EHF Satellite Communications on the Move: Baseband Considerations

J.B. Schodorf

14 February 2000

Lincoln Laboratory
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LEXINGTON, MASSACHUSETTS

Prepared for the Department of the Army
under Air Force Contract F19628-93-C-0002.

Approved for public release; distribution is unlimited.
This report is based on studies performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology. This work was sponsored by the Department of the Army, CECOM, under Air Force Contract F19628-95-C-0002.

The ESC Public Affairs Office has reviewed this report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

Gary Tunugian
Administrative Contracting Officer
Plans and Programs Directorate
Contracted Support Management

Non-Lincoln Recipients
PLEASE DO NOT RETURN

Permission is given to destroy this document when it is no longer needed.
ABSTRACT

This report investigates the challenges associated with the development of an EHF land mobile satellite communications system. Specifically, baseband signal processing and networking protocol issues are addressed. An attempt has been made to assess the problem at two levels simultaneously. On one hand is the global long range view, while on the other hand is a near term demo-oriented focus. Ignoring the very important problems of antenna pointing and tracking, the primary obstacle to be overcome by the system is signal blockage. At EHF, objects in the propagation path are virtually opaque and cast a dark “shadow” over the communications terminal, resulting in signal attenuations on the order of 20-30 dB or more. The duration of these shadow regions or blockage intervals will depend on a number of factors, including object size and vehicle speed. The proper long term solution to dealing with the blockage problem requires enhancements at most of the layers of the protocol stack. For example, forward error correction coding should be employed at the physical layer to mitigate the effects of relatively short blockage intervals (i.e., milliseconds to seconds). At the link layer automatic repeat request schemes are the best solution to ensure reliability over longer shadow regions (i.e., seconds to 10s of seconds). At the network layer dynamic routing protocols can be used to provide connectivity to blocked terminals via alternate terrestrial communications paths that include at least one unobstructed terminal. Finally, TCP splitting and protocol conversion offer a great deal of promise for achieving reliable end-to-end transport layer services without sacrificing TCP compatibility. Despite what may at first appear to be a relatively complicated long term solution to providing on-the-move services at EHF, a near term approach designed to demonstrate a basic capability is straightforward. For example, a number of the ideas and concepts just mentioned can be incorporated into a stand-alone software application designed to run over UDP. By addressing the aforementioned challenges at the application layer, many of the technical issues relating to modifying the other layers can be avoided at the initial expense of providing support for other commercial and military applications. In this report, each of the above areas are examined in more detail and areas requiring further research are identified.
ACKNOWLEDGMENT

There are a number of people that I would like to acknowledge for their support and contributions to this work. First, I would like to thank Ron Bauer and David Snider for their useful suggestions and encouragement.

In addition, I would like to thank Sam MacMullan and Steve Bernstein for giving graciously of their time in matters relating to forward error correction coding. I would also like to thank Jerry O'Leary for his input concerning the ARQ and voice processing parts of this report. Rich Greene also provided valuable insight into the area of tactical speech, for which I am extremely grateful. I am also indebted to Chris Karpinsky for his helpful suggestions and proofreading of the networking protocols chapter. Special thanks also go to Mark Gouker and Ron Bauer for taking the time to read this report in its and entirety and for making suggestions that greatly improved the content and clarity of the material.

I also want to acknowledge the following people for taking the time to answer questions, provide insight, and general guidance in writing this report: Ed Bucher, Ben Eaves, Ken Hetling, Steve Kolek, Eytan Modiano, Scott Stadler, Leslie Weiner and Peter Wu.

Finally, I would like to thank Mike Gleason, Thom Nelson, and Pete Stevens of CECOM and Gary Martin of PM Milsatcom for their continuing sponsorship of this work.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENT</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF ILLUSTRATIONS</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xi</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2. PROPAGATION MODELING</td>
<td>5</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>5</td>
</tr>
<tr>
<td>2.2 Path Loss Models</td>
<td>6</td>
</tr>
<tr>
<td>2.2.1 Foliage Attenuation</td>
<td>6</td>
</tr>
<tr>
<td>2.2.2 Weather/Atmospheric Effects</td>
<td>8</td>
</tr>
<tr>
<td>2.3 Signal Fading</td>
<td>8</td>
</tr>
<tr>
<td>2.3.1 Multipath Fading</td>
<td>11</td>
</tr>
<tr>
<td>2.3.2 Shadowing</td>
<td>12</td>
</tr>
<tr>
<td>2.3.3 Fading Due to Antenna Mispointing</td>
<td>14</td>
</tr>
<tr>
<td>2.4 Laboratory Simulation of the Propagation Channel</td>
<td>15</td>
</tr>
<tr>
<td>2.5 Conclusions</td>
<td>15</td>
</tr>
<tr>
<td>3. ERROR CONTROL AND TIMING</td>
<td>17</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>17</td>
</tr>
<tr>
<td>3.2 Channel Definitions</td>
<td>17</td>
</tr>
<tr>
<td>3.3 Strategies for Voice Communications</td>
<td>18</td>
</tr>
<tr>
<td>3.3.1 FEC Coding</td>
<td>18</td>
</tr>
<tr>
<td>3.3.2 ARQ Approaches</td>
<td>23</td>
</tr>
<tr>
<td>3.4 Strategies for Data Communications</td>
<td>28</td>
</tr>
<tr>
<td>3.5 MILSTAR Specifics</td>
<td>32</td>
</tr>
<tr>
<td>3.6 Timing Considerations</td>
<td>33</td>
</tr>
<tr>
<td>3.7 Conclusions and Recommendations</td>
<td>34</td>
</tr>
<tr>
<td>4. NETWORKING PROTOCOL CONSIDERATIONS</td>
<td>37</td>
</tr>
<tr>
<td>4.1 Introduction</td>
<td>37</td>
</tr>
<tr>
<td>4.2 TCP/IP and the Satellite Channel</td>
<td>37</td>
</tr>
<tr>
<td>4.2.1 TCP Modification/Replacement</td>
<td>39</td>
</tr>
<tr>
<td>4.2.2 Compatibility-Preserving Approaches</td>
<td>39</td>
</tr>
<tr>
<td>4.3 Routing Protocol Technologies</td>
<td>41</td>
</tr>
<tr>
<td>4.4 Conclusions and Recommendations</td>
<td>43</td>
</tr>
<tr>
<td>5. SUMMARY AND CONCLUSIONS</td>
<td>45</td>
</tr>
<tr>
<td>APPENDIX A – CONCEPT OF OPERATIONS</td>
<td>47</td>
</tr>
<tr>
<td>Figure No.</td>
<td>Illustration</td>
</tr>
<tr>
<td>-----------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1</td>
<td>Path loss prediction models.</td>
</tr>
<tr>
<td>2</td>
<td>Simulated Ricean fading envelopes for various values of the Rice factor, K.</td>
</tr>
<tr>
<td>3</td>
<td>Two-state Markov model that describes the switching process in a LMSC channel.</td>
</tr>
<tr>
<td>4</td>
<td>Losses due to antenna mispointing for various aperture sizes.</td>
</tr>
<tr>
<td>5</td>
<td>The two-state model proposed by Lutz et al..</td>
</tr>
<tr>
<td>6</td>
<td>Various signal envelopes produced by the Lutz channel simulator.</td>
</tr>
<tr>
<td>7</td>
<td>Histogram of transmission times in a tactical scenario.</td>
</tr>
<tr>
<td>8</td>
<td>Output BER with erasure decoding.</td>
</tr>
<tr>
<td>9</td>
<td>Output BER with error decoding.</td>
</tr>
<tr>
<td>10</td>
<td>Latency results for the three channels.</td>
</tr>
<tr>
<td>11</td>
<td>The effects of greater capacity on channel delay performance.</td>
</tr>
<tr>
<td>12</td>
<td>Performance comparison for FEC coding and ARQ.</td>
</tr>
<tr>
<td>13</td>
<td>Data throughput performance comparison for three types of error control.</td>
</tr>
<tr>
<td>14</td>
<td>Throughput vs. link margin for several different RS codes.</td>
</tr>
<tr>
<td>15</td>
<td>BER performance of an $R = 1/2$, $L = 7$ convolutional code.</td>
</tr>
<tr>
<td>16</td>
<td>A cluster with multiple satellite terminals.</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table No.</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Parameter values used to generate the signal envelopes in Figure 6.</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>Parameters that define the open, rural, and urban channels.</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td>Interleaving delays for the various channels.</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>Average delays, in sec., for different channel capacities.</td>
<td>26</td>
</tr>
<tr>
<td>5</td>
<td>Throughput results for ARQ.</td>
<td>26</td>
</tr>
</tbody>
</table>
1. INTRODUCTION

True on-the-move (OTM) satellite communications capability has been a strong desire of the U.S. Army for many years. Certainly, the ability to communicate OTM via satellite at UHF has been around for some time now. However, UHF satellite resources (i.e., bandwidth) are limited, and the Army is generally reluctant to depend on these systems for critical communications services. Instead, the Army has identified the MILSTAR system, which operates at EHF (i.e., 20.2-21.2 GHz downlink and 43.5-45.5 GHz uplink), as their primary satellite resource for future tactical operations. For example, MILSTAR is now tightly coupled with the Army’s mobile subscriber equipment (MSE) and tactical packet network (TPN) where it is used to provide range extension capabilities. From the Army’s perspective MILSTAR has a number of advantages, the most important of which is its robustness against hostile jamming. For these reasons there is a great deal of interest in the development of MILSTAR satellite terminals that have the ability to operate OTM. However, there are a number of challenges associated with communicating OTM at EHF. For example, propagation losses are significantly worse at EHF compared to other frequency bands. Consequently, EHF terminals typically use directive antennas (e.g., parabolic dishes) for increased gain. Of course, the use of directive antennas implies that they must be pointed accurately for maximum benefit, a significant task on a mobile platform. In addition, the shorter wavelengths at EHF imply that these signals are easily scattered by objects in the propagation path such as foliage or buildings. The need for error correction coding and other baseband signal processing schemes to help mitigate these effects are therefore equally important.

The purpose of this report is to investigate the relevant baseband signal processing issues associated with the development of a military land mobile satellite communications (LMSC) system that operates at EHF. Issues relating to antenna pointing and tracking are beyond the scope of this report. However, they have been addressed elsewhere [1]. In addition, it should be noted that the term LMSC is used to distinguish the proposed terminal from one that might be used on an airborne platform such as a plane or helicopter. While the development of an airborne terminal is not without its own challenges, the associated baseband signal processing issues are quite different from those of a ground-based mobile terminal. For example, airborne terminals will rarely experience the same blockage effects as ground-based terminals because airborne platforms are typically elevated above any obstructions. Of course, airborne terminals may have their own unique blockage problems, such as rotor blockage effects on a helicopter terminal. Nonetheless, algorithms for dealing with blockage experienced by a ground-based versus airborne terminal will differ significantly. Other work related to EHF satellite communications OTM includes the mobile paging experiments conducted by MIT Lincoln Laboratory. This work was important because it demonstrated a basic capability and focused the attention of the military community on the possibilities for EHF communications OTM. However, the current investigation is more ambitious in the sense that the proposed LMSC terminal will ultimately be capable of supporting communications more sophisticated than simple paging signals.

The fundamental conclusion of this report is that the proper long range solution to fielding an EHF LMSC system will require enhancements at most layers of the protocol stack. Moreover, algorithms and protocols implemented at each of the system layers should be capable of adapting to the dynamic channel conditions. However, an alternate solution that is more feasible in the near term would be to develop a stand-alone application for voice and data communications over the EHF LMSC channel. By incorporating blockage mitigation techniques at the application layer,
many of the tedious aspects associated with modifying or creating custom lower-layer protocols can be avoided. The disadvantage with stand-alone software applications is of course the fact that they are generally incompatible with other applications. Nonetheless, this approach represents a straightforward solution to the problem of demonstrating voice and data capability with an EHF LMSC system. In addition, it should be noted that many of the algorithms refined over the course of demonstrating the voice/data software application could eventually be incorporated into enhancements at other layers in the protocol stack as part of a more comprehensive, long term solution. In other words, the near term and long term approaches discussed above are not mutually exclusive. Rather, they are complementary with the near term software application representing a logical milestone on the path to a more sophisticated long term solution to the problem.

This report is somewhat long due to the fact that an attempt has been made to be as thorough as possible. However, the contents have been structured so that readers less interested in the technical details can skim the report and still come away with a general understanding of the significant issues and recommended approaches for addressing them. To get a quick overview of the report, readers should pay particular attention to the introduction and conclusions of each chapter. Moreover, figure captions were designed to be relatively self-contained and should provide additional insight.

The remainder of this report is organized as follows. In Chapter 2 the EHF LMSC propagation channel is examined briefly. The dominant effect of this channel is noted to be signal shadowing caused by objects in the propagation path such as buildings or foliage. In fact, the effects of shadowing are so severe (e.g., attenuations on the order of 30 dB below the mean signal level) that a system transmitter and receiver are effectively blocked from communicating with one another for the duration of the shadow interval. A useful software simulation technique that accurately reflects not only shadowing but other EHF LMSC channel phenomena is also presented in Chapter 2 and used to simulate typical channel received envelopes. This simulation capability is useful for the development and testing of algorithms designed to improve communications over the LMSC channel.

In Chapter 3 error control algorithms for the EHF LMSC channel are examined. In addition, problems associated with timing acquisition and tracking are also considered. With respect to error control, voice and data are treated separately due to their differing characteristics. However, for both classes of information, some form of ARQ was found to be the most appropriate means of addressing the relatively long channel outages due to shadowing. Pure FEC coding was found to introduce unacceptably large delays due to the interleaving required with this approach. Some FEC coding was found to be necessary for the purpose of addressing random bit errors due to noise, etc., in the channel good state when data packets were being transmitted. With regards to timing, it is postulated that acquisition in a shadow region is virtually impossible. However, time tracking can be maintained through typical outage periods by simply suspending for the duration of the channel outage the early/late correlations typically performed by the tracking circuitry.

In Chapter 4 network and transport layer protocol considerations are addressed. The most significant issues in this area relate to the performance of TCP over a satellite link. In general, TCP is known to perform poorly over geostationary satellite channels because the protocol suite was not optimized for the high delay-bandwidth product and high error rate conditions typical of the satellite environment. The two competing schools of thought on addressing this problem are summarized as follows. On one hand, backward compatibility with TCP is highly desirable given the significant commercial investment in this standard. There are several methods for improving TCP
performance while simultaneously achieving backwards compatibility, including link enhancement via specialized link layer protocols and TCP splitting via gateways. On the other hand, there are several proposals in the literature that call for modification or even replacement of TCP in order to achieve maximum performance gains over a satellite link. Of course, with these schemes alterations to end-user applications utilizing only TCP is necessary, implying significant cost.
2. PROPAGATION MODELING

Fundamental to the design of any communications system is an accurate understanding of the medium over which the communications signals will be traveling. For this reason, propagation modeling plays a very important role in the development of wireless communication systems. While propagation modeling continues to be an area of active research, especially at higher frequencies, the field is relatively mature with respect to most communications bands [2]. This chapter will highlight major results from propagation modeling research with an emphasis on the LMSC channel and its extension to the EHF band.

