NASA/TM—2018-219877

Metabrain for Embedded Cognition (MBEC)

Rigoberto Roche, Janette C. Briones, and Brittany N. Kowaleski
Glenn Research Center, Cleveland, Ohio

August 2018

NASA STI Program . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) Program plays a key part in helping NASA maintain this important role.

The NASA STI Program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI Program provides access to the NASA Technical Report Server—Registered (NTRS Reg) and NASA Technical Report Server—Public (NTRS) thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counter-part of peer-reviewed formal professional papers, but has less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., "quick-release" reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

For more information about the NASA STI program, see the following:

- Access the NASA STI program home page at http://www.sti.nasa.gov

- E-mail your question to help@sti.nasa.gov

- Fax your question to the NASA STI Information Desk at 757-864-6500

- Telephone the NASA STI Information Desk at 757-864-9658

- Write to:
  NASA STI Program
  Mail Stop 148
  NASA Langley Research Center
  Hampton, VA 23681-2199

NASA/TM—2018-219877

Metabrain for Embedded Cognition (MBEC)

Rigoberto Roche, Janette C. Briones, and Brittany N. Kowaleski
Glenn Research Center, Cleveland, Ohio

National Aeronautics and
Space Administration

Glenn Research Center
Cleveland, Ohio 44135

August 2018

Level of Review: This material has been technically reviewed by technical management.

Available from

NASA STI Program
Mail Stop 148
NASA Langley Research Center
Hampton, VA 23681-2199

National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
703-605-6000

This report is available in electronic form at http://www.sti.nasa.gov/ and http://ntrs.nasa.gov/

Metabrain for Embedded Cognition (MBEC)

Rigoberto Roche, Janette C. Briones, and Brittany N. Kowalewski*
National Aeronautics and Space Administration
Glenn Research Center
Cleveland, Ohio 44135

Summary

This study presents the application of hidden Markov models (HMMs) to determine specialized features without expert input. Specifically, the application of such a method for classification of high multipath fading is targeted for demonstrating the feasibility of such an approach. This is the first step in the development of a metabrain for embedded cognition (MBEC) suite that can be used to apply machine learning to various communication systems at NASA Glenn Research Center. The project explores the concept of fading and how it affects communication systems in a negative way. Currently, supervised learning methods are used to study the effects of fading on space links. However, such models rely on expert features to make predictions as to the state of a link and whether fading is present. This project offers the possibility of having the HMM learn what characteristics are important and make predictions based on those characteristics. This project explores HMMs, their theory and applications to various problems, as well as the underlying equations and assumptions. A preliminary result is presented and recommendations are made as to the use of such an approach for communication systems.

Nomenclature

EsNo	energy-to-noise ratio
ISS	International Space Station
HMM	hidden Markov model
MBEC	metabrain for embedded cognition
NEN	Near Earth Network
RF	radiofrequency
SNR	signal-to-noise ratio
TDRS	Tracking Data Relay Satellite
TSC	Telescience Support Center

Symbols

a	transition probability
A	transition probability matrix
b	emission probability
B	emission probability matrix
N	number of hidden states
n	observation per state
o	single observation
O	observation sequence matrix
P	probability matrix

---

*Summer Intern in Lewis' Educational and Research Collaborative Internship Project (LeRCIP), a graduate student at Towson University.

q	state
Q	set of hidden states
t	instance of a sequence
T	set of sequences
v	instance of a Viterbi probable state (most likely)
V	Viterbi state sequence matrix
l	possible models of (A,B)
a	previous state probability

Subscripts

F	end
i	initial state
j	final state
n	finite iteration dimension
0	start

1.0 Introduction

The world-class laboratory in space known as the International Space Station (ISS), shown in Figure 1, offers a unique science platform for which researchers from around Earth can experiment with the unknown. The ISS provides the conditions that facilitate key experiments for the advancement humanity's knowledge. This includes biological, psychological, material science, and communication technology amongst other fields. Researchers and scientists can create, operate, and continuously monitor and update experiments and equipment from a remote location. The planning and results of experiments underway are crucially important, and the data to produce conclusive findings, must be collected and carefully relayed back to the associated researcher on Earth.

