Independent Communication Messages: Methodology and Applications

J. Arlin Cooper
Sandia National Laboratories*
Albuquerque, NM 87185-0490
(505) 845-9168
FAX (505) 844-9225
acooper@sandia.gov

Abstract

Information flowing on communication buses is ordinarily "non-random" in the sense that data entities are not equally likely and independent. This is because they have relationships to each other and to physical occurrences to which they may be responding. Random data would convey no information or meaning.

From a different viewpoint, there can be applications for creating randomness characteristics, and four of these are described in this paper. Two examples derive from cryptology and the other two from safety. One cryptology application described is the generation of random numbers for use as, for example, keys, hash functions, nonces, and seeds. The other is for inter-message "padding" to resist traffic analysis by masking when data are being transmitted and when the channel is conveying no information.

One of the safety applications described is the "unique signal" approach used in modern nuclear weapon electrical safety. The other is the use of unique signals as non-weapon critical-operation control functions. Both of these safety applications require provisions to help assure randomness characteristics in any inadvertently occurring inputs. In order to satisfy these cryptology and safety needs, communication strategies are described that generate or selectively encourage independent (unrelated) symbols or messages.

Introduction

In a logical sense, information flowing on most communication buses ordinarily appears as "bits" ("zeros" and "ones"). Some forms of communication traffic analysis depend on estimating the relative numbers of zeros and ones in unknown traffic, estimating the relative numbers of zeros and ones in various communication logical positions of such traffic, and estimating the Markov characteristics (e.g., the number of ones followed by ones, the number of ones followed by zeros, etc.). While assumptions of random characteristics may give useful results on the average, there are situations where such analyses can be extremely misleading. This might be expected, since truly random bits would convey *no* useful information. One of the objectives of this paper is to outline an analysis approach that addresses such problems.

*Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin company, for the United States Department of Energy under Contract DE-ACO4-94AL8500.

There is a reverse situation leading to a set of synthesis problems. If it is desired to preclude any useful intelligence to an "outsider" in a communication sequence, it is desirable to *create* randomness characteristics. Although "pseudorandom" synthesis strategies are well known, so are after-the-fact analysis procedures, which can quickly detect pseudorandom characteristics. This means that protective difficulties may arise, depending on the amount of design structure involved in creating pseudorandomness. In this paper, new techniques for enhancing randomness will be described. Four applications will be discussed. Two examples derive from cryptology and the other two from safety.

On the "Analysis" of Non-Random, Unknown Communication Characteristics

The difficulty is obvious, since the unknown cannot be analyzed. The problem is that it is tempting to *over-analyze* by assuming characteristics that may not exist. For example, the assumption of random characteristics can be misleading. Under this assumption, the one/zero balance of unknown communication bits can be probabilistically computed as equal. Furthermore, the probability of a one/zero balance for a bit following a one (or a zero) would be computed as one-half. Similar assumption-based results could be invalid because of the inherent dependence in message traffic. Neither of these assumptions (equally likely and independent) will be true for meaningful messages.

Another common practice is to represent lack of knowledge by a "non-informative" probability distribution (e.g., a uniform distribution). But if the characteristics are unknown, this too is misleading, because every point across the range of uniformity is represented as exactly as likely as any other point. One type of representation that solves the above problems is a *possibilistic* distribution. A possibilistic analysis more accurately represents the available knowledge when knowledge is sparse or lacking altogether.

In order to show this with a simple example, the uniform probability density function (PDF) shown below indicates the likelihood of "ones" for a family of unknown communication channels or communication sources, under the tacit assumption that any probability from zero to one is equally likely. As a result, the mean value for the probability of a "one" is one-half.

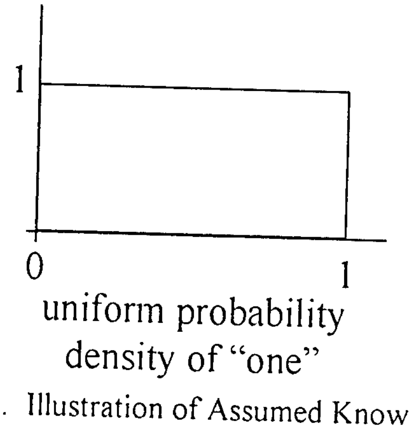

Figure 1. Illustration of Assumed Knowledge

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, make any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Consider four bits of information on a communication channel that are to be assessed for the probability of matching some unique pattern, say 0110. Assume uniform distributions between zero and one for the probability of matching each bit. Now we will use a Monte Carlo simulation to take four samples from the distribution and multiply the sample probabilities together. This will be repeated 20 times, each time computing a cumulative average, as shown in Fig. 2.