2.1 Introduction

The first propagation phenomenon considered here is path loss. Specifically, signal attenuation due to foliage and rain is examined and mathematical models that describe these losses as a function of frequency and path length are presented. The second phenomenon treated in this chapter is signal variability, often referred to as signal fading. Signal fading is most prevalent in wireless systems where either the transmitter, receiver, or both are moving. Statistical models that describe this fading process in a LMSC system are also examined. While path loss and signal fading occur in virtually all mobile wireless systems, EHF systems, due to their shorter operating wavelength, will experience these effects with different levels of severity compared to their lower frequency counterparts. Specifically, signal shadowing at EHF is so severe that the channel can be viewed as being in one of two states: a “good” state where the signal is not blocked and communications is possible, and a “bad” state where the signal is blocked and, in the absence of special algorithms, communications is not possible. Statistical descriptions of this behavior based on real data will be presented later in this chapter. Another phenomenon that will be investigated is fluctuation in received signal strength caused by antenna tracking errors. These fluctuations can be interpreted as fading events. However, characterizing the fades is difficult because the statistics vary depending on the quality of the antenna controller and the terrain in which the system is being operated. Finally, after having laid the groundwork by examining propagation effects, a software simulator for reproducing these phenomena is presented in Section 2.4. The simulator was originally proposed in [3] for use in modeling the behavior of a LMSC system at L-band. However, the model is general enough such that statistics that more accurately reflect the situation at EHF can be easily substituted. The development of a channel simulator is an important step towards the design and evaluation of algorithms proposed to compensate for propagation effects.

The most comprehensive EHF propagation experiments conducted to date involved NASA’s Advanced Communications Technology Satellite (ACTS). Results from these experiments will be referenced frequently in the following sections. Consequently, some background information on the ACTS propagation studies is in order. ACTS was launched by NASA into geosynchronous orbit in the fall of 1993 for the purpose of development of high-risk advanced satellite communications technology at K/Ka band. The system operates at 20 GHz on the downlink and 30 GHz on the uplink. As part of the ACTS program, extensive propagation experiments were conducted by a number of parties over a wide variety of locations in the United States using both fixed and mobile terminals [4, 5].
2.2 Path Loss Models

In any wireless communications system the received signal strength will vary significantly depending on the operating environment. Signal scattering and absorption by objects in the propagation path (e.g., trees and other vegetation) and signal attenuation from the atmosphere and precipitation are just some of the factors that contribute to path loss. Moreover, the effects of each of these phenomenon on received signal strength vary widely with wavelength. In fact, it is the size of objects in the propagation path relative to the signal wavelength that determines, in large part, their effect on the propagating signal. For example, whereas signal attenuation by foliage is not a significant problem at UHF frequencies, it is one of the dominating path loss effects at EHF.

A very general and widely accepted path loss model is given by [6]:

\[ L = \frac{(C \lambda)^2}{d^n} \]  \hspace{1cm} (1)

where \( L \) is the attenuation associated with the transmitter-to-receiver path, \( C \) is a constant that depends on the average path loss at a reference distance \( d_0 \) from the transmitter, \( \lambda \) is the signal wavelength, \( d \) is the distance between transmitter and receiver, and \( n \) is the path loss exponent. For free space, the constant is usually taken to be \( C = 1/4\pi \) and \( n = 2 \). However, for cellular communication systems, \( C \) is different and the path loss exponent is usually taken to be \( n = 4 \) [7] to account for the presence of significant ground reflections. On the other hand, in environments with dense trees the path loss exponent may exceed 8 [6].

2.2.1 Foliage Attenuation

The model in (1) is convenient because it is simple and quite general. However, the generality of (1) is also a drawback since it does not always accurately reflect losses for specific situations. For example, even with a relatively large path loss exponent, the model does not accurately reflect the significant losses experienced by EHF signals through short to moderate amounts of foliage (i.e., \( d < 50 \) meters). Consequently, other models must be examined.

A different approach to characterizing path loss due to foliage is to use models fitted to experimental data. One such model, discussed in [8], reflects a constant attenuation in dB per meter of path length. The empirical formula was fit to data that extended from 9.5 to 95 GHz and is given by:

\[ L_f = 1.102 + 1.48 \log_{10}(F) \]  \hspace{1cm} (2)

where \( L_f \) is the path loss in dB/meter and \( F \) is the operating frequency in GHz. The above expression yields losses on the order of 3 dB/meter at 21 GHz and 3.5 dB/meter at 44 GHz. One problem with models designed to reflect constant losses per meter of foliage, such as (2), is that they fail to account for the recombining of scattered energy over larger distances. In other words, models such as (2) are accurate over relatively short distances only. This fact follows from the observation that only part of the signal energy is absorbed by foliage while the rest is scattered. This scattered energy can recombine at further distances so that signal losses are less severe than predicted by a constant loss model. Another model that accounts for this recombining effect, the so-called modified exponential decay (MED) model, was developed by Weissberger after comparing several exponential decay models like the one in (2) to data sets available at frequencies from 230

6
MHz to 95 GHz [2]:

\[ L_f(dB) = \begin{cases} 
1.33F^{0.284}d^{0.588} & 14 \leq d \leq 400 \\
0.45F^{0.284}d & 0 \leq d \leq 14 
\end{cases} \] (3)

In (3) \( L_f \) is the path loss in dB, \( F \) is the operating frequency in GHz, and \( d \) is the depth, in meters, of dense, dry foliage in the transmission path. The difference in path loss for trees with and without leaves is approximately 3 to 5 dB [2]. Figure 1 illustrates path loss due to foliage as predicted by the models presented above at 20 GHz. For comparison purposes, the model given by (1) is also plotted with \( C = 1/4\pi \) and a path loss exponent of \( n = 8 \).

Note from the figure that (1) is entirely inappropriate for modeling path loss through foliage at EHF since it suggests there is little to no loss for \( d < 30 \) meters. Even though the model can be improved slightly by adjusting \( C \), it will never perform as well as those based on actual measurements, especially for smaller values of \( d \). Another observation to be made from Figure 1 is the relatively steep slope of the model in (2). Again, it is important to note that this model is less accurate for larger values of \( d \). Finally, the MED model seems best because it more accurately describes the situation at EHF: a relatively fast accumulation of losses over short distances with a gradual decrease in loss per meter over larger distances. Moreover, as the figure indicates, it takes only 5 to 10 meters of foliage before the losses exceed 10 to 15 dB. Since losses in excesses of 10 dB or so are not likely to be compensated for with raw link margin, bits transmitted during periods where the receive path is obstructed by foliage are practically erased. This result is in sharp contrast to the cellular bands (i.e., 800-900 MHz) where losses due to foliage are less significant and can generally be offset with a small amount of link margin. It should also be noted that for a satellite communications system, the fraction of the propagation path blocked by foliage will depend on elevation angle. Clearly, it is more likely that paths at lower elevation angles will experience more attenuation than those at higher elevation angles.
2.2.2 Weather/Atmospheric Effects

As opposed to lower frequencies, weather and atmospheric effects at EHF are significant [9]. Signal scattering and absorption by rain droplets are the dominating factors at 20 and 44 GHz. Attenuation due to rain is a function of the amount of rainfall, drop size, temperature, operating frequency, and path length through the rain. The results from two measurement campaigns in India were recently reported in [10, 11]. In both cases, a strong correlation between signal attenuation at millimeter wavelengths and rain rate was observed. A regression to a power law (i.e., \( L_r = aR^b \)) yielded the following best fit equations for the data in [10]:

\[
\begin{align*}
  L_r &= 0.087R^{1.05} & F &= 22 \text{ GHz} \\
  L_r &= 0.184R^{0.99} & F &= 31 \text{ GHz}
\end{align*}
\]

where \( L_r \) represents the attenuation in dB per kilometer (km) and \( R \) is the rain rate in millimeters (mm) per hour. The parameters in (4) agree closely with those reported in [11] where \( a = 0.07 \) and \( b = 1.10 \) for 20 GHz and \( a = 0.36 \) and \( b = 0.966 \) for 40 GHz. Substituting a “typical” rain rate of 20 mm/hr yields losses on the order of 2 dB/km and 6 dB/km at 20 GHz and 40 GHz, respectively. Obviously, rain rate statistics will vary from region to region, thus affecting average losses. Given these numbers, it seems likely that an adequate amount of link margin can be designed into an EHF LMSC system to compensate for losses due to rain.

2.3 Signal Fading

The term signal fading is used to describe random fluctuations in the received signal envelope induced usually by motion in the receiver, transmitter, or both. There are two general classes of signal fading. First there is multipath fading, which as the name suggests is caused by the constructive and destructive addition of multipath signals at the receiver. The second type of fading is shadowing, a phenomenon that reflects large scale variations in the received signal. Clearly, the rapidity of fluctuations due to multipath and shadowing will depend on vehicle speed and the local environment. At EHF the effects of multipath are less pronounced than at lower frequencies. This observation is due to the fact that the surfaces of objects in the propagation path that reflect signal energy and give rise to multipath propagation appear “rough” rather than “smooth” to millimeter waves, resulting in scattered energy instead of specular reflections. The energy per scattered ray is significantly smaller than in the specular case. Moreover, highly directive receive antennas are typically employed in EHF satellite communications systems resulting in fewer received multipath components, thus reducing the effects of multipath at the receiver. In the following subsections both multipath fading and shadowing are discussed in greater detail and within the context of an EHF LMSC system. In addition, a third class of signal fading behavior, caused by antenna mispointing errors, will also be examined.

2.3.1 Multipath Fading

As mentioned earlier, the term multipath fading is used to describe the phenomenon whereby multiple, reflected versions of a transmitted signal (i.e., multipath components) combine at a receiver in either a constructive or destructive fashion depending on their relative amplitudes and phases. When either the transmitter or receiver is in motion, the dynamic, random combining of multipath components leads to received signals that can vary by several 10s of dBs with relatively
small changes in spatial location. Statistically, multipath fading may be treated as a random process. In the event that no single dominant propagation path exists, the fluctuations in received signal power, $S$, are described by the central chi-square distribution [12]:

$$p(S|S_0) = \frac{1}{S_0} \exp \left( -\frac{S}{S_0} \right)$$

(5)

where $S_0$ is the mean received signal power, due entirely to multipath. This class of channel is often referred to as a Rayleigh channel because the received signal envelope follows the Rayleigh distribution. This statistical model holds well in practice for microwave cellular systems where typically no line of sight (LOS) path between transmitter and receiver exists [7]. For situations where a LOS path or even a dominant multipath component does exist, such as the satellite communications channel, the received signal power, $S$, is described by the non-central chi-square distribution [12]:

$$p(S|A, \sigma_d^2) = \frac{1}{\sigma_d^2} \exp \left\{ -\frac{A^2 + 2S}{2\sigma_d^2} \right\} I_0 \left( \sqrt{2S} \frac{A}{\sigma_d^2} \right)$$

(6)

where $A$ is the amplitude of the LOS component, $\sigma_d^2$ is the diffuse signal power, and $I_0$ is the modified zeroth order Bessel function of the first kind. This class of channel is often referred to as a Ricean channel because the received signal envelope follows a Ricean distribution. Ricean channels are frequently parameterized by the so-called Rice factor:

$$K = \frac{A^2}{2\sigma_d^2}$$

(7)

The Rice factor is simply the ratio of the power in the direct and multipath components. When no LOS component exists (i.e., $A = 0$), $K = 0$ and (6) reduces to (5) with the mean received signal power given by $S_0 = \sigma_d^2$. When $K = \infty$, the channel does not exhibit fading. Figure 2 compares simulated fading envelopes for several different values of the Rice factor, $K$.

At EHF the presence of multipath fading is less significant compared to lower frequencies. This observation follows from the fact that at millimeter wavelengths, most surfaces appear "rough" and result in scattered energy rather than specular reflections. A practical criterion for identifying a rough surface is given by [2]:

$$h \geq \frac{\lambda}{8 \sin \phi}$$

(8)

where $h$ is the height of undulations in the reflecting surface, $\lambda$ is the wavelength, and $\phi$ is the angle of incidence. From (8), it is clear that at frequencies in the GHz range, even minor surface irregularities (i.e., on the order of millimeters) will appear large and result in diffuse reflections. Since the diffuse reflections scatter the energy in all directions, and since highly directive (i.e., narrow beamwidth) antennas are typically used at EHF, only a small portion will be incident at the receiver. The reduced significance of multipath energy at EHF is reflected in larger values for the Rice factor as compared to lower frequencies. Whereas Rice factors reported for the L-band system studied in [3] average approximately 10 dB, the Rice factors reported from the ACTS data [13, 14, 15] average more than 20 dB.

Another useful way to characterize the multipath fading process is by its power spectral density (PSD). An accurate model for the PSD of a multipath fading channel is also useful in developing a
Figure 2. Simulated Ricean fading envelopes for various values of the Rice factor, $K$. (a) $K = 0$ (i.e., Rayleigh fading), (b) $K = 10$ dB, (c) $K = 20$ dB, and (d) $K = 50$ dB (i.e., virtually no fading). Typical values for $K$ at EHF are on the order of 20 to 25 dB, indicating the presence of only a modest amount of multipath energy at the receiver [13].
laboratory simulator for the channel. Accurate theoretical models for the PSD of multipath fading channels in cellular systems have been derived and verified in practice [7]. No such theoretical models exist for the EHF LMSC channel. However, in [3] the PSD for a LMSC channel was estimated based on real data collected in a variety of experiments at L-band. The PSD was observed to resemble an exponential decay on a log scale that is easily approximated with a lowpass filter.

2.3.2 Shadowing

Signal shadowing is caused by the complete or partial blockage of a transmitted signal (i.e., \(A = 0\)) by objects in the propagation path such as nearby buildings, hills, and mountains. As discussed in Section 2.2.1, path losses due to foliage are significant enough to classify trees as objects that give rise to shadowing in the EHF band. Although difficult to model mathematically, variations in the received signal power, \(S_0\), caused by shadowing have been observed to follow a log-normal distribution (i.e., a distribution whose values, when plotted on a log scale, appear Gaussian):

\[
p(S_0) = \frac{10}{\sqrt{2\pi} \sigma \ln 10} S_0 \exp \left[ -\frac{(10 \log S_0 - \mu)^2}{2\sigma^2} \right]
\]

(9)

with a mean, \(\mu\), and standard deviation, \(\sigma\), that depend on the carrier frequency and environment [2]. The effects of shadowing in the EHF band are more severe simply because the shorter wavelengths at these frequencies are more likely to be scattered than reflected, resulting in less energy at the receiver. Currently, the best sources for shadowing statistics at EHF are published results from the ACTS mobile propagation experiments [13, 14, 15], which report typical means from −15 to −20 dB and standard deviations in the range of 5 to 10 dB.