Since the ISS is involved with many ground facilities spread around Earth, from the United States, Russia, Canada, Europe, and Japan, an unclear or otherwise impossible communication reception is unacceptable for researchers and scientists deeply invested in ISS operations. Glenn hosts one of the Telescience Support Centers (TSCs) and is equipped to conduct scientific research onboard the station. For this reason, payloads and equipment belonging to Glenn are currently occupying the ISS, and up-to-date communication is especially important to oversee the data, voice, and video on payload operations, making the best use of the astronauts' time when interacting with the experiment on board.

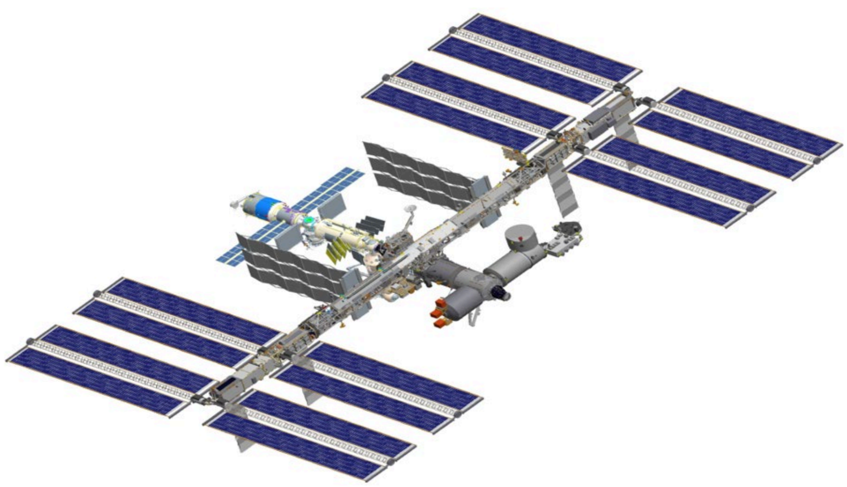

Figure 1.—International Space Station.

To prevent the waste of time and power in attempting to transmit data over a fading signal and determine optimal times for clear communication, a hidden Markov model (HMM) is used in conjunction with Python™ (PSF) programming as a classifier to determine for future reference when to stop transmission based on a less than optimal connection. A semisupervised version of reinforcement machine learning allows the application of previously viewed signal strength to consider and prevent further loss of data over future signal communications in space (Ref. 1).

The paper is organized as follows: Section 2.0 provides background information, Section 3.0 describes the proposed HMM algorithm, Section 4.0 provides some preliminary results, and Section 0 provides concluding remarks.

2.0 Background Information

The Cognitive Communication System Project is developing cognitive communication technologies to increase mission science return and improve resource efficiencies. The aim is to increase the autonomy, efficiency, and reliability of the Space Communications and Navigation for the next generation architecture by merging space communications and machine learning. The project focuses on three key areas:

(1) Cognitive link capability—These are point-to-point links aiming to achieve configurations that maximize throughput and avoid spectrum interference, sense incoming signals, and adapt the radio for the different space environments. Through this process, the cognitive engine analyzes all available link parameters and adjusts them so the most data can be brought down for any given pass, regardless of weather conditions, interference, or other factors.
(2) Decentralized networks—Here the focus is on how to move customer data between different assets in space down to the ground, in what is called drop data anywhere. In this context, data will be chunked and parsed to produce serializable packets that can be reconstructed from different nodes on the ground after segmented transmission to links of opportunity. This allows for the constellation to work as a unified system, without the need to preset schedules and timing, but rather be a data-driven system like cellular communications are on the ground today.
(3) Automatic scheduling of space networks, the Near Earth Network (NEN), and commercial services—This scheduling is achieved through what is referred to as "user initiative service," which means that the spacecraft has a request for service and the network knows how to satisfy that request. Current efforts focus on developing cognitive engines capable of making decisions from the network level all the way to the link level. The team is exploring the use of cutting-edge technology for implementing artificial intelligence solutions in size, weight, and power-contained assets. This focus of the paper will be on optimizing the cognitive links.

NASA's future mission plans continue to evolve by adding new technologies and capabilities to space communication systems ever accommodating for anticipated mission needs and objectives. These new technologies will include artificial intelligent and machine learning algorithms in an effort to improve automation and efficiency in these systems. This report investigates the advantages of using a specific model for maximizing data integrity and overall performance on a space-to-ground link. Previous research has addressed the performance of various machine learning and optimization techniques for decision making of link properties. In this paper, recent studies with a HMM are presented as potential candidates for deployment as future cognitive engines to control resources onboard NASA's satellite communication systems.

