Covert Binary Communications through the Application of Chaos Theory: Three Novel Approaches

Kyle J. BRADBURY
Electrical and Computer Engineering, Tufts University
Medford, MA 02155, USA

and

Joseph P. NOONAN
Electrical and Computer Engineering, Tufts University
Medford, MA 02155, USA

ABSTRACT

Today, most covert communications systems use a spread-spectrum approach to ensure that transmissions remain clandestine. This paper expands beyond traditional spread-spectrum schemes and into chaos theory in communications by presenting a novel design for a covert noncoherent binary communication system that uses chaotic signals. Three techniques are developed, with varying performance. Each system uses two chaotic signals with antipodal attractors as the information carriers.

Although the two chaotic signals used are continuously generated from random starting values without containing repetitious patterns, the receiver requires neither those initial values nor does it require synchronization with the transmitter. The chaotic signals used are both spread-spectrum in the frequency domain and undetectable using matched-filter receivers, thereby achieving a level of covertness. The signal-to-noise ratio performance is presented through simulated receiver operating characteristic (ROC) curves for a comparison to binary phase shift keying.

This system provides a binary communication scheme which is not detectable by standard matched filtering techniques and has noise-like spectra, requiring a new receiver configuration and yielding security.

Keywords: Chaos Communications, Digital Communications, Covert, Receiver Operating Characteristics

BACKGROUND

In the world of covert communications, there are three primary measures of a successful covert communications system: the probability of intercept, the probability of detection, and resistance to jamming. As technology progresses, there is always a need to remain the "most covert" by ensuring that communications technology is on the cutting-edge.

The preeminent method for of covert communication today uses a wide-band, spread-spectrum approach. Unlike traditional radio transmission schemes in which there is a carrier frequency which dominates the frequency spectrum of the signal, spread spectrum signals show no distinct peaks in the frequency domain, and appear more like random noise than information. In fact, signals are chosen to be as noise-like as possible. During the transmission of these signals, the same power levels of typical narrow-band transmissions are used, however, since the power is spread over such a large frequency range, the power spectral density is much lower than traditional narrowband transmitters.

The most common form of this technique today is to use direct sequences. Two sequences are chosen, usually having similar properties as random noise (known as pseudo noise), and usually being antipodal (one signal is the negative of the other - this maximizes the probability of detection for the given set of signals in additive white Gaussian noise). The digital information to be transmitted is then encoded using these sequences and typically single side band suppressed carrier (SSB-SC) amplitude modulated [3].

Due to the aforementioned properties of this scheme, for the receiver, spectral analysis is rendered useless, and other techniques are generally used to detect information in such signals, such as through the use of matched filters. This technique compares the received signal against specific known signal patterns through a correlation process. Using statistical hypothesis testing, a determination is then made as to whether a binary zero or a one is present.

This scheme has one obvious problem, and that is if the binary signal representations become known, then covertness is lost. A major improvement on this system would be to create a spread-spectrum system whose signals are undetectable through both spectral analysis and matched filtering.

THEORY

The novelty of this idea comes from the application of chaos theory and nonlinear dynamical systems in communications. A definition of chaos has not yet been agreed upon, but the general idea is that a chaotic system is one which, while being completely deterministic, is extremely sensitive to initial conditions. This property can yield chaotic signals that have the appearance of random data. This appearance, however, does not discount the underlying determinism, which is the property that is made use of

in order to obtain a level of covertness in the design of these communication systems. One chaotic mapping which is explored here is defined by the Logistic Equation:

$$x_{t+1} = kx_t(1 - x_t) \tag{1}$$

This is a discrete equation which is iterated. When the constant k equals 4, this equation produces chaotic behavior, and it is said to enter the "chaotic domain." Even in this domain, it retains its determinism.

Another important concept is the attractor, which, for a dynamical system (such as the Logistic Equation) is the equilibrium state to which the system converges [5]. In essence, after any initial transients in the system subside, the attractor provides a roadmap as to how the system will behave. The most common method of analyzing an attractor is by making a plot of the signal versus a delayed version of the signal; this plot is known as a phase-space plot.

Below in figure 2 is a plot of the attractor of the system described by the Logistic Equation. On the x-axis is x[n], and on the y-axis is x[n+1]. For any value of x in the range of this equation (which is from -1 to 1), this phase-space plot determines what the next value will be.