In [3] a Total Shadowing Model that effectively combined the densities given by (5), (6), and (9) was proposed to describe a LMSC propagation channel at L band. A time-share parameter \(0 \leq X \leq 1\) was introduced such that a fraction \(X\) of the time the received signal power is unaffected by shadowing and described by (6), while the remaining fraction \((1 - X)\) of the time the LOS component is totally blocked (i.e., \(A = 0\)) and the received signal power follows (5) with the average power, \(S_0\), given by the log normal density in (9). Expressed mathematically:

\[
p(S) = X \cdot p(S|A, \sigma_d^2) + (1 - X) \int_0^{\infty} p(S|S_0) \cdot p(S_0) \cdot dS_0
\]

\[
= X \cdot \frac{1}{\sigma_d^2} \exp \left\{ -\frac{A^2 + 2S}{2\sigma_d^2} \right\} \cdot I_0 \left( \frac{A}{\sigma_d} \right) + (1 - X) \int_0^{\infty} \frac{1}{S_0} \exp \left( -\frac{S}{S_0} \right) \times
\]

\[
\frac{10}{\sqrt{2\pi} \sigma \ln 10} S_0 \exp \left[ -\frac{10 \log S_0 - \mu)^2}{2\sigma^2} \right] dS_0
\]

(10)

According to (10), the fading behavior of the channel consists of two dominant modes or states. In the unshadowed state (i.e., the “good” channel state) the channel is characterized by the presence of a LOS component, which implies high received power and Ricean fading, while in the shadowed state (i.e., the “bad” channel state) the channel is characterized by the absence of a LOS component, which implies low received power and Rayleigh fading. The time-share parameter, \(X\), is a long-term average that describes the fractional amount of time spent in each state. The short-term characteristics of the switching process are accurately described by a two-state Markov model [3]. The situation is depicted in Figure 3. According to the model, the mean duration, in bits, of a
Figure 3. Two-state Markov model that describes the switching process in a LMSC channel. When the channel is in the good state \( G \), there is a probability \( p_{GG} \) associated with remaining in that state and a crossover probability \( p_{GB} \) associated with the transition to the bad state \( B \) such that \( p_{GG} + p_{GB} = 1 \). Likewise, there is a probability \( p_{BB} \) associated with remaining in the bad state and a probability \( p_{BG} \) associated with switching from the bad state to the good state such that \( p_{BB} + p_{BG} = 1 \).

good or bad channel state is given by:

\[
G_b = \frac{1}{p_{GB}} \quad B_b = \frac{1}{p_{BG}}
\]  
(11)

and the probability that a good or bad channel state lasts longer than \( n \) bits is given by:

\[
p_{G}(>n) = p_{GG}^n \quad p_{B}(>n) = p_{BB}^n
\]  
(12)

In addition, the time-share parameter, \( X \), can be expressed in terms of the Markov model parameters:

\[
X = \frac{G_b}{G_b + B_b} = \frac{p_{BG}}{p_{BG} + p_{GB}}
\]  
(13)

Finally, note that the above parameters can be expressed in terms of alternate units such as meters or seconds, depending on the vehicle speed, \( v \) (meters/second), and/or bit rate, \( R \) (bits/second). For example, the mean duration, in meters of a good or bad state, is given by:

\[
G_m = G_b \frac{v}{R} \quad B_m = B_b \frac{v}{R}
\]  
(14)

In [3] the Markov model parameters were estimated by fitting the statistics to actual recorded data. The validity of the model described by (10) - (13) for the EHF LMSC channel was verified over the course of the ACTS mobile propagation experiments and values for the various model parameters, including \( X, K, \mu, \sigma, p_{GG}, p_{BB}, G_b \), and \( B_b \) were reported in [13].

2.3.3 Fading Due to Antenna Mispointing

Characterizing the fluctuations in received signal strength due to antenna mispointing is difficult because of the wide variety of factors that contribute to pointing errors. These factors include
Figure 4. Losses due to antenna mispointing for various aperture sizes. (a) Uplink losses (i.e., 44 GHz) and (b) Downlink losses (i.e., 20 GHz).

terrain, vehicle type and speed, antenna beamwidth, and the antenna controller. In general, for a given pointing error, $\theta$, the loss in received signal strength, $L_p$, in dB is given by:

$$L_p(\theta) = -12 \left( \frac{\theta}{\theta_{3dB}} \right)^2$$

(15)

where $\theta_{3dB}$ is the 3 dB beamwidth of the antenna, which is typically approximated by:

$$\theta_{3dB} = 70 \left( \frac{\lambda}{D_a} \right)$$

(16)

where $\lambda$ is the signal wavelength in meters and $D_a$ is the antenna diameter (i.e., aperture), also in meters. From (15), it is clear that smaller antenna beamwidths (i.e., larger antenna apertures) result in greater losses for a fixed pointing error. Figure 4 illustrates pointing error losses at 44 GHz (i.e., MILSTAR uplink) and 20 GHz (i.e., MILSTAR downlink) for several different antenna apertures. The pointing error itself is defined as the product of the vehicle angular velocity and the latency of the antenna controller. Hence, as mentioned previously, additional information is required to characterize the pointing error. As an example [16], consider the simulated data for a HMMWV on the Churchville course B at Aberdeen Proving Grounds. Five percent of the time the roll rates exceed 22°/sec up to a maximum of 67°/sec. Further, assuming an antenna controller with a latency on the order of 20 ms, the corresponding pointing errors for a 12-inch aperture terminal will be between 0.4° and 1.2°, yielding downlink losses between 0.3 and 1.5 dB and uplink losses between 1.0 and 7.0 dB. Note that these numbers are based on a preliminary and unsophisticated design of the antenna pointing control system.

In [15], a modification to the Total Shadowing Model in (10) that reflects losses due to pointing errors was proposed. The model is based on the observation that received signal power during the ACTS mobile propagation experiments was concentrated at two distinct levels during unshadowed
conditions. This characteristic motivated the following bimodal density function for the amplitude of the LOS component, A:

\[ p(A) = (1 - Y)\delta(A - A_1) + Y\delta(A - A_2) \]  

(17)

where \( Y \) is another time-share parameter similar in nature to, but independent of, \( X \) in (10). The time-share parameter \( Y \) and the two amplitude levels \( A_1 \) and \( A_2 \) are obviously scenario dependent. Unfortunately, no intuition that relates this model to the physical situation is provided in [15]. However, one explanation for the bimodal nature of the data is that the mobile terminal encountered relatively constant angular velocities in each of the propagation experiments. This would lead to a fixed pointing error due to vehicle turns and hence a fixed loss in received signal power. This theory also suggests that the applicability of the model in (17) to tactical environments is limited.

In summary, there are numerous factors that affect signal strength and variability as measured by a receiver. Moreover, these factors will vary widely with the operating environment. Given the range of expected channel conditions in an EHF LMSC system, the use of channel simulators, both for assessing system performance and as a design tool in the development of new systems, seems prudent.

2.4 Laboratory Simulation of the Propagation Channel

Perhaps the most accurate way to simulate a propagation channel is not to simulate it at all, but rather use recorded strips of actual data. However, with this method it is practically impossible to control channel conditions, motivating the need for laboratory simulation techniques. A simple yet useful simulator was proposed in [3] and is duplicated in Figure 5. The simulator is based on the Total Shadowing Model given by (10) and was originally proposed for reproducing received
Table 1

Parameter values used to generate the signal envelopes in Figure 6. These statistics were reported by Rice et al. in [13] and are based on real data collected during the ACTS mobile propagation experiments.

<table>
<thead>
<tr>
<th>$K$ (dB)</th>
<th>$\mu$ (dB)</th>
<th>$\sigma$ (dB)</th>
<th>$X$</th>
<th>$G_m$</th>
<th>$B_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>21.1</td>
<td>-20.8</td>
<td>-0.09</td>
<td>0.298</td>
<td>18.6</td>
<td>4.2</td>
</tr>
<tr>
<td>4.0</td>
<td>-20.3</td>
<td>-0.13</td>
<td>0.464</td>
<td>13.2</td>
<td>6.4</td>
</tr>
<tr>
<td>16.2</td>
<td>-13.7</td>
<td>11.40</td>
<td>0.101</td>
<td>34.3</td>
<td>3.3</td>
</tr>
<tr>
<td>19.8</td>
<td>-11.9</td>
<td>3.30</td>
<td>0.095</td>
<td>65.7</td>
<td>2.8</td>
</tr>
</tbody>
</table>

signal envelopes subjected to propagation effects observed in a LMSC system at L-band. However, by substituting the appropriate model parameter values, the simulator is easily extended to EHF. The model works as follows: a complex Gaussian process is filtered to produce a desired multipath fading characteristic. In the event that Ricean fading is desired (i.e., non-shadowed propagation), a LOS component is added based on a given Rice factor $K$. In the event that Rayleigh fading is desired (i.e., shadowed propagation), the fading envelope is scaled based on given log-normal statistics. The switching between non-shadowed (i.e., good) vs. shadowed (i.e., bad) states is based on the parameters of the two-state Markov model (i.e., $p_{GG}$ and $p_{BB}$). The fading coefficient generated by the above process is used to scale a signal waveform, $s(t)$. Finally, noise is added to produce a simulated received signal, $r(t)$.

The simulator in Figure 5 can be used to produce a wide range of channel conditions depending on the input statistics. In practice, these statistics would have to be measured over a variety of operating environments. The results of just such a measurement campaign conducted at 20 GHz with ACTS were reported in [13]. Based on these statistics, some of which are reproduced in Table 1, the simulator in Figure 5 was used to generate representative signal envelopes. The results are depicted in Figures 6a-d. For convenience the envelopes were normalized such that the mean signal value in a good state was 0 dB. However, the envelopes could just as easily have been normalized to a value more appropriate for a different scenario (e.g., rain).

2.5 Conclusions

In summary, a few generalizations regarding the EHF LMSC propagation channel can be made. First, while path losses due to weather and other atmospheric effects are more severe at EHF compared to lower frequencies, additional link margin can typically be used to compensate in these situations. On the other hand, the effects of shadowing by foliage, buildings, hills, etc. at EHF are so severe that any attempt to overcome them with additional link margin is impractical. A relatively simple yet accurate description of the fading behavior of the EHF LMSC channel is given by the so-called two-state model. This model describes the channel at any given time as being in either a “good” state (i.e., non-shadow region) where a LOS path exists and communications is possible or a “bad” state (i.e., shadow region) where no LOS path exists and, in the absence of specialized algorithms, communications is not possible. Given the harsh effects of signal shadowing at EHF, it seems reasonable to explore algorithmic strategies, such as coding, interleaving, and specialized protocols for overcoming them.
Figure 6. Various signal envelopes produced by the Lutz channel simulator. Envelopes depicted in Figures (a)-(d) were generated based on the statistics in rows 1-4, respectively, of Table 1. These statistics were derived from ACTS data and reported in [13].
3. ERROR CONTROL AND TIMING

This chapter of the report is concerned with error control for the EHF LMSC channel. The primary focus will be on algorithms that can be used to facilitate communications despite the signal blockage that occurs in shadow regions.

3.1 Introduction

The two primary classes of error control to be considered here are forward error correction (FEC) coding and automatic repeat request (ARQ) protocols. Simply stated, FEC coding is a process of adding redundancy to a bit stream according to some rule known to both the transmitter and receiver. In the event that the data stream is somehow corrupted by channel impairments, this redundancy can be used to compensate. As the name implies, FEC coding algorithms operate on the forward link. No feedback or return link is necessary. As opposed to FEC coding, ARQ schemes address the presence of errors in a block of data by requesting a retransmission of the packet rather than attempting to correct it at the receiver. Of course, the existence of a feedback path (i.e., a return link) is required with such an approach. Typically, coding is still used in ARQ strategies, but only to alert the receiver to the presence of errors, not to correct them. Since the probability of an undetected error is usually much smaller than the probability of a decoding error, ARQ schemes are an inherently more reliable form of error control than FEC coding. Detailed explanations of specific error control algorithms are beyond the scope of this report. Instead, a basic familiarity with FEC coding and ARQ protocols is assumed. Additional information on specific procedures can be found in [12, 17, 18].

In the following sections the performance of both forms of error control for the EHF LMSC channel is investigated. The primary performance metrics are throughput and delay. Voice and data communications are treated separately, due to the differing characteristics between these two classes of information. The fundamental conclusion drawn here is that the use of FEC coding for the purpose of correcting errors/erasures caused by signal shadowing is ill-advised. Long interleaving delays and high coding overhead are the primary drawbacks with this approach. On the other hand, ARQ and certain combinations of ARQ and FEC coding, where the FEC coding schemes are designed only to correct random errors in the channel good state, give acceptable performance. Before the details of the analysis are presented, it is first necessary to define several representative channels based on the parameters introduced in Chapter 2.

3.2 Channel Definitions

In practice, an EHF LMSC system user might expect to face a continuum of channel conditions. However, for ease of analysis and clarity of presentation, it is useful to define several typical channels. Specifically, three channels will be considered: the open channel, the rural channel, and the urban channel. All three channels are assumed to have the same data rate, 4800 bps, and vehicle speed, 15 meters per second (i.e., approximately 30 mph). The distinguishing characteristic of the channels is the value of the good state time-share parameter, \( X \). For the open channel \( X = 0.90 \), for the rural channel \( X = 0.65 \), and for the urban channel \( X = 0.35 \). Moreover, for the assumed vehicle speed, the average length of time in the channel good and bad states are 10 seconds and 1 second, respectively, for the open channel; 7 seconds and 3 seconds, respectively, for the rural channel; 17
Table 2
Parameters that define the open, rural, and urban channels. In all cases the channels have a data rate of 4800 bps and a constant vehicle speed of 15 meters per second (i.e., approximately 30 mph). In addition, the random BER in the channel good state is assumed to be 0.01.

<table>
<thead>
<tr>
<th>Type</th>
<th>X</th>
<th>$G_t$ (sec)</th>
<th>$B_t$ (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>open</td>
<td>0.90</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>rural</td>
<td>0.65</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>urban</td>
<td>0.35</td>
<td>3</td>
<td>7</td>
</tr>
</tbody>
</table>

and 3 seconds and 7 seconds, respectively, for the urban channel. Table 2 summarizes the channel parameters. In all cases, it is assumed that the bit error rate (BER) is 0.01 in the channel good state. This BER corresponds to a bit energy to noise ratio, $E_b/N_0$, of approximately 6 dB assuming DPSK modulation, a Rice factor of $K = 20$ dB, and no additional margin. Initially, no losses due to antenna mispointing are assumed. However, the impact of antenna mispointing errors on the performance results are discussed briefly at the end of the voice and data subsections. For simplicity, point-to-point communications links are assumed. In addition, unless otherwise stated it is assumed that only one of the two communications terminals is mobile and the fixed terminal is presumed to be free of any obstructions that might contribute to signal blockage.

3.3 Strategies for Voice Communications

Before discussing error control, a few observations concerning tactical voice communications should be made. First, adequate voice quality and intelligibility can be obtained in the presence of bit errors. While the maximum tolerable bit error rate is largely a function of the vocoding algorithms, in general, uncorrelated bit errors on the order of 1% are acceptable. Moreover, uncorrelated packet\(^1\) losses on the order of 1% can also be tolerated in addition to the bit errors [19]. Second, typical transmission times in tactical scenarios are quite short. Figure 7 illustrates this point. The figure is a histogram of utterance lengths collected from a mock tank battle conducted at the National Training Center (NTC) in 1999 [20]. The histogram is comprised of over 650 speech transmissions and yields an average duration of slightly less than 3 seconds. These characteristics of tactical speech will be exploited subsequently by the error control algorithms.