3.0 Fading Problem

Current space links to ground assets are achieved in two ways. The first way is direct to ground. In this case, the satellite sends data links to a ground station that tracks its position while there is a transfer of information. The second way is via a Tracking Data Relay Satellite (TDRS). This is a telemetry and data relay satellite parked in geosynchronous orbit. This system can relay signals to ground stations that are beyond the line of sight of the data originating asset. This way, satellites can communicate with their respective ground stations without having to be in a constant line of sight to their targets.

Fading is a phenomenon that can occur when radio waves are blocked or reflected from the source by an interferer. This could be a physical object that is reflective to radiofrequency (RF) waves in the frequency of transmission or another signal close to the center frequency of the transmitter that is jamming the channel by adding noise to the modulation. Fading is characterized by rapid movement of the signal power per unit time. The vastly changing distance at a given moment from a ground station to the ISS, environmental factors including weather and cloud coverage, and physical blockages of the antenna by the ISS structure itself as well as the surrounding variable circumstances apparent in space can contribute to fading. The encoded data may be impossible to understand or decode, depending on the distortion of the incoming signal.

3.1 Signal-to-Noise Ratio (SNR)

In space, messages and data communication are sent via radio waves, and for any radio receiver, signal variations can be measured relative to noise performance. To measure the sensitivity performance of a receiver and its ability to separate the signal from the unwanted noise, a SNR is used. The greater the SNR is, the greater the difference between the signal and the interference, meaning that the receiver has a greater sensitivity performance. Despite the sensitivity of the receiving satellite, several other factors, including distance, weather, and obstacles, will still cause fading. A fading condition will cause a radio signal to experience intervals of the following:

a. Rapid amplitude fluctuation and/or oscillation
b. Very little power
c. Static and/or otherwise rendered unable to be used for data decoding

Figure 2 is an SNR graph showing a visual example of an interval that is classified as being in fading. During this interval, the ISS would want to stop transmission of a signal, as decoding of data would be negatively affected during that time.

3.2 Distance, Weather, Spatial Factors, and Structural Blockages

Fading is an extreme burden in space communication; however, factors that cause fading are often unable to be controlled or prevented. Considering satellite signals in general, the distance between the ISS and Earth is an obvious obstacle. While transmitters from the ISS are able to send a powerful signal and ground stations work effectively as very sensitive pieces of equipment, the approximate 254-mile distance to the ISS will be host to numerous hazards to a clear signal along the way. Radio waves at certain frequencies cannot pass through every substance they encounter. Earth’s atmosphere contains numerous substances, some of which pose a threat to a traveling wave. Rough weather on Earth, such as thick

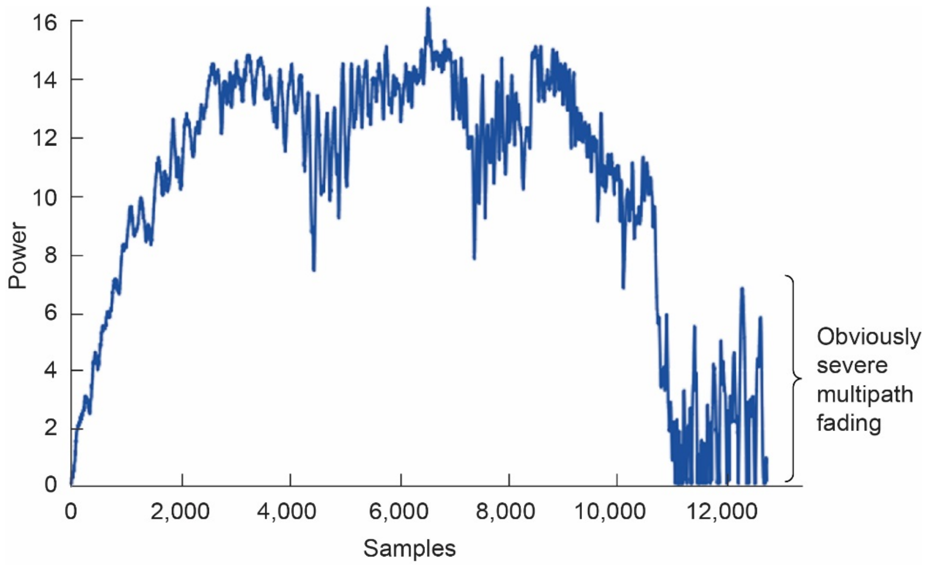

Figure 2.—Power (signal-to-noise ratio) versus samples direct to ground link event experiencing periods of fading.