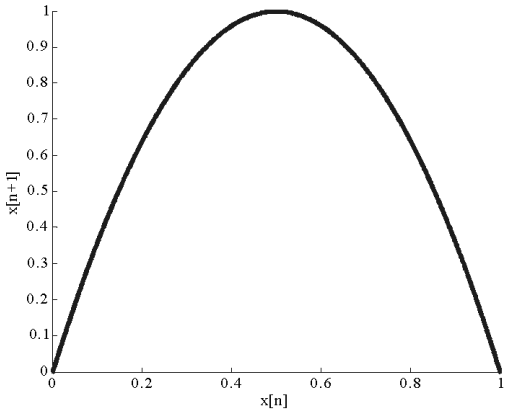


Figure 1 - Plot of the Logistic Equation Attractor

Knowing a system's attractor reveals the determinism of the system, and it is from this concept that the methods presented here were developed.

DEVELOPMENT

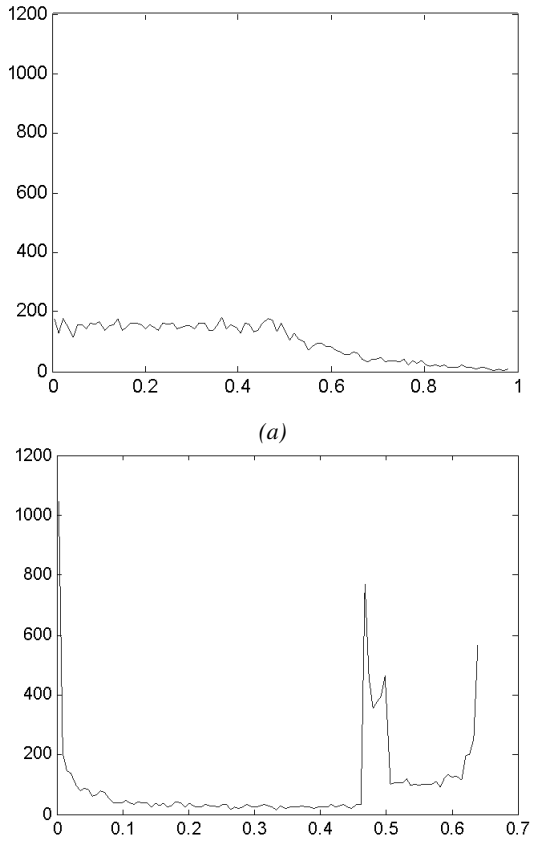
In total there were three systems developed, the first two leading to the third, which was the most practical. Each design presented new insight into the problem and a greater understanding of how to achieve the goal.

The inspiration and first approach came from [1] and their MEan SquAred Histogram (MESAH) algorithm.

MESAH algorithm [1]

Given data set x_i, $i = 1, 2, N$
Let $x_{max} = max(x_1, ..., x_N)$
Let $x_{min} = min(x_1, ..., x_N)$
For $i = 1$ to $N - 1$
$value_i = \frac{1}{2}((x_i - x_{max})^2 + (x_{i+1} - x_{min})^2)$
Get the histogram of the value array

One observation that was made in [2] was that the Logistic mapping tends to transition from one-to-zero and zero-to-one rather often. The MESAH algorithm enhances the detection of chaos by looking for such high-to-low and low-to-high transitions in the input signal. Figure 2 demonstrates the power of this algorithm. In (a) the MESAH output for the random white Gaussian data. There is no distinct peak for this set of random data, however, in the chaotic data's MESAH output in (b), there are two distinct peaks: one at 0, and one at approximately 0.47.



(a)

(b)

Figure 2 - Comparison of MESAH Output for (a)Random and (b)Chaotic Signals

This algorithm made it possible to distinguish chaos from noise. McDonough [2] made the suggestion that finding two chaotic signals with distinct MESAH outputs could provide a basis for a covert binary communications system. What would have to be found were two signals which, when passed through the MESAH detector would yield histograms which were uncorrelated. This would allow for a binary communication system to be established which may be able to function in lower signal-to-noise ratios (SNRs). It was determined, however, that the MESAH algorithm, without alteration, did not yield any histograms that were uncorrelated. For that reason, the MESAH algorithm was modified to help to "uncorrelate" some of the MESAH outputs.