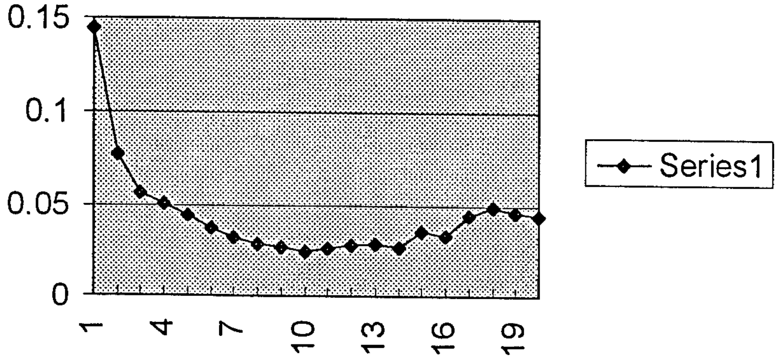

Figure 2. Cumulative Average of Four Random Samples on {0, 1}

An interesting feature of this simulation is that although the average will eventually reach 1/16, the initial sample gives about 1/7. The result shows that even if uniform distributions are a correct model, the first sample that we see may not be average.

Another way to visualize the problem is to observe averages of various numbers of Monte-Carlo sampled density functions. Figure 3 shows averages of up to 12 uniform distributions (most peaked curve) from one (least peaked curve). As expected from the Central Limit theorem, the average of a large number of independent linear distributions over {0, 1} approaches a Gaussian (Normal) distribution with mean 0.5. However, the plot shows that a large number of samples is required before the "theoretical" value is approached. Also, the necessary Central Limit Theorem assumptions are not easily met.

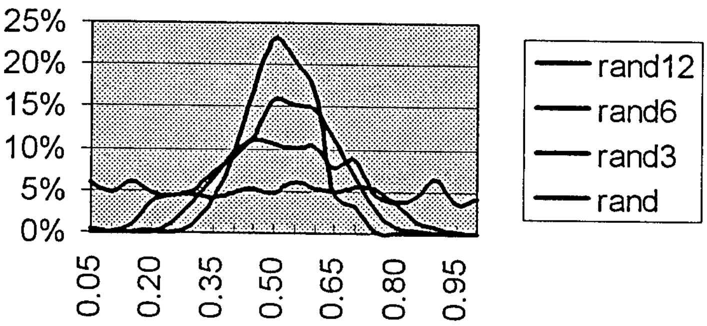

Figure 3. Monte-Carlo-Generated Averages of Various Numbers of Uniform Distributions

As a result of these observations, we conclude that point-estimate probabilities and probability distribution functions are misleading indicators of exposure to an unknown data stream.

But if probabilities and probability functions are not valid indicators of our state of knowledge, is there *any* possible indicator? The answer is yes. Consider a "possibilistic" distribution [Ref. 1] for the occurrence. Figure 4 shows the possible values for probability of a zero (or one). The ordinate is a possibility "membership." The maximum membership is one by convention, and the area integral is not constrained, as it would be in a PDF.

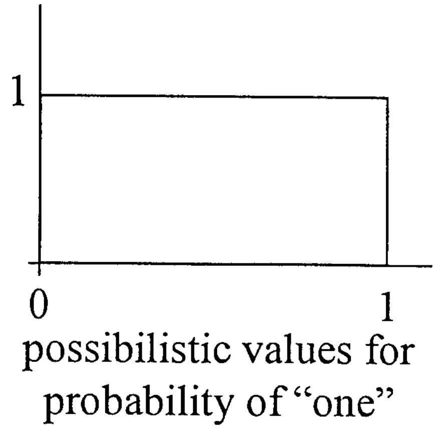

Figure 4. Illustration of One-Bit Possibility Function for No Assumed Knowledge

This distribution conveys the information that any value is possible, but no more extensive knowledge is available, and the mean value remains unknown. This accurately represents the available knowledge; the PDF does not. The possibilistic representation of the probability of a particular four-bit pattern is shown in Figure 5.

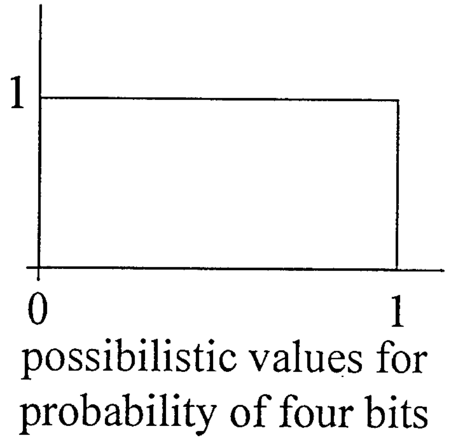

Figure 5. Illustration of Four-Bit Possibility Function for No Assumed Knowledge

Although Fig. 5 appears identical to Fig. 4, it is actually obtained from Fig. 4 by possibilistic multiplication [Ref. 2]. This result differs sharply from Fig. 2, where we depend on averages of random functions to approach the result 1/16.

These observations would be of little value if more sophisticated problems were not tractable. We will illustrate an analysis that has some similarities to those that must be accomplished for judging the safety of "unique signals" (UQSs) used to pre-arm nuclear weapons.