clouds, heavy rain, and flaky snow can add noise to the signal. In the event that the radio wave encounters a physical object, the signal may distort altogether. There are also objects in space that can distort or reflect a signal. In transmission, if an object, such as the ISS's own solar panels as shown in Figure 1, gets in the way, the radio wave may be blocked or spread out into several small waves instead of the aimed narrow beam that was expected. In either case, fading will prevent the reception of the entirety of the data that were encoded in the signal, so again, the ISS would want to stop transmission during that time.

4.0 Hidden Markov Model (HMM) Overview

To classify intervals of fading and no fading during a time-continuous incoming transmission, an HMM can be adapted to fit the signal. An HMM is a sequence model, or sequence classifier, which functions by assigning a label or class to each value in a sequence. As a descendent of a Markov model, an HMM is a probabilistic model that maps each assigned label to an observation sequence based on the most probable label sequence. Unlike Markov models, an HMM is useful when a sequence is not directly observable, but has an associated emission sequence that is observable (Ref. 2).

4.1 Properties and Compounds

The foundation of an HMM coincides with a general Markov model, namely a Markov chain. A Markov chain is a probabilistic graphical model for which a sequence traveling from a start to an end state is assigned a number of probabilities. As a Markovian process, transition probabilities among states abide by the Markov assumption; the probability of any given next state in the process depends only on the previous state in the sequence.

From an initialized start state, the sequence now has various transition probabilities to the next state that will be observed. Altogether, the transition probabilities form a matrix, A, with each entry representing a transition probability. If a start state, q_0, and end state, q_F, are not explicitly defined, there is an additional set of probabilities over the states which initialize the probability of starting in any given state. The basic functionalities of a Markov chain include

$\pi = \pi_1, \pi_2, \ldots, \pi_N$ — is an initial probability distribution over the states. π_i is the probability that the Markov chain will start in state i. Some states, j, may have $\pi_j = 0$, meaning that they cannot be initial states. Also, $\sum_{i=1}^{n} \pi_i = 1$.

Markov assumption: $P(q_i | q_1 \ldots q_{i-1}) = P(q_i | q_{i-1})$

$Q = q_1 q_2 \ldots q_N$ — is a set of N states.

$A = a_{01} a_{02} \ldots a_{n1} \ldots a_{nn}$ — is a transition probability matrix for A, with each a_{ij} representing the probability of moving from state i to state j, such that $\sum_{j=1}^{n} a_{ij} = 1 \quad \forall i$.

q_0, q_F — are special start and end (final) states that are not associated with observations.

To apply a Markov chain, all the states are a sequence of events that is directly observed in the real world. An HMM, however, applies by assigning probabilities for a hidden sequence of events that cannot be directly observed. A hidden sequence of states can be thought of as causal states because they emit, or cause, an observable event. The formulas associated with an HMM are a derivation from the Markov chain:

$Q = q_1 q_2 \ldots q_N$ — is a set of N states.

$A = a_{11} a_{12} \ldots a_{n1} \ldots a_{nn}$ — is a transition probability matrix for A, with each a_{ij} representing the probability of moving from state i to state j; such that $\sum_{j=1}^{n} a_{ij} = 1 \quad \forall i$.

$O = o_1 o_2 \ldots o_T$ — is a sequence of T observations, with each one drawn from a vocabulary of $V = v_1, v_2 \ldots v_v$.

$B = b_i(o_t)$ — is a sequence of observations likelihoods, also called emission probabilities, with each expressing the probability of an observation, o_t, being generated from a state i.

q_0, q_F — are special start and end (final) states that are not associated with observations, together with transition probabilities $a_{01} a_{02} \ldots a_{0n}$ out of the start state and $a_{1F} a_{2F} \ldots a_{nF}$ into the end state.