3.3.1 FEC Coding

Note that when the channel is in a good state, the requirements for acceptable voice communication (i.e., random BER of 1% or less) are inherently satisfied. Of course, some form of error control must still be used since the channels, as defined above, would result in more than 1% packet loss due to signal outage (i.e., the time-share parameter for each of the channels is less than 99%). Therefore, the objective of any coding strategy designed to facilitate voice communication over the EHF LMSC channel is to mitigate the effects of channel blockage 99% of the time while

\(^1\)The general term "packet" is used throughout this section of the report to refer to a block of data. In the speech processing community a block of speech samples to be processed by a vocoder is also sometimes referred to as a frame.
simultaneously maintaining an output BER of 0.01. In order to meet this objective, some form of interleaving is required. Interleaving is a simple process of temporally scrambling the data such that those bits affected by channel blockage are distributed somewhat uniformly among all the packets, thus allowing traditional coding schemes to effectively correct the errors in each packet. The implicit assumption here is that the packet size (i.e., codeword length) is small compared to the average blockage interval, so that in the absence of interleaving a typical packet would tend to be either all good (i.e., no bits blocked) or all bad (i.e., all the bits are blocked). Of course, if the packet is all bad, then no amount of FEC coding will recover the data. Therefore, interleaving is necessary to decorrelate the channel and essentially enforce the long term (i.e., average) channel behavior over the short term. Unfortunately, interleaving introduces a delay between the transmitter and receiver due to the buffering required at each end of the link. However, the only alternative to interleaving is to use packet sizes large enough to ensure that each one is average in the sense that approximately $X\%$ of the bits in each packet are from the channel good state and $(1 - X)\%$ are from the bad state. For the channel conditions defined in Table 2, this would require codewords that are tens of thousands of bits (i.e., many seconds) long and therefore far too complex to be practical.

The degree of interleaving required to sufficiently scramble the channel bits is a function of not only the average good and bad state durations but also the packet size. For the remainder of this report, the assumed packet size will be approximately 500 bits. This choice of packet size reflects a tradeoff between so-called packetization delay, which in order to be minimized requires shorter packets, and overhead efficiency, which in order to be maximized requires longer packets. With 500-bit packets and a 2400-bps vocoder, the packetization delay is approximately 200 ms.\(^2\) Moreover,\(^2\) The 200-ms packetization delay assumes that all 500 bits in each packet are from the vocoder. This assumption is accurate for the ARQ case. However, for the case of FEC coding, only a fraction $R$ of each packet will contain information bits while the rest will represent coding redundancy. Consequently, the packetization delay for FEC coding will be approximately $R \frac{500}{2400}$ seconds, where $R$ is the code rate.

\(^2\)
Table 3

Interleaving delays for the various channels. The results were determined via simulation and assume the use of convolutional interleavers.

<table>
<thead>
<tr>
<th>Type</th>
<th>X</th>
<th>Interleaver delay (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>open</td>
<td>0.90</td>
<td>25</td>
</tr>
<tr>
<td>rural</td>
<td>0.65</td>
<td>55</td>
</tr>
<tr>
<td>urban</td>
<td>0.35</td>
<td>65</td>
</tr>
</tbody>
</table>

assuming relatively modest sized packet headers (i.e., 40-50 bits), which contain information such as packet sequence number, the overhead can be kept to approximately 10% or less. Note the distinction here between overhead associated with packet headers, which will be incurred by all error control schemes, and overhead associated with coding redundancy, which will be incurred only by error control schemes that use FEC coding and will vary depending on the code rate. Table 3 summarizes the results of an experimental determination of the interleaver lengths required for each of the three channels. In the table, end-to-end interleaver delay is recorded based on the use of convolutional interleavers in simulations designed to produce 99% of the packets with a good-to-bad bit distribution equal to the channel time-share parameter X. It is interesting to note that for a fixed packet size, the degree of channel interleaving required is inversely proportional to vehicle speed. In other words, the higher the mobile speed, the less interleaving required. This observation follows from the fact that, relative to a fixed time interval (i.e., a packet length), a fast moving vehicle will experience a higher rate of fluctuation in received signal power due to shadowing compared to a slower vehicle (i.e., Gi and Bi decrease as the vehicle speed increases). Unfortunately, adapting the interleaving strategy to suit varying channel conditions (i.e., vehicle speeds) is virtually impossible. Hence, the interleavers would need to be designed for worst case scenarios, which will lead to, in many cases, unnecessarily long channel delays.

Assuming sufficient interleaving, the remaining issues to be addressed are the choice of coding scheme and appropriate code rates to be used for each of the channels. The coding strategy considered here is to use Reed-Solomon (RS) codes with error and erasure decoding. RS codes are popular for two main reasons. First, they are maximum distance separable, implying that they are optimal from a coding efficiency viewpoint. Second, efficient decoding algorithms exist for RS codes. The RS codes considered here are length n = 64 symbols, shortened from a (256,k) parent code so that codeword symbols each consist of 8 bits. Hence, all codewords have a length of 512 bits. As with any coding scheme, the use of erasure decoding algorithms is desirable since twice as many erasures can be corrected as errors. For the EHF LMSC system considered in this report, it will be assumed that erasures are declared by monitoring the received signal power. In other words, those bits with a received power below some threshold will be flagged and treated by the decoder as erasures. The appropriate code rate to use will depend on the shadowing statistics of the channel. Below, the maximum supportable code rates for each of the three channels defined in Section 3.2 are determined for two different scenarios: a transponded satellite (e.g., some commercial systems) and a processing satellite (e.g., MILSTAR). With respect to FEC coding, the fundamental difference between transponded and processing satellites is that with transponded systems, blockages (i.e., erasures) that occur on the uplink can still be detected on the downlink. On the other hand, a processing satellite will hard-limit the uplink so that any “soft” information is lost on the down-
link, requiring uplink bit errors due to blockage to be treated as errors instead of erasures.\footnote{In principle, there is no reason that a processing satellite could not perform decoding on-board. This strategy would allow the use of erasure-based algorithms on the uplink and yield improved performance. However, MILSTAR only demodulates the uplink, no decoding is performed. Hence, this model will be assumed for the case of a processing satellite.} The implication of this difference is that, in general, lower rate codes will be required to achieve the same level of performance with processing satellites compared to transponded satellites.

Although MILSTAR is a processing satellite, it is still worthwhile to examine the error control performance over a transponded system because many commercial satellites are transponded. As mentioned earlier, the advantage of working with transponded systems is that erasure decoding algorithms can be used. Figure 8 illustrates the maximum code rates that can be supported over each of the three channel types with erasure decoding. In the figure, output BER is plotted against the channel time-share parameter $X$. In the analysis, probability of erasure is assumed to be approximately $1 - X$. For example, the open channel has a time-share parameter of $X = 0.9$, implying that 90% of the time the channel is in a good state, or conversely, 10% of the time the channel is in a bad state where the bits are erased. The results were calculated using the simple formula [21]:

$$P_e = \sum_{j=r}^{n} \binom{n}{j} \frac{j}{2(n-1)} (1-p_e)^{n-j}$$

(18)

\textbf{Figure 8.} Output BER with erasure decoding. These results show that in order to support voice service (i.e., maximum BER = 0.01) a code of $R = 42/64$ is required for the open channel (i.e., $X = 0.9$), $R = 28/64$ for the rural channel (i.e., $X = 0.65$), and $R = 4/64$ for the urban channel (i.e., $X = 0.35$).
Figure 9. Output BER with error decoding. These results show that in order to support voice service (i.e., maximum BER = 0.01) a code of \( R = \frac{22}{64} \) is required for the open channel (i.e., \( X = 0.9 \)) and \( R = \frac{2}{128} \) for the rural channel. For the urban channel, the probability of error is greater than 1/2, implying that no FEC coding scheme can be used for this channel.

where \( P_e \) is the output probability of error, \( p_s \) is the input symbol error rate (i.e., approximately 8 times the input bit error rate, \( p_b \)), and \( r \) is a measure of the redundancy in the codeword. If erasure correction is used, then \( r = n - k \), where \( n \) is the codeword length and \( k \) is the length of the information sequence. On the other hand, if error correction is used, then \( r = \frac{n-k}{2} \) (i.e., the error correcting capability is equal to one-half the erasure correcting capability) [21]. The criteria for determining the appropriate code rate is that the output BER is less than the maximum acceptable BER for voice (i.e., 0.01), given the time-share parameter (i.e., probability of erasure) for the channel. Note from the figure that the maximum supportable code rate for the open channel is \( R = \frac{42}{64} \), for the rural channel \( R = \frac{28}{64} \), and for the urban channel \( R = \frac{4}{64} \).

As mentioned earlier, the fact that a processing satellite hard-limits on the uplink implies that erasure decoding can not be used in an end-to-end fashion. Consequently, error decoding must be used. Figure 9 illustrates the maximum code rates that can be supported over the channels with error decoding. As with Figure 8, output BER is plotted against the channel time-share parameter \( X \). In this case, the input probability of symbol error is taken to be \( \approx 0.875(1 - X) + 0.08X \). This relationship reflects the fact that with 8-bit code symbols the maximum BER, which occurs in shadow regions, is 0.875, while in the channel good state the symbol error rate is 0.08 (i.e., \( 8p_b \)). In this analysis, the probability of erasure is taken to be 0. Note from the figure that a rate \( R = \frac{22}{64} \) code is needed to support voice service over the open channel with error decoding. For the rural channel, the maximum code rate is \( R = \frac{2}{128} \) (i.e., 1024-bit packets are required with approximately 98% coding overhead). In both cases, these code rates are well below those required
when erasure decoding is used. Unfortunately, since the urban channel yields a probability of error greater than 1/2, no amount of FEC coding will be sufficient. This observation can be explained by noting that the error correcting capability of a block code is directly related to its minimum distance so that any combination of \( t \) errors can be corrected as long as:

\[
d_{\text{min}} > 2t
\]  

(19)

When the channel probability of error is greater than 1/2, \( t \) will be greater than \( n/2 \), and since the minimum distance of a codeword will never exceed its length, \( n \), FEC coding fails in this case. Finally, it should be noted that MILSTAR low data rate (LDR) channels use 2400-bps vocoders and support data rates up to a maximum of only 4800 bps. For this reason, the lowest code rate that can be supported for voice service over a MILSTAR LDR channel is \( R = 1/2 \). Unfortunately, this implies that in the absence of higher rate channels, such as those provided by the medium data rate (MDR) system, or lower rate vocoders (e.g., 1200-bps MELP), voice service could not be supported in any of the environments defined in Section 3.2 with error decoding.

For the special case of a processing satellite where one terminal is mobile and the other is fixed and presumably free of obstructions, it may make sense to use different FEC coding schemes depending on whether the mobile terminal is transmitting or receiving. For example, as indicated in Figure 9, relatively high rate codes are required when error decoding is used, as would be necessary for mobile terminal transmissions that are hardlimited by a processing satellite. On the other hand, when the mobile terminal is receiving, erasure decoding can be applied, which according to Figure 8 suggests that higher rate codes could be used on this portion of the link. Use of different rate codes on symmetric fixed-to-mobile and mobile-to-fixed links would result in asymmetric information rates to and from the mobile terminal. Specifically, the mobile terminal could receive at higher rates compared to those at which it could transmit. However, in many cases this type of performance might be acceptable. For example the mobile terminal might belong to a voice net where the intent is to monitor a situation by simply listening to the other participants. Depending on the specifics of the operation, the mobile might send short, text-based messages in response to the voice traffic it receives.

In the following subsection, an alternate approach, based on ARQ protocols will be examined. In addition to improved throughput and reduced latency, advantages of ARQ include the fact that this approach is inherently adaptive and can be used to exploit the short transmission times of tactical speech.

### 3.3.2 ARQ Approaches

In a commercial telecommunications system, ARQ would not even be considered as a viable form of error control for voice service. The reason is that delays associated with requesting a repeat transmission and then waiting for this repeat transmission to arrive are too long to guarantee commercial-grade service. This problem is even worse in systems that use satellite links, where the roundtrip time is relatively long. However, tactical users may be willing to accept some delay in exchange for the robustness that is provided by an EHF satellite communications system. Given this fact, together with the relatively poor performance of FEC coding, it makes sense to explore ARQ as an error control option for voice service. It will be shown in this section that the use of ARQ results in significantly less end-to-end delay when compared with the FEC coding approach described previously. The reason ARQ outperforms FEC with respect to latency is that no interleaving is required with ARQ. Instead, packets that are corrupted by shadowing are simply retransmitted.
This form of error control results in latencies that are more closely matched to the channel outage times, as opposed to 10 times longer, as with FEC. Other potential advantages of the ARQ approach to error control include the following. First, latencies are directly proportional to transmission time so that shorter transmissions experience less overall delay. Second, in contrast to FEC coding where higher data rates are necessary to support the coding overhead required for the rural and urban channels, ARQ can be used in these environments regardless of the channel data rate. Lastly, ARQ represents an inherently more adaptive form of error control for the EHF LMSC channel. In other words, while FEC algorithms must typically be designed for the worst case, ARQ has the advantage that when the channel is well behaved, no action is taken, thus improving latency and throughput. These aspects of ARQ will all be revisited in the simulation results that are presented below.

Based on the channel definitions in Section 3.2, simulations have been conducted using the selective repeat ARQ protocol. As mentioned earlier, with the ARQ approach channel latencies are directly proportional to the transmission time. Consequently, a range of transmission times were simulated for each of the three channels. In the simulations, 500-bit packets were used and repeat requests were based on whether or not the channel was in a bad state. Perfect detection of the channel bad state was assumed. Latency results are depicted in Figure 10 assuming MILSTAR LDR channel conditions (i.e., 4800 bps). As the figure illustrates, delays are proportional to transmission length. This observation stands to reason since longer transmissions will be subject to more signal outages and the subsequent delays are cumulative. Clearly, the worst delays are associated with the urban channel where the fractional amount of time spent in the channel bad state is the largest, followed by the rural and open channels.

As mentioned earlier, ARQ has the advantage that it can be used even when the channel data rate is quite low. However, when higher data rates are available, the extra capacity can be used to improve performance with respect to channel delay. Figure 11 illustrates this result for the urban channel where the effect is most dramatic. In the figure, simulation results for four different data
rates are presented. The data rates are each characterized relative to the source (i.e., vocoder) rate: $1 \times$, $2 \times$, $4 \times$, and $16 \times$. For the MILSTAR system where the vocoder rate is 2400 bps, these data rates correspond to 2400 bps, 4800 bps, 9600 bps, and 38400 bps, respectively. The improvement in delay performance for higher data rate channels can be explained as follows. In an ARQ scheme, packets are stored in a buffer until they are positively acknowledged. Assuming a constant rate source (e.g., a vocoder), the contents of this buffer will grow during periods when the channel is not available. The advantage of higher rate channels is that multiple packets can be sent simultaneously, allowing the transmitter to “catch up” with the source when the channel is suddenly available after an outage. With lower rate channels the transmitter either never catches up or catches up at a slower rate, exposing the buffer contents to more channel time, and hence the possibility of more delays.

Characterizing the average delay for tactical speech transmissions can be accomplished by integrating the delay curves in Figure 10 over a typical transmission profile, such as depicted in Figure 7. Such an approach will emphasize the shorter transmissions and yield a single number, in seconds, for the average delay. Table 4 summarizes the results of such an exercise. In the table, several different channel rates are compared for each of the three channels defined in Section 3.2. Note from the table the benefits of extra channel capacity (i.e., increased data rates).