Revised MESAH Algorithm

Given data set $x_i,\ i = 1, 2, ..., N$
Let $x_{max} = max(x_1, ..., x_N)$
Let $x_{min} = min(x_1, ..., x_N)$
for $i = 1\ to\ N - 1$
$value_i = x_i - kx_{i+1}$
Get the histogram of the value array

This new algorithm's output ranged from positive to negative values, while the original algorithm only allowed for positive values. This differentiated a high-to-low transition from a low-to-high transition. Also, a tuning factor was added, the constant k, to be able to adjust the algorithm to minimize the correlation between the chosen chaotic signals. Those signals that were chosen for this method were the Logistic Eq.(2) and Lorenz Eq.(3) equations:

$$x_{t+1} = 4x_t(1 - x_t) \tag{2}$$

$$x_{t+1} = (x_t - 2)^2 \tag{3}$$

These two equations are iterative mappings. A random starting value is chosen, and through iteration the mapping determines the next values in the sequence. These mappings yield values in different ranges. The Logistic Mapping yields values from 0 to 1, while the Lorenz Mapping yields values from 0 to 4. For the system being developed to be covert, the signals should have a similar standard deviations in their output for a zero and a one, in order to equalize the energy in the signals. Therefore, these two mappings were scaled to be between -1 and 1. These scaled Logistic Eq.(4) and Lorenz Eq.(5) mappings are as follows (these will be used as binary zero and one representations, as indicated by the respective subscripts):

$$f_1(x_t) = -2{x_t}^2 + 1 \tag{4}$$

$$f_0(x_t) = 2{x_t}^2 - 1 \tag{5}$$

Sending the signals described by the two above mappings through the revised MESAH algorithm, with k = 0.25, yielded the plots in figure 3. These two histograms had a correlation of 0.04.

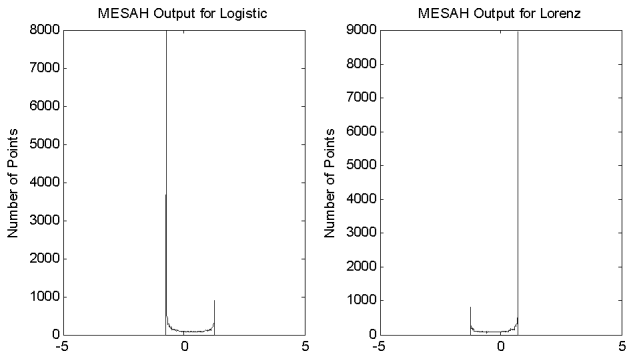


Figure 3 - Revised MESAH Output Scaled Logistic and Lorenz Equations

To simulate this system, a random string of binary bits were created in Matlab and each '1' was represented by a set of N iterations of Eq.(4), and each '0' was represented by N iterations of Eq.(5). Additive white Gaussian noise (AWGN) was then added to that signal, and the received signal was put through the revised MESAH algorithm. The two histograms shown in figure 3 would be used for comparison through correlation. If the received signal's MESAH output for a given bit was more closely correlated to the MESAH output of equation (4), then the decision would be made that a '1' was received, otherwise, it would be decided that a '0' had been received.

This scheme did work for noiseless signals, but performed quite poorly even in the slightest presence of noise. The results can be seen receiver operating characteristic (ROC) curves shown in figure 7. Since binary phase shift keying (BPSK) is the optimal binary signaling scheme for AWGN, the MESAH scheme was compared to the performance of BPSK. It was found that there was a 33 dB difference in performance between the two systems. This made it clear that a new scheme had to be explored.

It was during the implementation of this first design that two interesting facts came to light. The first was that the pure intensity histograms of both of the signals chosen for the zero and one representations were identical. Between the identical intensity histograms and the signals' flat Fourier Transform, this limited the ways that the signal could be distinguished from one another, adding to the potential covertness. The greatest hope came when looking at the phase-space plots of each of these signals, where the attractors of the two signals are antipodal - the mapping defined in Eq.(4) is essentially an upside down parabola, while Eq.(5) is the same parabola mirrored about the horizontal axis. The correlation between these two phase-space plots was found to be approximately -1.