$\pi = \pi_1, \pi_2, \ldots, \pi_N$ — is an initial probability distribution over states. π_i is the probability that the Markov chain will start in state i. Some states, j, may have $\pi_j = 0$, meaning that they cannot be initial states. Also, $\sum_{i=1}^{n} \pi_i = 1$.

The sequence, Q, of states becomes hidden and now has an associated observation sequence, O. As a Markovian process, the Markov assumption is still present in addition to output independence, and this means that the probability of an output observation only depends on the hidden state that produced the observation, not previous states or observations. Functionally, HMM algorithms are used to analyze three different types of problems: likelihood, decoding, and learning.

4.2 Forward Algorithm

The first problem, likelihood, becomes present when the known values are the model with parameters (*A*,*B*) and the observation sequence, *O*, but the desired likelihood of the observation sequence, *O*, given the model with *A* and *B* is unknown. To determine the probability of *O* given the model, $\lambda = (A,B)$, the goal is to calculate the probability of the hidden state sequence that produces output *O*. Because the hidden state sequence is unknown, all possible hidden state sequences given by the observation sequence, *O*, would be summed by their weighted probabilities. With *N* hidden states, there would be N^T possible hidden state sequences, and the total observation likelihood would be too difficult to compute. Thus, the forward algorithm is applied. The forward algorithm is a dynamic programming algorithm, which scores the probability of a possible hidden path producing the observation sequence using a forward trellis. Each score in the trellis assigns a probability of being in a state after the first specific number of observations and given the model, $\lambda = (A,B)$. The formulas for the forward algorithm are explained as

$$\alpha_t(j) = \sum_{i=1}^{N} \alpha_{t-1}(i) a_{ij} b_j(o_t)$$

This can be represented as

$$\alpha_t(j) = P(o_1, o_2, \ldots, o_t, q_t = j \mid \lambda)$$

In the formulas, to find the probability that the *t* numbered state in the sequence is the state *j*, all paths that lead to this current point in time are summed. Note that the previous state probability, α, is initialized with a random float from 0 to 1, since there is no knowledge of the probability of the previous state in the first iteration.

4.3 Viterbi Algorithm

The second problem, decoding, again begins with known values $\lambda = (A,B)$ and observation sequence *O*, but instead the desirable is the most probable sequence of hidden states *Q*. For each possible hidden state sequence that could produce the observation sequence, the goal is to find the one with the highest probability of occurring. The Viterbi Algorithm is another form of dynamic programming that takes advantage of a trellis procedure by recursively going through the observation sequence and taking the most probable hidden state sequence path that would put the pointer at that specific observation. Mathematically, the Viterbi trellis is

$$v_t(j) = \max_{i=1}^{N} v_{t-1}(i) a_{ij} b_j(o_t)$$

This can be expressed as

$$v_t(j) = \max_{q_0, q_1, \ldots, q_{t-1}} P(q_0, q_1, \ldots, q_{t-1}, o_1, o_2 \ldots o_t, q_t = j \mid \lambda)$$

In the formula, the HMM is in state j after taking the most probable hidden state path $q_0,\ldots q_n$ and observing $o_1,\ldots o_t$, and given $\lambda = (A,B)$. Note that v_t is initialized with a random float from 0 to 1, since there is no knowledge of the probability of the previous state in the first iteration.

4.4 Backward Algorithm

The third problem, learning, involves knowing only the set of possible hidden states in the HMM and an observation sequence, O, and the goal is to learn the model parameters, A and B. Essentially, the HMM will be trained based on the given sequence, O, and will make initial estimates for the transition and emission probabilities A and B. To help with the learning problem, the forward algorithm can be applied due to the iterative nature. The forward algorithm is helpful because its goal is to calculate the probability of the hidden state sequence that produces output O. However, the hidden state sequence is not directly observed in an HMM and the model parameters are unknown, so with each forward probability, the probability mass of the different hidden state paths involved must be considered. The backward probability can be used to take into account the remaining observations in O, starting from the next state to the end of the sequence, given the current state. The forward-backward algorithm will need to make initial estimates for transition and emission probabilities, decide the observation sequence that can be generated with the given probabilities, and update to a closer estimate, learning from the deviations to the given sequence O. The backward probabilities are calculated as