One last point to be made with respect to channel delays for ARQ is that the delays are interspersed with the speech. Unfortunately, this degrades the subjective quality of the transmission. One solution to this problem is to perform buffering at the receiver so that the listener does not perceive any pauses in the transmission. However, determining the proper amount to buffer is virtually impossible because the receiver does not know the original transmission length a priori. In lieu of optimal buffering, which depends directly on the original transmission length, at least two choices are available. First, buffering based on the average expected transmission length could be

Figure 11. The effects of greater capacity on channel delay performance. Results for several different bandwidths are compared for the urban channel. Note that the most significant improvements are obtained with relatively little excess channel capacity.
Table 4

Average delays, in sec., for different channel capacities. Results were calculated by averaging delay results over a typical voice utterance profile.

<table>
<thead>
<tr>
<th>Type</th>
<th>1×</th>
<th>2×</th>
<th>4×</th>
<th>16×</th>
</tr>
</thead>
<tbody>
<tr>
<td>open</td>
<td>1.04</td>
<td>1.01</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>rural</td>
<td>3.41</td>
<td>3.02</td>
<td>2.89</td>
<td>2.88</td>
</tr>
<tr>
<td>urban</td>
<td>11.93</td>
<td>8.60</td>
<td>7.90</td>
<td>7.43</td>
</tr>
</tbody>
</table>

Table 5

Throughput results for ARQ. For each channel, the number of retransmissions per packet, $A_p$, was determined via simulation.

<table>
<thead>
<tr>
<th>Type</th>
<th>$X$</th>
<th>$\eta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>open</td>
<td>0.90</td>
<td>0.79</td>
</tr>
<tr>
<td>rural</td>
<td>0.65</td>
<td>0.52</td>
</tr>
<tr>
<td>urban</td>
<td>0.35</td>
<td>0.20</td>
</tr>
</tbody>
</table>

performed. This approach would result in relatively small buffering delays but performance would degrade rapidly for longer transmissions. The second approach is to employ some sort of adaptive buffering scheme [22]. While buffering will improve the subjective quality of the transmission by eliminating or at least reducing the pauses, it will unfortunately increase the end-to-end channel delays beyond those values reported in Table 4.

Aside from delay, another performance parameter of interest is channel throughput. Recall that for the case of FEC coding, the throughput is easily approximated by the code rate, $R$. For the case of ARQ, the throughput is determined by

$$\eta = \frac{1}{A_p} \left( \frac{k}{k + h} \right)$$

(20)

where $\eta$ is the throughput, $A_p$ is the number of transmission attempts per packet, $k$ is the number of information bits, and $h$ is the number of header bits in the packet. As mentioned earlier, the total packet size (i.e., $k + h$) is taken to be 500 bits. Moreover, 50-bit headers are assumed so that $k/(k + h) = 0.9$. Ideally, the first term in (20) is exactly equal to the channel time-share parameter, $X$. However, due to the fact that a retransmission is requested even for packets that are only partially blocked, $1/A_p$ will always be strictly less than $X$. Table 5 summarizes the throughput performance for each of the three channels with ARQ. In each case, $A_p$ was determined experimentally via simulation.

In comparing the two forms of error control for voice service, several observations can be made. First, from a performance viewpoint, the ARQ approach offers superior throughput with less overall delay compared to FEC coding. This fact is illustrated in Figure 12. In the figure, throughput vs. delay is plotted for FEC coding with erasure decoding and ARQ for each of the three channel types defined in Section 3.2. In the case of ARQ, multiple values for the channel data rates (i.e., 1×,
2x, 4x, and 16x) are shown. Moreover, theoretical maxima for throughput are given by the solid vertical lines while the theoretical minimum for delay, assuming a geostationary satellite channel, is given by the solid horizontal line. Note that in all cases ARQ outperforms FEC coding with respect to both throughput and delay. In addition, ARQ has the advantage that it can be used over low data rate channels, even in urban environments where the time-share parameter, \( X \), is quite small. For FEC coding to be effective in these environments, a relatively high rate channel is necessary to support the large coding overhead. In addition, ARQ has the advantage that it is an inherently adaptive form of error control. In other words, while FEC algorithms must typically be designed for the worst case, ARQ has the advantage that when the channel is well behaved, no action is taken, resulting in improved performance.

As mentioned in the channel definitions section of this report, two key assumptions have been used in the preceding analysis: no losses due to antenna mispointing are present and only one end of the bidirectional link is mobile. Although these issues have not been studied in detail, a few comments regarding the implications of relaxing these assumptions can be made. For example, the net result of antenna mispointing errors will be an increase in the BER in the channel good state. If FEC coding were to be used for error control, then lower rate codes compared to the ones in Figures 8 and 9 would have to be used. However, given that the fades caused by antenna mispointing are relatively short compared to the fades caused by signal shadowing, additional interleaving would probably not be necessary. On the other hand, with the ARQ form of error control a relatively high rate code would likely be necessary to ensure that the threshold BER (i.e., 0.01) is maintained in the channel good state. Moreover, a modest amount of interleaving may also be required. Nonetheless, the ARQ form of error control will almost certainly still outperform FEC coding even when antenna mispointing is included in the analysis. For the case where mobility is permitted at both ends of the link, the net result would be a decrease in the time-share parameter,
$X$, for each of the channels. This decrease would degrade the performance of both forms of error control discussed in this section. However, it is likely that ARQ would still outperform FEC coding.

3.4 Strategies for Data Communications

The primary difference between voice and data service is that whereas some errors can be tolerated with voice transmissions, data transmissions must be virtually error free. For this reason, ARQ is almost always used to ensure reliability for data transmissions. Hence, the remaining issues to be addressed here are if and how FEC coding should be used with ARQ to yield the best performance. Again, the performance parameters of interest are delay and throughput. However, for data transmission, throughput is probably the more important of the two, as some delay can be tolerated with data. The same channels defined in Section 3.2 are assumed here. Moreover, 500-bit packets will be assumed. Recall that the choice of 500-bit packets was dictated largely by the need to minimize the packetization delay for voice. While no such concern is relevant with data transmission, the use of similar-sized packets should facilitate the handling of both voice and data by the same link layer protocol. Three approaches to error control will be evaluated in this section:

1. ARQ only with no FEC coding
2. ARQ with FEC coding and no interleaving
3. ARQ with FEC and interleaving

The primary difference between the last two strategies relates to the way in which the FEC coding is used. In one case, only a modest amount of FEC coding is used to correct random errors in the channel good state. No attempt is made to correct errors caused by signal shadowing. Hence, no interleaving is required. In the other case, FEC coding is used to correct errors caused by signal shadowing. Hence, interleaving is required to decorrelate the burst errors. In the results that follow, error decoding, as opposed to erasure decoding, is assumed. As discussed previously, this model is the most appropriate for two-way communication over a processing satellite. Erasure decoding may be used in transponded systems or broadcast situations, in which case both of the last two techniques listed above would exhibit better performance compared to their error decoding counterparts. Nonetheless, their performance relative to each other should be the same regardless of whether error or erasure decoding is used.

The first approach listed above, ARQ only with no FEC coding, is appealing because of its simplicity. The idea is that only those packets affected by signal shadowing need to be retransmitted. Moreover, in theory the receiver can detect the presence of signal shadowing by simply monitoring the received signal power and comparing it to a threshold. Due to the absence of coding overhead and interleaving delay, one might expect this approach to display superior performance. However, as will be shown later in this section, the presence of a relatively large number of random errors in the channel good state severely degrades performance.\(^4\)

The second approach listed above is similar to the first in the sense that repeat transmissions are requested based on the detection of signal shadowing at the receiver. However, a relatively high rate FEC code is also used to address the presence of random errors in the channel good state. Although the code does contribute overhead, the reduction in repeat requests caused by random

\(^4\)Although not stated explicitly, some form of random error detection (e.g., a CRC code) is assumed to be in use for the purpose of ensuring high reliability.
errors more than offsets this redundancy. The net result is improved throughput relative to the ARQ-only technique with the same delay performance.

The third approach listed above is perhaps best characterized as the “brute-force” method because FEC coding and interleaving are used to address the signal shadowing problem directly. The approach is similar to the FEC coding technique described in Section 3.3.1 for voice service. The primary difference between that technique and the one proposed here is that ARQ is used in addition to the FEC coding to increase the reliability well above the 99% level required for voice. From a delay viewpoint this approach is the worst of the three because of the long interleaving times. Although delay is usually less of a concern compared to voice, the delays reported in Section 3.3.1 are significant even for data, especially relatively short data files or messages. Moreover, as discussed later in this section, the throughput efficiency achieved with this method of error control is less than that obtained with the second approach described above.

A comparison of the throughput performance for each of the three error control techniques has been conducted. As with the voice case, throughput was determined according to (20). However, as opposed to the voice case, retransmissions for data can be triggered by either signal shadowing or random bit errors. Hence, the number of transmission attempts per packet is given by:

$$A_p = A_s A_e$$

(21)

where $A_s$ is the number of transmission attempts per packet due to signal shadowing and $A_e$ is the number of transmission attempts per packet due to random bit errors in the channel good state. Recall that for voice the random BER in the channel good state was small enough to support acceptable quality voice transmissions. Hence, no retransmissions were assumed to be caused by random errors and $A_e = 1$. For data, random bit errors in the channel good state will necessitate retransmissions. Moreover, the number of transmission attempts per packet due to random errors is a function of the channel BER and the packet length, and is given by [23]:

$$A_e = (1 - p_b)^{-(k+h)}$$

(22)

where $p_b$ is the random bit error rate and $k + h$ is the total packet length (i.e., information bits plus header). In fact, given this relationship between $p_b$, $k$, $h$, and channel throughput, an expression for $k_{opt}$, the number of information bits per packet that maximizes throughput for a given $p_b$ and $h$, can be determined [23]:

$$k_{opt} = -h \ln(1 - p_b) - \sqrt{-4h \ln(1 - p_b) + h^2 \ln(1 - p_b^2)}$$

(23)

$$2 \ln(1 - p_b)$$

According to (23), the optimal number of information bits per packet when $p_b = 0.01$ and $h = 50$ is approximately $k_{opt} = 40$.

Figure 13 summarizes the results of the throughput analysis assuming $p_b = 0.01$. As in Section 3.3.2, $A_s$ was determined via simulation. Note from the figure that the first proposed approach (i.e., ARQ only with no FEC coding) performs poorly. Moreover, these results represent the best achievable with this particular form of error control given $p_b$ and $h$, since $k$ was optimized according to (23) (i.e., $k = 40$ and $k + h = 90$). If, as with the other two methods, 500-bit packets were used, the results would be significantly worse because this packet size is much too large given the relatively high channel BER. For the second approach (i.e., ARQ with FEC and no interleaving), RS codes with length 512-bit codewords were assumed. The code rate was optimized experimentally to
maximize throughput. It is interesting to note that despite the coding overhead, this approach still produces the best results because the code reduces the channel BER to a level more appropriate for the packet size. With the third approach, an adaptive FEC coding scheme that could be used over a range of channel conditions (i.e., $X$) was evaluated. Although adaptive schemes have not been discussed in this report, they do exist and are relatively simple to implement. The most popular schemes are commonly referred to as hybrid ARQ (HARQ) Type II techniques. The basic idea with these approaches is that a different FEC code is used for each retransmission. Moreover, these codes are carefully selected so that the receiver can combine multiple retransmissions in order to obtain more powerful (i.e., lower rate) codes. The end result is that the code rate is effectively adapted based on the channel conditions. In practice, arbitrary code rates can not be used. Rather, there exists a special relationship among the code rates that can be achieved with any given HARQ Type II scheme. For convenience, this relationship is ignored here and the code rates optimized for voice in Section 3.3.1 are used. While not necessarily achievable in practice, use of the optimal code rates represents a best case performance scenario for this form of error control. From the figure it is clear that with respect to throughput performance the HARQ Type II approach is only marginally better than the ARQ-only technique when channel conditions are highly favorable (i.e., large $X$). Moreover, as demonstrated in the preceding section, the channel delay associated with this form of error control is dominated by interleaving, which is on the order of 10 times the channel outage length, as opposed to the other two approaches where channel delay is relatively close to the outage time.

As with the case for voice communications, several comments should be made here regarding some of the assumptions stated in Section 3.2. First, with respect to errors caused by antenna mispointing, the net result is an increase in BER for the channel good state. This will adversely affect all three forms of error control discussed in this section. However, it is likely that their relative performance will remain unchanged. Second, in the event that both ends of the link are mobile,
Figure 14. Throughput vs. link margin for several different RS codes. These results show that when additional link margin is available, significant throughput improvements can be achieved by using either higher rate codes or perhaps no coding at all (i.e., ARQ only).

the net result would be a decrease in the channel time-share parameter, $X$. As with the antenna mispointing, all three forms of error control examined in this section would be adversely affected but their relative performance should stay the same. Finally, in the event that additional margin is available (e.g., favorable weather conditions, etc.) the BER in the channel good state would be smaller, suggesting that higher rate codes could be used for increased throughput efficiency. The reasoning behind this assertion is that when, relative to the packet length, the BER is sufficiently small enough so as not to trigger retransmissions, additional FEC coding is unnecessary and only contributes to the algorithm overhead. Figure 14 quantifies the possible gains when additional link margin is available. In the figure, the throughput of several different codes are plotted as a function of link margin relative to the 6 dB threshold figure that has been used throughout this report. The figure shows that when no link margin is available (i.e., $E_b/N_0 = 6$ dB), the rate $R = 38/64$ RS code gives the highest throughput. However, if even 1 dB of additional link margin is available, a rate $R = 48/64$ code can be used and a corresponding increase in throughput efficiency is achieved. With 4 dB in additional link margin no coding is necessary (i.e., ARQ is the only form of error control) and maximum throughput is obtained. Two points that must be considered when interpreting the results in Figure 14 are as follows. First, the results are specific to 512-bit packets with 50-bit headers. In the event that different packet and header sizes were used, the exact throughput numbers would change. However, one would still expect to see an increase in potential throughput via higher rate codes as a function of available link margin. Second, the results in Figure 14 demonstrate throughput in the absence of blockage. Hence, to get the approximate throughput efficiency when blockage is present, one would need to multiply the throughputs in the figure by the channel time-share parameter $X$. 

31
3.5 MILSTAR Specifics

Several MILSTAR-specific issues relating to the implementation of the error control algorithms discussed above must still be addressed. First, assuming that some form of ARQ is used for error control, provisions must be made for a return channel. Use of full duplex channels is probably the easiest means for providing a return channel. However, full duplex channels will often represent an inefficient allocation of resources due to the fact that the data rates necessary to support acknowledgment traffic are typically much less than those required by the information traffic. One alternative available in the MILSTAR LDR system is to use the C1 channel for acknowledgment traffic. The C1 channel is a 300-bps channel originally intended to provide end-to-end in-band control capability. Although this capability will not exist in the MDR system, it may serve as a convenient means to conserve LDR resources. Also on the subject of acknowledgments, it should be noted that both the MILSTAR uplink and downlink will be highly correlated with respect to their shadowing behavior. This fact impacts the acknowledgment strategy in the following way. Typically, ARQ systems rely on positive and negative acknowledgments (i.e., ACK and NAK, respectively) to inform the sender as to the status of their message. However, when both channels are highly correlated, as they are here, it is quite likely that the shadowing event that led to the generation of a NAK would also prevent the NAK from reaching the other end of the link. Hence, the strategy recommended for ARQ over MILSTAR is to ACK messages only and treat the absence of an ACK as a NAK. There are still some inefficiencies with this approach because there will be times when a packet is successfully received just prior to a shadow event. In these cases the ACK will not get through due to shadowing and the transmitter will unnecessarily retransmit the packet. However, it is expected that these situations will be rare.