With this knowledge, a second method was developed based on comparing the behavior of the signals as they approach the attractors in phase space. The chosen chaotic signals are extremely localized in phase space. AWGN, on the other hand, is not as localized. Therefore, the second method was to compare the phase space plots of each bit of received data, in order to determine if it matches the mapping of Eq.(4) or the mapping of Eq.(5). To accomplish this a two-dimensional histogram was taken of the received signal, as shown in figure 4, where the phase-space was divided into an $n \times n$ grid, and the value of each data point was the number of points of the received signal that fell into that "box." The upper two plots (note the axes are $x[t+1]$ vs $x[t]$) show the phase-space histograms of, on the left, Eq.(4), and on the right Eq.(5). These upper histograms represent the signals without noise. When 40 dB of noise is added, as in the lower two plots, one can see that the phase-space plots become less localized, but retain their basic shape.

For this design, the transmitter sends the signals in exactly the same way as for the MESAH scheme, however, the phase-space histogram is found for every bit, and the histogram which is produced is then correlated with the two noiseless phase-space plots shown below as the upper two plots. Whichever correlation was higher was chosen

as the received bit, be it a zero or a one.

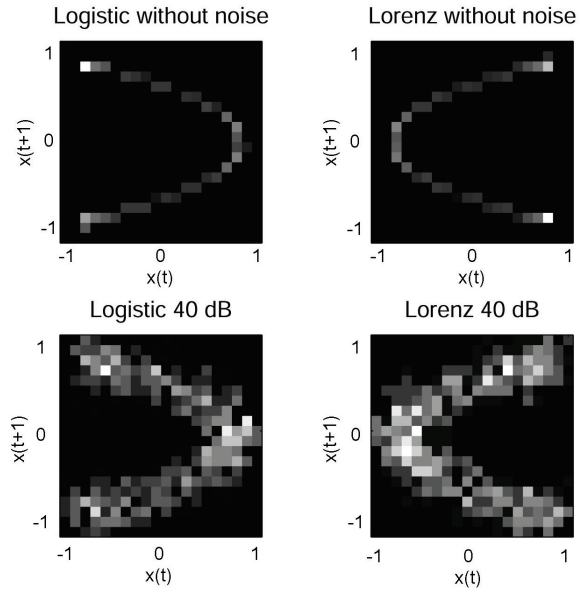


Figure 4 - Attractor Dimensionality Method (Box Counting)

This system worked in noiseless environments and performed better than the MESAH scheme, improving the SNR performance by about 13 dB. However, this technique was still 20 dB worse than BPSK. One lesson this method did teach was that having to histogram the signal led to larger sample sizes and poorer performance at lower SNRs.

THE DESIGN

As in the previously explained methods the transmitter intakes a binary data string and produces N samples for each bit. The transmission begins with a random number between -1 and 1, then each of the remaining N-1 data points are determined by iterating using Eq.(4) if a one is desired, or using Eq.(5) if a zero is desired. This data is then amplitude modulated and transmitted.

The transmitted data will likely encounter AWGN, and therefore the receiver will see, for each sample x_i, the addition of noise, n_i. Therefore, the received signal is $\sigma_i = x_i + n_i$.

Given this sample, it could have either come from the '0' or the '1' mapping. Therefore, one of the original mappings, in this case the one-mapping was chosen, is applied to the received value to predict the next point x_{i+1} (the second mapping would simply be the negative of that predicted value). This same method is applied to every point from x_1 to x_{N-1}. This data series of length $N-1$ is then correlated with the actual received values x_2 to x_N. This is a comparison of the values predicted based on one attractor against the actual received values. If this correlation is greater than zero, then it is statistically likely that the received data represented a bit value of '1' and the decision would be that a '1' had been received; otherwise, the decision would be that a '0' had been received. Figure 5 summarizes the entire system in both block-diagram and algorithmic form.

