSEC encryption; spreading waveforms using frequency hopping, direct sequence, or some combination to mitigate interference and provide a degree of covertness; and electronic protect measures such as spectrum monitoring, cancellation, and filtering based on position, orientation, and time. Many of these approaches are not new to communications in general, but are not normally applied to satellite C2, either due to historical reasons, or over concerns with compromising standard operating procedures. However, this paper has demonstrated that these advanced concepts can improve system resiliency while still preserving key operations such as post-launch initialization and anomaly operations (including prior to establishment of attitude control), and integrated ranging and time transfer capability.