Other aspects of the error control implementation are related to the MILSTAR operational specification. For example, serial or stream-based encryption has been specified for use over the MILSTAR system. However, the serial encryption process will severely degrade the performance of mobile terminals because even relatively short channel outages will cause the cryptographic devices to lose synchronization with one another and force a time-consuming reacquisition process. Clearly, block encryption techniques are more compatible with the packet oriented error control schemes (i.e., block FEC coding and ARQ) discussed here. Switching from serial to block encryption can be accommodated with a hardware/software change in the end-user equipment.

Another issue to be addressed is the fact that an exact form of error control has been officially designated for use over MILSTAR. Specifically, a rate \( R = 1/2 \), constraint length \( L = 7 \) convolutional code is to be used. Moreover, several different interleavers are specified as part of the system. As with the encryption, the official MILSTAR coding scheme can be used with a mobile terminal but with degraded performance compared to some of the approaches discussed above. This fact is illustrated in Figure 15 [24]. In the figure, output BER is plotted versus input BER. Because the MILSTAR FEC code is binary, the input bit error rate as a function of the channel time-share parameter, \( X \), is given by \( 0.5(1-X)+0.08X \). Using this relationship and assuming that the maximum tolerable output BER is 0.01, it is clear that the smallest value of \( X \) is 0.95. Or, stated another way, the only environment in which a mobile MILSTAR terminal restricted to the official \( R = 1/2 \), \( L = 7 \) convolutional code could successfully operate is one where the signal blockage comprises only 5% of the channel time. Moreover, assuming that the required interleaving time is roughly 10 times the outage period, the maximum tolerable shadow duration for the MILSTAR waveform is only on the order of one second. Clearly, when restricted to official MILSTAR error control techniques, the utility of a mobile MILSTAR terminal is limited. For purposes of demonstration, it should not
Figure 15. BER performance of an $R = 1/2$, $L = 7$ convolutional code. Results are plotted versus input BER, which can be directly related to the channel time-share parameter, $X$.

be difficult to experiment with alternate forms of encryption and error control as outlined in the preceding sections.

3.6 Timing Considerations

Timing acquisition and tracking are very important considerations in an EHF LMSC system. In general, timing acquisition is performed by processing a series of increasingly "fine" synchronization patterns that repeat at known intervals. Consequently, as long as there exists sufficient link margin to process the synchronization patterns, timing acquisition should be possible. From the perspective of an EHF LMSC system, this suggests that as long as the terminal is not in a shadow region and the antenna is being pointed with sufficient accuracy, system timing can be acquired while OTM. On the other hand, given the typical signal levels experienced in shadow regions as discussed in Chapter 2, timing acquisition will, in general, not be possible in these cases.

Once timing has been acquired, a number of factors may conspire to introduce errors, including oscillator drift and ephemeris errors. Typically, synchronization is maintained through the use of an early/late gate circuit. This circuit computes correlations against the synchronization pattern. The first correlator has a reference delayed (i.e., late) by a small amount, $\delta$, relative to the synchronization pattern while the other uses a reference signal that is advanced (i.e., early) by an equal amount relative to the synchronization pattern. Moreover, an error signal, nearly proportional to the difference between these two correlator outputs, is generated. Since the autocorrelation of the synchronization pattern is an even function, the difference in magnitude between the early and late correlators (i.e., the error signal) should be zero when timing is synchronized. If the timing is off such that the error signal is not zero, then the clock is either advanced or retarded depending on the sign of the error.

The fundamental issue with respect to the EHF LMSC system is how to maintain synchronization in a shadow region. Clearly the traditional early/late approach is no longer sufficient because
the synchronization pattern is not available, implying that the circuit is operating on noise and that timing updates are now likely to cause the system to actually lose synchronization as opposed to maintain it. In fact, if it were possible to detect the presence of a shadow region, perhaps via estimation of the received signal level, it would usually be desirable to temporarily suspend operation of the early/late gate circuit. Moreover, assuming that timing errors accumulate at a relatively slow rate, it may be acceptable to simply “coast” through a shadow region without timing updates. As long as synchronization could be maintained to within some threshold, the acquisition process would not be invoked at the end of the shadow region. Of course, this approach is highly dependent on the nature of the timing errors and the duration of the shadow regions. One way to improve the likelihood that such an approach would work is to use highly accurate and stable oscillators such that errors due to oscillator drift are minimized. Unfortunately, this approach implies an increase in cost for the terminal. Another approach, currently under investigation at MIT Lincoln Laboratory [25], is to compensate for timing errors due to oscillator drift and ephemeris in a deterministic way and only employ an early/late gate circuit for other random miscellaneous errors. If this approach were to be adopted by an EHF LMSC terminal, then certain timing errors could be corrected even in shadow regions, increasing the length of time that synchronization could be maintained during these intervals. Clearly, more work is necessary to characterize the timing requirements of an EHF LMSC terminal and the performance of various algorithms for satisfying these requirements.

3.7 Conclusions and Recommendations

In summary, the FEC coding strategy for both voice and data communications over the EHF LMSC channel requires relatively long interleaving times (i.e., 10s of seconds) and large amounts of coding overhead. Moreover, these results are fixed in the sense that FEC coding parameters are typically chosen based on the worst expected case and offer no through put or latency improvements when channel conditions exceed expectations. In addition, the FEC coding approach does not exploit the relatively short transmission times of tactical speech. On the other hand, ARQ protocols were found to offer superior through put and delay performance compared to FEC coding for both voice and data. For voice, pure ARQ with no additional FEC coding was found to be sufficient due to the fact that some errors can be tolerated in voice transmissions. However, for data, a modest amount of FEC coding designed to minimize repeat requests caused by noise in the channel good state is necessary. High rate RS codes are recommended because of their excellent performance and relative ease of implementation.

With regard to MILSTAR-specific issues, several points should be made. First, the MILSTAR waveform can be used with no modification whatsoever to achieve OTM capability. However, operation would be restricted to only the most benign (i.e., virtually blockage-free) environments. Second, the best solution for providing OTM services with MILSTAR is to bypass or “turn off” the MILSTAR coding, interleaving, and encryption features and use the techniques described in this section of the report: ARQ with FEC and block encryption. Third, if for some reason users were not allowed to bypass the coding and encryption algorithms specified as part of the MILSTAR waveform (i.e., backwards compatibility were mandated), the basic OTM capability of the system could still be improved somewhat by implementing the error control techniques described above “on top” of the MILSTAR waveform. Of course, there would be a great deal of overhead associated with doing so and the performance would still not be optimal due to the propensity of the serial encryption devices to lose synchronization with one another during blockage intervals.
From an implementation viewpoint, there are at least two approaches to providing error control for an EHF LMSC system. The proper long term approach would be to develop a custom link layer with the appropriate mix of FEC and ARQ as described above. Traditionally, error control protocols are implemented at the link layer and serve to condition the underlying physical channel for reliable communications at higher layers in the protocol stack. However, the development of a custom link layer is not trivial. For example, there are numerous issues related to interfacing the protocol with higher level network protocols. An alternate approach, recommended for the near term, is to implement the error control at the application layer. With this approach, tedious link layer development issues can be avoided leading to a faster implementation that may be more appropriate for demonstration purposes. Moreover, making changes to the error control algorithms based on experimental results is easier when this functionality resides at the application layer because the impact on other parts of the system is limited.
4. NETWORKING PROTOCOL CONSIDERATIONS

It is often useful to view a communications system as being comprised of several layers, where the layers represent a convenient way of partitioning the various functions of the system. This layered partitioning is an especially useful interpretation of networked communication systems, where in addition to well known tasks like modulation and demodulation, the various nodes are also responsible for functions such as traffic routing and flow control. Moreover, in cases where the communications system supports specific applications, the application functions are often viewed as simply another layer in the overall system. There are numerous models for layered communication systems. Generally, the layers can be grouped into 3 main categories: the lower layers (e.g., the physical and link layers) where data modulation and error control algorithms are typically implemented, the middle layers (e.g., the network and transport layers) where routing and other traffic management protocols are executed, and the higher layers (e.g., session and application layers) where control session processing (i.e., session setup and tear-down) and application related functions are performed.

4.1 Introduction

Because the proposed LMSC system will typically represent only one segment in a larger defense communications network, it seems prudent to examine the system from a broader, network-oriented perspective. Such an investigation is the purpose of this section of the report. There are several areas of interest. First, the impact of the LMSC physical layer (i.e., the LMSC propagation channel as discussed in Chapter 2) on middle and higher layer protocols should be assessed. Given the likelihood that standard Internet protocols (e.g., TCP and IP) will be used on at least some if not most network segments, and given that these protocols may perform inefficiently when confronted with a geostationary satellite link, special attention will be paid to approaches for enhancing or augmenting TCP to increase the overall network efficiency in the presence of a LMSC channel. In addition, the strategies examined thus far for overcoming the effects of signal blockage (i.e., FEC coding and ARQ schemes) in a LMSC system are algorithms that reside typically in the lower layers (i.e., the physical and link layers, respectively) of a communication system. However, middle and higher layer protocols are often designed to yield a robust internetwork and may be exploited by the proposed EHF LMSC system in situations where FEC coding and ARQ strategies alone are not effective (e.g., prolonged blockage intervals). Specifically, routing protocols designed to support mobility (e.g., mobile IP) and ad-hoc or two-tier network protocols are of interest. These protocols may provide a capability whereby terminals that are not blocked from communicating with the satellite can be used to assist blocked or otherwise disadvantaged terminals by rerouting and/or caching information for these terminals. These issues are addressed in the following sections.

4.2 TCP/IP and the Satellite Channel

The most prevalent protocol suite in both military and commercial networks today is TCP/IP. IP is a network layer protocol and hence, is responsible primarily for routing data. This task is

\footnote{The reference model of open systems interconnection (OSI) developed as an international standard for data networks by the International Standards Organization (ISO) includes seven layers: the physical layer, the data link layer, the network layer, the transport layer, the session layer, the presentation layer, and the application layer.}
typically accomplished through the use of routing tables that contain connectivity information for the various network nodes. These tables are used to select the appropriate output line for an incoming packet based on the packet destination information contained in the IP packet header. TCP is a reliable, end-to-end connection-oriented transport layer protocol, and as such has numerous responsibilities. These responsibilities include breaking messages from higher layers at the data source into packets and reassembling the packets back into messages at the destination, executing error recovery if the lower layers are not suitably error free, performing flow control and multiplexing/demultiplexing multiple sessions together. Another common transport layer protocol is UDP. UDP is a connection-less, inherently unreliable protocol that does nothing more than provide an interface to the network layer protocol IP. Although unreliable, UDP is popular because of its simplicity and is often used in real-time networking applications (e.g., voice over IP) where the mechanisms necessary to ensure reliability (e.g., packet retransmissions) are prohibitively time consuming. The primary advantage of employing standard transport/network layer protocols such as TCP/IP is that applications and other protocols that comply with the standard can be run over the network regardless of other system specifics. Examples of popular applications that make use of TCP include the file transport protocol (FTP), Telnet, web browsing via the hypertext transfer protocol (HTTP), and email via the simple mail transfer protocol (SMTP).

Even though the suite was designed and optimized for terrestrial networks, TCP/IP will operate over a wide variety of link conditions. However, when the assumptions around which the protocol suite is based are violated, the result is reduced efficiency and quality of service (QoS). Unfortunately, the LMSC channel represents a situation for which TCP/IP is not especially well suited. For example, in most terrestrial networks the BER is quite low (e.g., 10^-10 or less) and latency is short (e.g., less than 20 ms). On the other hand, satellite systems are characterized by links where the raw BER is much higher (e.g., 10^-2 to 10^-6) and, in the case of geostationary satellites, the latency is significantly longer (e.g., a roundtrip time of approximately 500 ms). Other factors that have been observed to adversely affect TCP performance include limited channel bandwidth and highly disparate forward and return link channel bandwidths.

The reduced efficiency of TCP/IP over a geostationary satellite link was investigated in [26] and the most significant contributors to inefficiency were identified to be the TCP ARQ and flow control algorithms. The ARQ scheme employed by TCP is the go back N approach, which is well known to be inefficient when the channel roundtrip time is large, as with a geostationary satellite system. The TCP flow control mechanism is responsible for adjusting the packet transmission rate to minimize congestion on the link. Flow control algorithms typically operate by monitoring incoming ACKs and taking appropriate action in their absence. Of course, packets can be lost for a variety of reasons including congestion, data corruption, and link outage. Unfortunately, the TCP flow control algorithm, Slow Start, does not distinguish among these possibilities. Instead, the algorithm assumes that lost packets are due to congestion, which is a legitimate assumption for the networks that TCP was originally designed and optimized to serve. However, in situations where the BER is large or link outage is common (e.g., an EHF LMSC satellite channel), the flow control mechanism will react to these events by decreasing the transmission rate, thus reducing the link utilization. In addition, the Slow Start algorithm does not allow a TCP source to transmit at a high rate until it can be sure that the transmissions will not result in congestion. As its name implies, the algorithm achieves this objective by starting transmission with a small packet size (i.e., window) and waiting for the ACK. As the ACKs are received, the window size is slowly increased until some threshold is reached. Clearly, in cases where the channel roundtrip time is long, the
algorithm is especially slow since it requires that ACKs be received before additional packets are transmitted. This is a distinct problem for the proposed LMSC system where the channel may only be available for short periods of time and hence, the desire is to transmit data as quickly as possible during these times to maximize the link utilization.

The problems associated with realizing efficient communications using TCP/IP in non-ideal environments have been recognized for some time. Consequently, numerous approaches have been proposed for addressing these difficulties. Most approaches proposed thus far fall into one of two broad classes: those that require modification to end-user applications and those that don’t.

4.2.1 TCP Modification/Replacement

One approach to improving the performance of TCP over a satellite link is to extend or augment TCP with modifications that address specific shortcomings. For example, TCP-Vegas [27] is a variant of TCP that contains a special flow control algorithm designed to improve performance under certain conditions. Another example is the transport protocol associated with the space communications protocol standards (SCPS), SCPS-TP [28]. SCPS-TP is a transport protocol based largely on TCP and UDP, but with a number of extensions designed to improve performance in a space environment. Depending on the configuration options selected, SCPS-TP can be either identical to TCP or contain any one of a large number of enhancements. A more radical approach along these lines is to remove TCP altogether and replace it with a transport protocol more suitable to the satellite channel. This is the idea behind the development of Berkeley’s Satellite Transport Protocol (STP) [29]. Of course, the advantage of TCP modification/replacement is clearly the performance improvements achieved. However, there is a severe drawback to this approach, namely that protocol extensions are elective by nature. This fact implies that both hosts participating in a connection are required to implement the extensions before performance improvements can be realized by either. This point should not be taken lightly given the difficulties associated with enforcing a uniform standard, even within a single organization.