$$\beta_t(i) = P(o_{t+1}, o_{t+2}, \ldots o_T \mid q_t = i, \lambda)$$

5.0 Proposed Solution

Recall that when the ISS transmits radio waves, the signal is eventually received by a ground station and is processed as a continuous sequence. The incoming signal transmission can be directly observed, but areas with fading and nonfading conditions cannot be directly determined. By considering the signal as the observation sequence, O, with an unknown hidden state sequence and model parameters, an HMM can be applied to the data. The goal is to use the model as a classifier to decide when a signal is fading or not fading and either stop or continue transmission for a given interval. An example, in graphical view, of a fading interval in a signal can be found in Figure 2; the data shown are from an actual event of a signal that was received in the past. Past data such as this will be used to train the model, hence, it will enter the learning problem of an HMM. Given a sequence that was already experienced in the past, characteristics can be analyzed to classify fading.

For a known signal that has already been received and processed, the HMM algorithm can begin training to recognize where fading had occurred in relation to self-identified characteristics at that point. This is the learning method by which training of the HMM can be achieved, and thus building a model classifier where previous known data is used to make predictions based on probabilities. Because the HMM is first trained, but then expected to process data that it has not seen or processed before, this is considered supervised machine learning (Ref. 3).

The HMM is built by estimating parameters A and B and trained using the forward-backward algorithm with a cost function that iteratively evaluates the model parameters against the training data by gradient descent until the epsilon difference is very small. Since the model is expected to identify intervals of fading on incoming real-time signals, once the supervised portion is complete, the HMM is trained with past data. This process is illustrated in Figure 3.

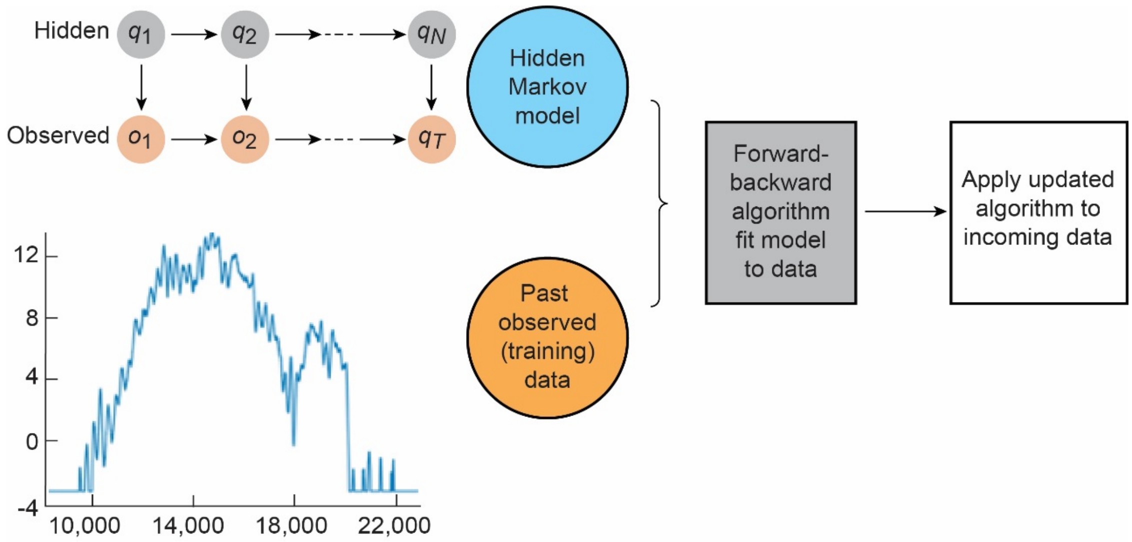

Figure 3.—Training and implementation of hidden Markov model algorithm.