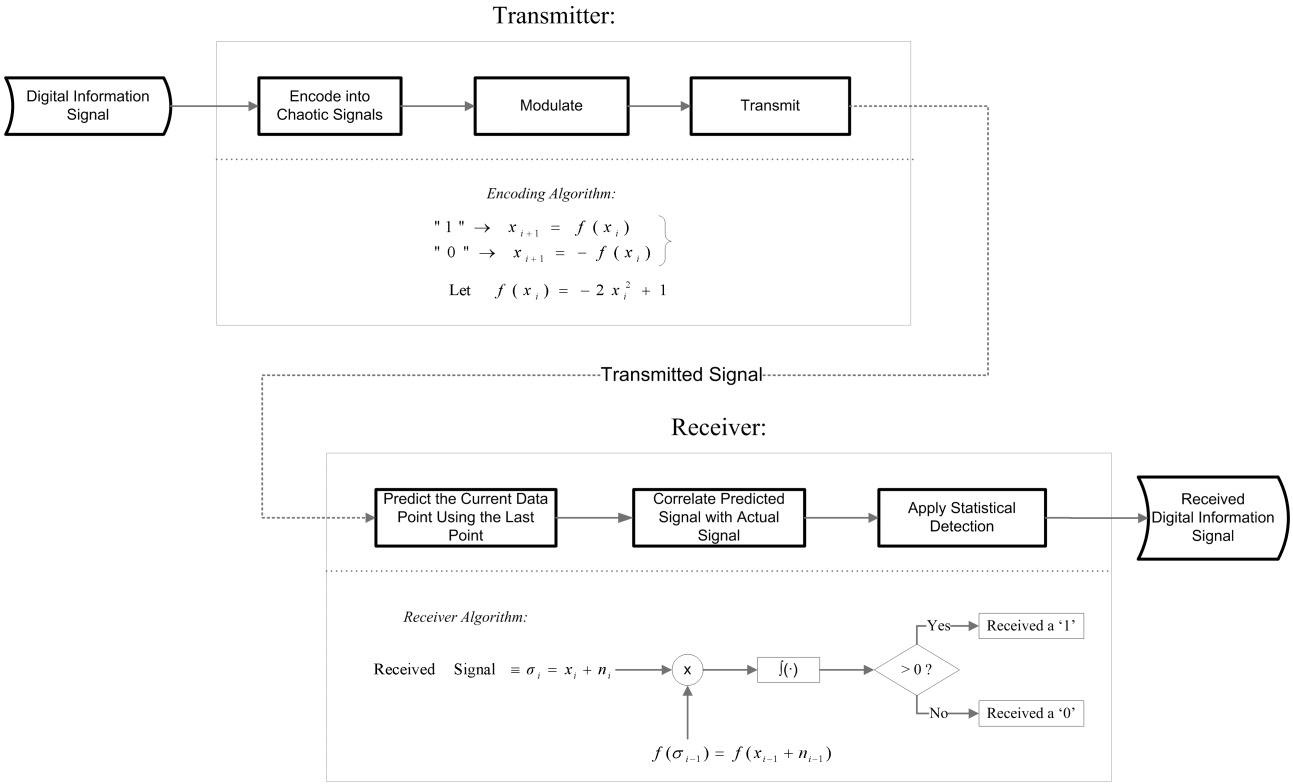


Figure 5 - Attractor Predictor Method Flowchart

RESULTS

The attractor predictor method was analyzed through Matlab simulations in order to judge its covertness in the time and frequency domain, its SNR performance, and its computational complexity.

This scheme was tested for time domain covertness by simulating a matched filter receiver and applying that receiver to transmitted data from the attractor-predictor approach. A time series of one set of iterations for a '1' and a '0' was provided as the two basis signals. Then, a noiseless transmission was simulated and the probability of error for that matched filter receiver was determined. With 100,000 bits simulated and a sample size of 40 samples per bit, the probability of error was 0.4357, which would prove matched filters useless in receiving this data.

In order to ensure the frequency domain covertness of this system, noiseless representations of ones and zeros were produced using the attractor predictor encoding method, and the Fourier Transforms of these signals are presented in figure 6. There are no prominent peaks in this transform, and have the potential to appear as noise to spectral analyzers.