4.2.2 Compatibility-Preserving Approaches

Given the desire to minimize the impact on the large installed base of TCP end-user equipment, a number of alternatives to the TCP modification/replacement schemes discussed above have been proposed. These approaches fall into two main categories: the use of physical/link layer schemes that condition the link so as to be more suitable for use with TCP and gateway/proxy based approaches that engage in protocol conversion over certain network segments. Note that these two approaches are not mutually exclusive and may be combined for added performance gains. Such a combination is employed by the wireless IP suite enhancer (WISE) [30], a collection of software algorithms and protocols designed explicitly for improving the performance of TCP/IP over satellite links while simultaneously preserving compatibility with TCP-based application software.

Physical and Link Layer Enhancements

The philosophy behind these approaches is that by addressing certain channel conditions (e.g., high BER) at the physical and link layers the system performance can be improved with no TCP modifications to end-user equipment. For example, FEC coding is a physical layer approach that can be used to reduce the channel BER to a level commensurate with TCP expectations. The benefits of a reduced BER include a potential reduction in the number of TCP repeat requests and
also possibly less intervention by TCP flow control mechanisms. At the link layer, a selective repeat ARQ scheme could be employed, thereby superseding the less efficient go back \(N\) strategy used by TCP. Finally, there are numerous link layer protocols that in addition to special FEC and ARQ schemes incorporate other desirable features such as packet fragmentation and fast retransmission. Examples of specialized link layers appropriate for use with TCP over a satellite link include the Lincoln Laboratory Link Layer (L4) protocol and certain so-called TCP-aware link layer protocols. While physical and link layer enhancements can be used to indirectly improve TCP performance, they do little to address the impact of Slow Start.

**Gateway Based Solutions**

Solutions that make effective use of gateways are an attractive means for improving the performance of TCP based networks over satellite links. In the sense that they are being used in this report, gateways may be defined as entities placed at the boundaries of a particular network segment (e.g., a satellite link). In the simplest case, gateways are used to divide an end-to-end TCP connection into three separate segments: source-to-gateway, gateway-to-gateway and gateway-to-destination. In addition, the gateway is responsible for converting the end-to-end (i.e., source-to-destination) traffic from TCP to an alternate protocol more suitable for the intermediate network segment (i.e., the gateway-to-gateway segment), then converting the data back to TCP at the other end. This process is also sometimes referred to as TCP splitting. The advantages of using gateways for TCP splitting as defined above are twofold. First, performance improvements via specialized intermediate protocols can be achieved without the need for every host to either modify or replace their version of TCP. Instead, only the gateways are required to undergo this process. The result is that protocols most suitable for specific link conditions are used over the appropriate segments in a manner that is virtually transparent to the end-users. The second advantage of TCP splitting is that the effects of Slow Start are less significant since the algorithm is confined to links for which it is better suited (e.g., the source-to-gateway and gateway-to-destination links). The only drawback to TCP splitting is that in order to convert traffic between TCP and other protocols the gateway must be able read the TCP packet headers. Unfortunately, this implies that TCP splitting will not work in systems where the TCP headers are encrypted unless the gateway is a trusted system and is allowed to decrypt and re-encrypt this information. Another process similar to TCP splitting is TCP spoofing. The differences between splitting and spoofing are subtle and unfortunately rarely well-defined in the literature. However, spoofing generally has a negative connotation within the networking community. Many people consider splitting and spoofing to be the same thing and use the word spoofing only to distinguish a malicious form of splitting from a beneficial or intentional one. WISE was mentioned earlier as an example of a TCP-splitting approach. However, it should also be noted that the SCPS suite could be used as part of a gateway-based solution. Although the splitting mechanism is not inherently built into the SCPS framework, there is nothing prohibiting these protocols from being used in this way.

Within the military community there is a strong desire to remain compatible with the standard Internet protocols, TCP and IP. This desire is understandable since it allows military system designers to leverage the sizable installed base of TCP/IP and exploit the vast developmental capabilities of the commercial sector, thus decreasing their own costs and design cycles. Moreover, for many military systems, TCP/IP compatibility is easy to achieve and poses no additional burdens on the system. In fact, TCP/IP is probably just as popular and widespread within the military as it is in the commercial sector. Given this widespread desire for TCP compatibility, a great deal of
effort has been placed on the development of transparent splitting schemes that utilize gateways for protocol conversion. The most significant problem associated with these approaches is that they are not compatible with encryption schemes that encrypt the TCP header. It should be noted that WISE is fully compatible with application layer encryption such as Secure Socket Layer (SSL). Moreover, many firewall systems also require access to TCP headers. Consequently, several equipment manufacturers have begun implementing a nonstandard version of IPsec\(^2\) that leaves the TCP headers unencrypted. If this practice were to become more popular, WISE would certainly benefit. Finally, it should be noted that WISE was developed with a static satellite channel in mind. While changes to parameters in the L4 protocol may be required to achieve maximum efficiency over the LMSC channel, these changes are minor and easily implemented.

Despite the desire for TCP compatibility, in some cases the security concerns associated with splitting may preclude this feature from use in a future EHF LMSC system. However, there are advantages to outright replacement of TCP with a more suitable protocol such as SCPS-TP or STP. The first advantage is of course the ability to use end-to-end encryption techniques that encrypt the TCP header. Another benefit is the ability to address specific channel anomalies by optimizing the replacement protocol's behavior during these events. An example of this is the "persist mode" used in SCPS-TP to deal with channel outages. Finally, although more work would be necessary to develop applications that utilize the replacement protocols, these applications would undoubtedly be more efficient than commercial applications written for TCP and used in an environment for which TCP was not optimized. One relevant example of this approach is the Army's Force Battle Command Brigade and Below (FBCB2) system which is being designed to run over UDP due to the limitations of TCP in a wireless environment.

4.3 Routing Protocol Technologies

As mentioned earlier, routing protocols make use of tables that contain connectivity information for the various network nodes. These tables are updated periodically with connectivity information and traffic statistics in order to ensure efficient traffic routing. This practice introduces a measure of robustness into the network by providing a mechanism for routing around nodes that are not working properly or offline for some reason. Clearly, this kind of robustness is desirable in an EHF LMSC system where terminals will constantly be moving in and out of shadow regions that effectively prevent satellite-to-terminal communications. However, conventional routing algorithms, such as those contained in IP, are not appropriate for a LMSC system because they have been optimized for fixed networks where the nodes do not change location and the connectivity is planned a priori by a system administrator. On the other hand, there have been a number of recent proposals for routing algorithms aimed at providing support for mobility. For example, IP version 6 supports a process known as neighbor discovery [31] whereby nodes on the same link discover each other's presence, find routers, and maintain connectivity information about the paths to active neighbors, thus facilitating a certain degree of node mobility. Another, more comprehensive example is Mobile IP [32], a protocol that allows a mobile node to roam, changing its point of attachment to a fixed network (i.e., the Internet) while continuing to be identified by its home address. In its most basic form, Mobile IP allows transparent interoperation between a mobile node and its correspondents

\(^2\)IPsec is short for IP security, a set of protocols developed by the IETF to support the secure exchange of data at the IP layer. Standard IPsec represents a case where WISE will not be able to split a connection since the TCP header information is encrypted.
by forcing all traffic destined for a mobile node to be routed through its home (i.e., permanent) address.

While Mobile IP and certain related technologies are certainly of interest, perhaps the most applicable work currently being done is in the area of mobile ad-hoc network (MANET) protocols [33]. Ad-hoc networks are autonomous systems in which the nodes of the network are mobile and free to move about arbitrarily. The network may operate in isolation or in a “stub” configuration whereby the system interfaces with a fixed network via gateways. Assuming that the network nodes communicate with one another in a wireless fashion, the connectivity of the ad-hoc network will depend on such factors as relative location, antenna coverage patterns, co-channel interference, and transmission power levels. Moreover, as these factors change over time, the networking protocols “adapt” to facilitate changes in the network connectivity. While similar to Mobile IP, the MANET system concept is different in the following fundamental way. Mobile IP supports roaming within a fixed network, while MANET protocols are applicable when the whole network is mobile and the topology is quite dynamic. The mobile network concept supported by MANET protocols is the one most applicable to the proposed EHF LMSC system because, as discussed in Appendix A, it is believed that the system will be used to provide communications support to a large number of mobile platforms whose interconnectivities will change frequently.

Within the context of a multiuser LMSC system, MANET protocols could be exploited to assist disadvantaged (i.e., obstructed) terminals by simply routing traffic to these users via an alternate path comprised of at least one unobstructed terminal. Of course, this approach assumes that at least a small fraction of the LMSC system users will have an unobstructed view of the satellite at any given time. The approach further assumes that the disadvantaged terminal, while blocked from communicating with the satellite, is still able to communicate with at least one unobstructed terminal. This secondary communications path could be provided by augmenting the LMSC system terminal with the capability to transmit and receive on alternate frequencies. Although MANET technologies are still relatively new, the huge increase in commercial demand for mobile wireless networking services is expected to drive a rapid development in protocols and applications. This development within the commercial sector is desirable since it could perhaps be exploited in some ways by a military LMSC system.

The application of MANET technology to a military EHF LMSC system is consistent with a recently proposed view of the proper role of satellite communications in a military communications infrastructure. In [34], the concept of using a portable satellite communications terminal as a shared resource to support lower echelon operations was developed. This concept is based on the fact that a group of users, referred to as a cluster, will wish to communicate primarily amongst themselves in order to coordinate their actions and achieve their common objectives. This local communications was noted to be best supported by LOS radios, which provide adequate range communications capabilities in relatively small, lightweight packages. However, these clusters will also wish to communicate with other clusters (e.g., their sustaining base) that may be beyond LOS distance and hence out of range of their local communications gear. In these cases, satellite communications is more appropriate. The idea is that each cluster has a satellite communications terminal that serves as a cluster head or concentrator node to facilitate communications between clusters. For a user in one cluster to communicate with a user in another distant cluster, links must be established not only between the cluster satellite terminals but also between the users and their respective satellite terminals. Extending this vision to a mobile environment in light of the MANET concepts discussed above implies equipping clusters with multiple satellite terminals, thus increasing
the probability that at least one unobstructed communications path exists between geographically separated clusters. This approach is analogous to using spatial diversity (e.g., multiple antennas) in commercial cellular systems to overcome the effects of multipath fading. Figure 16 illustrates the concept.

4.4 Conclusions and Recommendations

At the transport layer the fundamental issue to be addressed is the relatively poor performance of TCP over satellite links. The approaches for dealing with this problem fall into two main categories: those that sacrifice TCP compatibility through modification or replacement of TCP and those that preserve TCP compatibility through physical and link layer enhancements or splitting via gateways. At this point, it is still not clear which approach makes the most sense with regards to an EHF LMSC system. On one hand, TCP compatibility would seem to be a highly desirable feature in any communications system due to the proliferation of this protocol in commercial and
military networks. On the other hand, there are still a number of concerns related to security and implementational complexity associated with TCP splitting via gateways, the most promising approach to achieving TCP compatibility in an EHF LMSC system. Moreover, at least one military application, FBCB2, is being designed to run over UDP, not TCP. Given these facts, the recommended near term approach for an EHF LMSC demonstration is to sacrifice TCP compatibility for improved performance by designing a custom stand-alone application that will run over UDP. This approach makes the most sense because it achieves the goal of demonstrating a basic EHF LMSC capability without precluding any of the above approaches to dealing with the TCP performance/compatibility tradeoff. For example, many of the lessons learned as a result of experimentation with the stand-alone application could be applied to the development of a protocol optimized for EHF LMSC communications. This protocol could be used to replace TCP in an end-to-end fashion as described in Section 4.2.1. Another option would be to use the protocol for TCP enhancement without splitting as described in Section 4.2.2. Finally, the protocol could serve as an alternate satellite-segment protocol in a TCP splitting scheme, if that approach was deemed viable.

With respect to the network layer, there seems to be a great deal of potential in work related to MANET protocol development. However, this technology is not mature enough to be exploited in the near term. Nonetheless, MANET research should continue to be monitored in hopes that it may someday be applied to improve the performance of a future EHF LMSC system.
5. SUMMARY AND CONCLUSIONS

This report has attempted to identify and provide preliminary insight into solving the baseband signal processing issues related to the development of an EHF LMSC system. System aspects ranging from the physical layer to the transport layer have been considered. The fundamental conclusion of this investigation is that a successful long-term implementation will require development at most of the layers of the protocol stack. Moreover, the algorithms and protocols should be capable of adapting to the wide range of channel conditions that may potentially be experienced over time. However, a more feasible near term approach to demonstrating OTM capability at EHF is to develop a stand-alone voice and data application that runs over UDP. By focusing near term efforts on the application layer, many of the tedious aspects associated with the implementation of error control at lower layers can be avoided. Moreover, the use of UDP represents a convenient means for temporarily bypassing TCP performance issues. Below is a brief summary of conclusions from each of the report chapters.

- **Propagation Modeling:** Due to their short wavelengths, EHF signals tend to be scattered by objects in the propagation path, leading to severe signal attenuations (i.e., 20-30 dB below the mean signal level). Signal shadowing is the term used to describe these large scale variations in the received signal strength, which represent the dominant propagation effect for the EHF LMSC channel. The duration of shadow events or blockage intervals depend on a number of factors, including the size of the obstruction and the vehicle speed. However, according to the ACTS propagation experiments, average blockage intervals due to shadowing are on the order of seconds. In [3], a Total Shadowing Model was proposed to describe the fading behavior in a LMSC system. According to the model, the fading behavior of the channel consists of two dominant modes or states. In the unshadowed state (i.e., the "good" channel state) the channel is characterized by the presence of a LOS component, which implies high received power and Ricean fading, while in the shadowed state (i.e., the "bad" channel state) the channel is characterized by the absence of a LOS component, which implies low received power and Rayleigh fading. The time-share parameter, \( X \), is a long-term average that describes the fractional amount of time spent in each state. The short-term characteristics of the switching process are accurately described by a two-state Markov model [3]. Also in [3] a simple channel simulator was described. Although proposed originally for use in simulating L-band signals, the simulator is easily extended to model the situation at EHF. In fact, several simulated EHF received signal envelopes are presented in Chapter 2.

- **Error Control and Timing:** Several different forms of error control for both voice and data communications were investigated. The fundamental conclusion of the investigation was that the use of FEC coding for the purpose of correcting errors/erasures caused by signal shadowing is ill-advised. Long interleaving delays and high coding overhead are the primary problems with this approach. On the other hand, by relying on ARQ to address burst errors due to shadowing, the overall latency introduced by the error control is closely matched to the shadow duration, as opposed to pure FEC coding approaches where interleaving introduces delays on the order of 10 times the outage period. For the case of voice transmission where some bit errors are tolerable, no FEC coding is necessary to correct the random bit errors due to noise, etc., in the channel good state. However, for data communications, the tolerable random bit error rate is significantly lower compared to voice. Hence, a modest amount of
FEC coding, for the purpose of addressing the presence of random bit errors in the channel
good state, is necessary in addition to the ARQ.