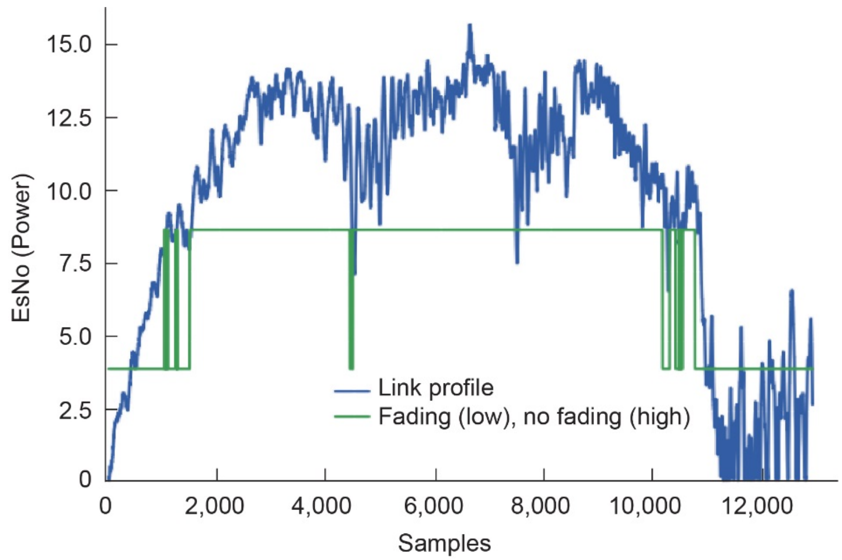

Figure 4.—Preliminary study of hidden Markov model predicting fading over link profile of unseen data. EsNo, energy-to-noise ratio.

6.0 Preliminary Results

The results show the preliminary feasibility study on a link experiencing fading. The model was tested on a profile of unseen data to determine whether it was able to detect the presence of fading on a real-world space downlink. The results from one of these tests are shown graphically in Figure 4. Figure 4 shows the link experiencing fading and no fading. The classification of fading on the link profile was done using unseen data.

The model was evaluated with test data and the classification report illustrated in Figure 5 was created to show its performance after training. The classification report of HMM on test data shows the fading, marked as 0, has 85 percent accuracy whereas no fading, marked as 1, showed 98 percent of accuracy in the classification.

	precision	recall	f1-score	support
0	0.85	0.97	0.91	15889
1	0.98	0.86	0.91	19492
avg / total	0.92	0.91	0.91	35381

Figure 5.—Classification report of hidden Markov model on test data.

```
[[15489   400]
 [ 2733 16759]]
```

Figure 6.—Confusion matrix of hidden Markov model on test data.

The confusion matrix of HMM shown in Figure 6 show the model misclassified 400 samples for class 0 (no fading) and 2,733 samples for class 1 (fading). These 400 samples were labeled as fading but the model misclassified them as no fading and the 2,733 samples were in fact not fading but the model misclassified them as fading.

7.0 Conclusion

This study presented the application of hidden Markov models (HMMs) as a possible method for characterization of fading in a communications link using specialized features without expert input. The feasibility of such an approach was addressed with a preliminary study and the results of the study were presented. The model was able to correctly predict fading 92 percent of the time on the test data. This study serves as the first step in the development of a metabrain for embedded cognition (MBEC) suite. This suite will be able to establish automatic feature detection for unknown data if a target function or target action is indicated by the user. The HHM represents the automatic feature detection portion of such an effort.

There are several advantages to the use of this method. One of the main reasons to use this method as an approach for feature detection is the removal of bias from the human engineered features. This presents the opportunity of obtaining a model that has been trained on data-driven solutions alone, without the restriction of expert-driven, arbitrary scoping that can sometimes limit the knowledge extracted from the data.

There are several disadvantages to the use of such a method. One such disadvantage is the fact that the model can detect features on every sample and come up with large, computationally expensive solutions without any real gain with respect to the expert-drive feature selection. This is a problem that can be easily addressed using semisupervised methods to limit the scope of the model and compare that to the current one. By trying both, an assessment can be made as to which one to use, especially in size, weight, and power-constrained environments.

Overall, this study demonstrated that the HMM approach is a feasible method for characterization of fading on a link and detecting features from data in communication systems. Further research is needed to optimize these solutions and advance the technology readiness level of this approach on NASA Space Communications and Navigation systems.

References

1. Qualitative Reasoning Group: Principles of Operations for Deep Space 1. Qualitative Reasoning Group at Northwestern University online database, 2018. http://www.qrg.northwestern.edu/projects/vss/docs/index.html Accessed Dec. 5, 2017.
2. Jurafsky, Daniel; and Martin, James H.: Speech and Language Processing. Second ed., Ch. 6, Prentice-Hall, Upper Saddle River, NJ, 2009.
3. Sutton, Richard S.; and Barto, Andrew G.: Reinforcement Learning: An Introduction. Second ed., The MIT Press, Cambridge, MA, 2016. Ch. 3 and Ch. 6.