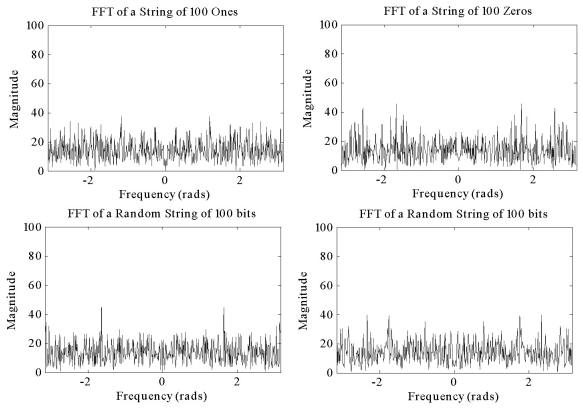


Figure 6 - Fourier Analysis of Encoded Data

In order to analyze the SNR performance of this receiver (and each of the other methods presented in this report), simulations of each of the chaos receivers were developed to transmit and receive a random set of data in AWGN and determine the receiver operating characteristic (ROC) curves, which can then be directly compared to the optimal system for Gaussian noise: binary phase shift keying (BPSK), and differential binary phase shift keying (BDPSK), the noncoherent analog to BPSK. As previously stated, the MESAH and Box-counting methods were 33 dB and 20 dB, respectively, away from BPSK. The Attractor-Predictor method performs 10dB worse than BPSK. However, it performs only 7dB worse than differential BPSK (DBPSK), which is a fairer comparison, as it is also noncoherent. The ROC curves are shown in figure 7.

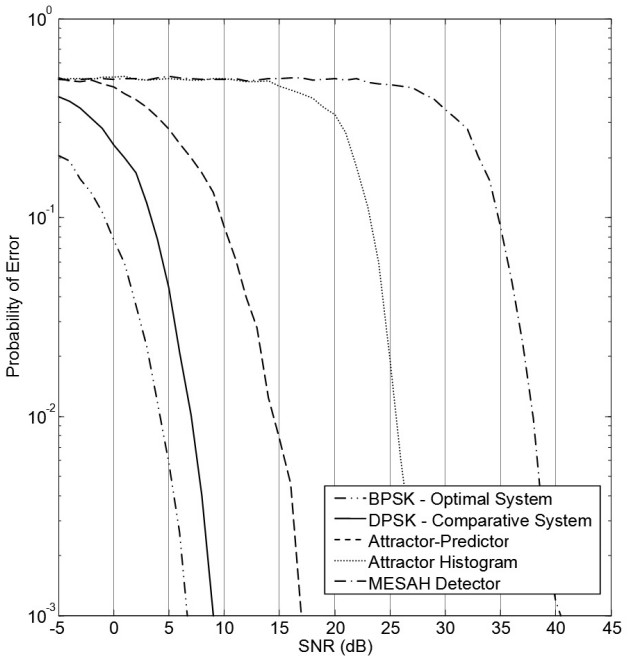


Figure 7 - Receiver Operating Characteristic Curves

In order to compare the computational intensity of the attractor-predictor method to BPSK, the reception of two identical random sequences of 1,000,000 bits were simulated, sent using the same number of samples, and the time required to received those bits using a BPSK receiver and the Chaos receiver, respectively was determined. The chaos-based approach required on average 1.5 times more time to receive a given signal than BPSK.

CONCLUSION

In summary, the receiver design is noncoherent, as no knowledge of phase is required, and no knowledge of the starting values of the chaotic signals is necessary. In terms of covertness, all generated signals are spread-spectrum in the frequency domain, containing no distinguishing peaks or noticeable structure. In the time domain, signals of sample size 10 or greater are only slightly correlated, and those 40 or greater are uncorrelated, eliminating matched-filtering as a viable receiver option. In addition, the signal reception time is only 50 percent greater than that of BPSK. As for the performance of this receiver in low SNRs, the system performance does degrade gracefully with the introduction of noise, however, its overall performance is 10dB away from that of BPSK and 7 dB away from DBPSK. Due to the covert nature of this system and the relative simplicity of the algorithm's implementation, a tradeoff in SNR performance is acceptable.

REFERENCES

[1] J. P. Noonan McDonough, P. V. and G. R. Hall. A new chaos detector. *Journal of Computers and Electrical Engineering*, 21(6), 1995.

[2] P. V. McDonough. *A New Chaos Detector and Applications to Communications*. PhD thesis, Tufts University, 1993.

[3] Randy Roberts. Introduction to spread spectrum. Spread Spectrum Scene Online. Pegusus Technologies, 2006.

[4] Steven H. Strogatz. *Nonlinear Dynamics and Chaos*. Perseus Book Publishing, LLC, 2000.

[5] Garnet P. Williams. *Chaos Theory Tamed*. John Henry Press, 2006.