Another important consideration is timing acquisition and tracking. While timing acquisition
in a shadow region is virtually impossible, time tracking is not. The most important feature
of a timing system for an EHF LMSC terminal is that the tracking algorithm be adaptive in
the sense that its behavior is altered depending on whether or not the terminal is operating
in a shadow region. For example, timing synchronization is usually maintained via early/late
gate circuits that compensate for timing errors by processing a synchronization signal. Un-
fortunately, in a shadow region this synchronization signal is not present and the algorithm
designed to help maintain timing may actually introduce errors in these situations. Therefore,
a mechanism whereby the early/late algorithm is either suspended or altered in some way
during blockage periods is desirable. Moreover, by compensating for different timing errors in
different ways, tracking capability in shadow regions may be improved. For example, timing
errors introduced by oscillator drift, satellite ephemeris, and vehicle motion are all somewhat
predictable and may be corrected in a deterministic fashion, even during blockage intervals.
Consequently, a timing system that continuously corrects predictable timing errors in a deter-
ministic way and compensates for random errors via a traditional early/late algorithm when
channel conditions permit seems like the best approach.

- Networking Protocols: The fundamental problem at the transport layer is that the most
  popular transport layer protocol, TCP, is not well suited for use over satellite links. On the
  other hand, given the installed base of TCP and the military's desire to exploit commercial
development as much as possible, TCP compatibility is highly desirable. Consequently, a
number of schemes for providing reliable end-to-end transport layer services over a satellite
link while simultaneously maintaining TCP compatibility have been proposed. Another
approach to providing efficient transport layer services over a satellite link is to abandon TCP
compatibility concerns and use a protocol optimized for the space environment. These proto-
cols have the advantage that they work well over satellite links. However, their use requires
modifications to the protocol stacks of every node in the network.

Recent innovations in routing protocol technologies have enormous potential in a future EHF
LMSC system. Specifically, MANET protocols, which adapt to support dynamic network
topologies, are of interest. Assuming that a multiuser EHF LMSC system is comprised of
terminals capable of communicating with one another via LOS radio, MANET protocols could
be exploited to assist a disadvantaged (i.e., obstructed) terminal that wishes to communicate
over the EHF LMSC system by simply routing traffic to this user through an alternate path
comprised of at least one unobstructed terminal. Of course, this approach assumes that
at least a small fraction of the LMSC system users will have an unobstructed view of the
satellite at any given time. The approach further assumes that the disadvantaged terminal,
while blocked from communicating with the satellite, is still able to communicate with at least
one unobstructed terminal. This system concept is loosely analogous to the spatial diversity
practices used in cellular systems to mitigate the effects of multipath fading.
APPENDIX A
CONCEPT OF OPERATIONS

In this appendix, the term concept of operations is used loosely to include: the intended users of the system, the type and amount of information that will flow over the system, and the environments in which the system will be expected to operate. Simply stated, the primary function of an EHF LMSC system should be to provide mobile connectivity among the various command posts (CPs), tactical operations centers (TOCs), and subordinate units, all of which may be widely dispersed over the battlefield.

Presently, the Army's primary fighting unit is the division. The hard core communications requirements of a division are those associated with battle command, an ongoing cycle of assigning missions and tasks, acquiring feedback, evaluating information, and making decisions. Battle command at the division level is conducted by a division commander and supporting staff, referred to as a command group [35]. The command group operates from a command post (CP), a physical location that houses the facilities and system interfaces that allow commanders to direct the battle. Another entity associated with battle command is the tactical operations center (TOC). Functionally, TOCs are quite similar to CPs. However, the primary difference between a TOC and a CP is that whereas CPs reside at the division level, TOCs are typically associated with a brigade (Bde) or battalion (Bn). Consequently, there may be as many as 25 to 30 TOCs in a division compared to 2 or 3 CPs.

To a command group, the primary benefit of EHF satellite communications, aside from the protection of the link, is the beyond line of sight (BLOS) connectivity that it provides. This connectivity is crucial to have among the various CPs, TOCs and their subordinate units, which may be highly dispersed over the battlefield. While CPs and TOCs typically operate as static entities, it is necessary for them to move as the battle evolves. Moreover, mobility may be viewed as a countermeasure to the increasing range and precision of modern weaponry and surveillance systems, implying that CPs and TOCs will move often as a simple form of survivability.

Another area in which the Army might benefit from an EHF LMSC system is cavalry operations [36]. The primary role of cavalry is to perform reconnaissance and provide security in close operations. In addition, cavalry may also counter enemy cavalry, cover a retreat, or pursue a retreating enemy. In order to conduct reconnaissance, cavalry units are often spread relatively thin with various elements separated by large distances. In addition, reconnaissance missions are often conducted at the outskirts of a division deployment, far from the CP/TOC. For these reasons, satellite communications would seem to be a logical choice to provide connectivity among the various cavalry elements as well as between cavalry and division CPs/TOCs. One observation made in [37] is that future cavalry reconnaissance missions will likely make use of robotic vehicles. Given the dangerous nature of cavalry reconnaissance operations, remote controlled robotic vehicles may provide a safer means of achieving these objectives. Moreover, with no human lives at stake, robotic vehicles may be employed in even riskier situations, such as operations deep in enemy territory. In addition to supporting the transmission of sensor information collected by robotic vehicles, an EHF LMSC system also seems ideally suited for use in controlling the vehicle from a remote location.

When considering the communications capability that should be supported by an EHF LMSC system, a couple of observations should be made. First, human limitations in processing certain forms of information while OTM make it unnecessary to deliver the same communications support to a mobile platform as a static one. For example, experiments have shown that soldiers get
physically ill after repeated attempts at consuming visual information displayed on a computer monitor while OTM [38]. Second, many of the battlefield functional area (BFA) systems used in CPs and TOCs make use of information stored in local databases. Moreover, this information needs to be updated periodically to ensure that decisions are being made based on the most accurate information available. Together these observations suggest that there is little need to support communications rich with graphics and multimedia while OTM. Rather, attention should be focused on supporting database updates such as situational awareness (e.g., maps, weather, friendly and enemy position information), weaponry, and other asset status information, etc. With this information in hand, the decision-making cycle can be reduced substantially after the CP/TOC stops and the command group has an opportunity to evaluate the situation. In addition to database updates, the ability to receive limited intelligence information (e.g., troop movements, incoming fire) while OTM is highly desirable, as is the ability to issue orders to subordinates and receive force report-back information. The kind of information one would expect to be generated from a cavalry reconnaissance mission includes text-based reports, voice, imagery, and perhaps video. Collectively, these communications requirements are relatively modest and could likely be supported with links that range in capacity from 10s of Kbps to 100s of Kbps.
APPENDIX B
EHF VS. UHF

In this section an attempt will be made to establish the relationship between UHF and EHF satellite communications by examining their relative strengths and weaknesses and exploring how they may typically be used in tactical situations.\(^1\) The military's satellite communications architecture is often viewed as consisting of a "hard core" surrounded by a "soft shell." Within the hard core are the most important communications requirements, including command and control, situational awareness, force report-back, selected intelligence, and selected strategic messaging. Due to the importance of this information, hard core communications needs must be serviced by highly robust systems with a large degree of survivability. On the other hand, soft shell requirements include items such as logistics, administration, and support for contingency operations. Because they are less critical, these needs are best served by more general purpose systems that while not as robust or survivable are likely to be cheap and widely available.

Due to the large amounts of bandwidth allocated in the EHF region of the electromagnetic spectrum, the MILSTAR satellite system provides a great deal of antijam (AJ) capability to system users. Moreover, other features such as antiscintillation make the system highly survivable. For these reasons, MILSTAR is typically seen as the military's primary hard core communications service provider. On the other hand, soft shell requirements could perhaps be served by two separate options: military UHF satellite systems and commercial satellite systems. While these systems are not as robust as MILSTAR, they do possess many desirable attributes. For example, the longer wavelengths at UHF imply that these signals are more likely to penetrate objects such as foliage and buildings compared to EHF signals. One approach to the EHF vs. UHF dichotomy is to develop system concepts that attempt to exploit the strengths of both systems while simultaneously minimizing their weaknesses.

One idea that exploits the potential in both EHF and UHF systems is to equip certain high priority mobile platforms with both EHF and UHF capability. Due to their superior propagation characteristics, the UHF systems could be used to support interactive communications such as voice calls, while EHF systems could be used to support noninteractive communications such as file transfers. This partitioning of interactive communications via UHF and noninteractive communications via EHF seems logical for several reasons. First, the bandwidth associated with any single channel, whether it be UHF or EHF, may not be adequate to fulfill the platform's communications needs. By using two channels, these needs may be more readily satisfied. Second, the use of two distinct systems introduces a measure of redundancy that helps guard against disasters related to equipment failure.

Building further on the idea that a mix of UHF and EHF may be the most appropriate means for providing satellite communications OTM, there are numerous additional options for partitioning the mix. For example, given the scarcity of UHF satellite channels, it seems likely that UHF could not be used to provide voice channels to all those in need. Therefore, a more appropriate use of this resource may be to provide a service that can be shared by a large user population. One service that falls into this category is paging. Although there are numerous commercial paging services available, for coverage and accessibility reasons it makes sense for the military to incorporate this capability into their satellite communications systems. Given the low data rate of paging signals,

\(^1\)SHF satellite communications are largely ignored here because these systems have traditionally been used to provide connectivity to the higher, more strategically oriented echelons.
a relatively small number of shared UHF channels should be sufficient to support a large user population. This idea has been proposed numerous times, and even demonstrated via the UHF payload on MILSTAR I [39]. As an adjunct to an EHF satellite system, the benefit of UHF paging capability is that in the event system users are blocked from communicating by obstructions in the propagation path they could be made aware of someone else's desire to communicate with them and take the steps necessary to establish the communications link (e.g., move out from under foliage, etc.). Another idea for efficient utilization of UHF channels is to use them for protocol signaling. For example, many data communications protocols rely on packet acknowledgments to ensure reliable delivery. These acknowledgments are relatively short and could perhaps be sent over a small number of UHF channels shared by multiple EHF system users. The advantage of such an arrangement would be the improvement in overall system performance achieved by virtue of the fact that protocol signaling was being conducted over a channel less susceptible to outages caused by obstructions in the propagation path. Finally, it would also be desirable to perform time tracking based on a UHF signal. This way, in the event of an EHF channel outage, time tracking could be maintained so that reacquisition would be faster when the channel became available. Moreover, timing information could be obtained by monitoring a single UHF downlink channel, a task that could be performed simultaneously by an unlimited number of users. It should be noted that each of the last two ideas (i.e., UHF for protocol signaling and time tracking) would require that both UHF and EHF payloads be located on the same satellite and, in the case of UHF protocol signaling, that these payloads be capable of exchanging data with one another. It should also be noted that, because EHF systems will typically be used to satisfy hard core communications requirements, no system concept that allows for UHF to provide certain capabilities should be without mechanisms for providing these same capabilities via EHF. As an example, the use of UHF channels for protocol signaling should yield improved performance. However, this practice also introduces a potential vulnerability in the sense that UHF channels are susceptible to jamming, and the loss of this channel would essentially deny communications services to the users. Therefore, the ability to duplicate this functionality at EHF, albeit with possibly degraded system performance, is critical.
APPENDIX C
GLOSSARY

ACK: ACKnowledgment
ACTS: Advanced Communications Technology Satellite
AJ: AntiJam
ARQ: Automatic Repeat reQuest
Bde: Brigade
BER: Bit Error Rate
BFA: Battlefield Functional Area
BLOS: Beyond Line Of Sight
Bn: Battalion
CP: Command Post
CRC: Cyclic Redundancy Check
dB: deciBel
EHF: Extremely High Frequency (30-300 GHz)
FBCB2: Force Battle Command Brigade and Below
FEC: Forward Error Correction
FTP: File Transfer Protocol
GHz: GigaHertz (10^9 Hz)
HARQ: Hybrid Automatic Repeat reQuest
HTTP: HyperText Transfer Protocol
Hz: Hertz
IETF: Internet Engineering Task Force
IP: Internet Protocol
IPsec: IP security
ISO: International Standards Organization
Kbps: Kilobits per second
km: kilometers
L4: Lincoln Laboratory Link Layer
LDR: Low Data Rate
LMSC: Land Mobile Satellite Communications
LOS: Line Of Sight
MANET: Mobile Ad hoc NETwork
Mb: Megabits
MDR: Medium Data Rate
MED: Modified Exponential Decay
MELP: Mixed Excitation Linear Prediction
MHZ: MegaHertz (10^6 Hz)
MILSTAR: MILitary Strategic and TActical Relay
MIT: Massachusetts Institute of Technology
ms: milliseconds
MSE: Mobile Subscriber Equipment
NAK: Negative AcKNowledgment
NASA: National Aeronautics and Space Administration
NTC: National Training Center
OSI: Open Systems Interconnection
OTM: On The Move
PSD: Power Spectral Density
QoS: Quality of Service
RS: Reed-Solomon
SCAMP: Several Channel Anti-jam Portable

SCPS: Space Communications Protocol Standards

SCPS-TP: SCPS-Transport Protocol

SHF: Super High Frequency (3-30 GHz)

SMTP: Simple Mail Transfer Protocol

SSL: Secure Socket Layer

STP: Satellite Transport Protocol

TCP: Transmission Control Protocol

TOC: Tactical Operations Center

TPN: Tactical Packet Network

UDP: User Datagram Protocol

UHF: Ultra High Frequency (300-3,000 MHz)

WISE: Wireless IP Suite Enhancer
REFERENCES


16. M. A. Gouker, MIT Lincoln Laboratory. private communication.


20. R. J. Greene, MIT Lincoln Laboratory. *private communication.*


25. K. J. Hetling, MIT Lincoln Laboratory. *private communication.*


38. T. Newsome, Command and Control Directorate, CECOM. *private communication.*

EHF Satellite Communications on the Move: Baseband Considerations

Jeffrey B. Schodorf

Lincoln Laboratory, MIT
244 Wood Street
Lexington, MA 02420-9108

Commander
US Army CECOM
AMSEL-RD-ST-SS-SP
Fort Monmouth, NJ 07703

None

Approved for public release; distribution is unlimited.

In this report, baseband signal processing and networking protocol issues associated with the development of an EHF land mobile satellite communications system are investigated. Ignoring the very important problems of antenna pointing and tracking, the primary obstacle to be overcome by the system is signal blockage. At EHF, objects in the propagation path are virtually opaque and cast a dark "shadow" over the communications terminal, resulting in signal attenuations on the order of 20-30 dB or more. The duration of these shadow regions or blockage intervals will depend on a number of factors, including object size and vehicle speed. Recommended solutions to this problem include the following: forward error correction coding at the physical layer to mitigate the effects of relatively short blockage intervals (i.e., milliseconds to seconds), and automatic repeat request schemes at the link layer to ensure reliability over longer shadow regions (i.e., seconds to 10s of seconds). In addition, dynamic routing protocols can be used at the network layer to provide connectivity to blocked terminals via alternate terrestrial communications paths that include at least one unobstructed terminal. Finally, TCP splitting and protocol conversion offer a great deal of promise for achieving reliable end-to-end transport layer services without sacrificing TCP compatibility.