First, assume a 24-bit unique pattern (these are actually called UQS "events" for reasons we will explain later). We need to assess the probability of matching that pattern with a single occurrence of 24 bits from a communication channel. A possibility function such as shown in Fig. 5 is not helpful, because it just bounds the probability between zero and one. A probability function such as shown in Fig. 2 (converging toward a value of $6x10^{-8}$) is not appropriate, because it is only valid for extensive averages.

We will introduce a triangular possibility function for each bit probability, peaked at ½ as the most "possibilistic" value (Fig. 6).

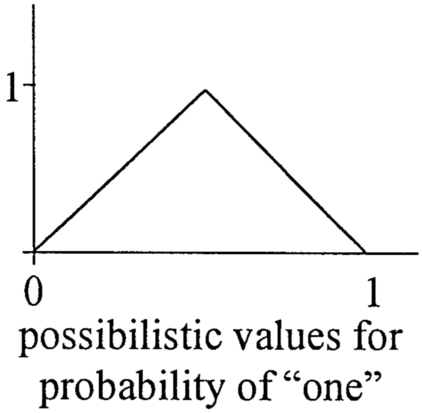

Figure 6. Triangular Possibility Function for a Single Bit

For multiple independent bits, the peak values of probability simply multiply. For dependent values, the result can be greater (if the dependence characteristics match the unique signal) or less (if they are counter to the unique signal). For illustration, assume that the multiplication has an additional factor "d" ($0<d<2$) to represent the effective dependence. The resultant peak value is then $(d/2)^{24}$. If d cannot be determined, the result becomes the same as that shown in Fig. 5. However, if we are able to learn enough from communication characteristics to better bracket d, the information is useful. For example, if $0.7<d<1.4$, we obtain the trapezoidal possibilistic function shown in Fig. 7 (drawn on a logarithmic scale). Although this dependence is similar to actual results compiled from communication statistics, there are communication strategies that reduce the effective dependence. For example, if unique signals are communicated using separate transmissions for each event, the dependence in inadvertent communication is reduced since samples received inadvertently are from more unrelated message bits [Ref. 3]. This is shown in Fig. 8.

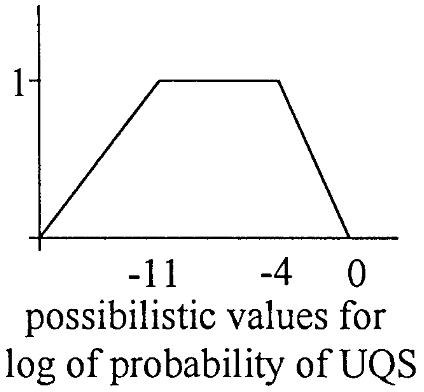

Figure 7. Possibility of UQS Considering Dependence

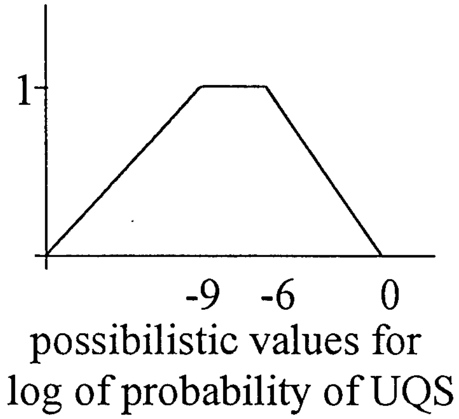

Figure 8. Possibility of UQS Using Separate Communication Entities

Applications for Synthesis

There are many applications for synthesis of random-appearing sequences. Some cryptology applications are for use as, for example, keys, hash functions, nonces, and seeds. In all of these cases unpredictability of what might be generated is of value. Another application is for inter-message "padding" to resist traffic analysis. The objective is to give no indication that messages are not being communicated. In all of the above applications, randomness properties are an attribute, but predictability is not. A common approach is to use "pseudorandom" generators, constructed through linear algebra operations. The advantages are that many common randomness properties are satisfied, and the generation is easily constructed mathematically and through logic circuits. The disadvantages are that some randomness properties are not met, and analysis can efficiently detect pseudorandom characteristics.

Safety applications for synthesized randomness include the "unique signal" approach used in modern nuclear weapon electrical safety and the use of unique signals as non-weapon critical-operation control functions. Both of these safety applications require provisions to help assure randomness characteristics (but not pseudorandom characteristics) in any communicated patterns.

Observations Concerning Synthesis of Random Patterns

One of the important properties of random patterns is that the number of "ones" should be as nearly equal as possible to the number of "zeros." The reason for this is that the maximum probability of inadvertently generating an equal number of ones and zeros is thereby minimized [Ref. 3]. Similarly, the number of ones followed by ones should be as nearly equal as possible to the number of ones followed by zeros, and the number of zeros followed by ones should be as nearly equal as possible to the number of zeros followed by zeros. These constraints (equivalent to protecting against first-order Markov dependence) can be visualized by a three-dimensional pattern such as that shown in Fig. 9. The abscissas represent, for example, the percentage of ones and the percentage of ones followed by ones. The maximum likelihood of inadvertently matching a pattern (illustrated by the peak of the three-dimensional image) is minimized by achieving the balance indicated.

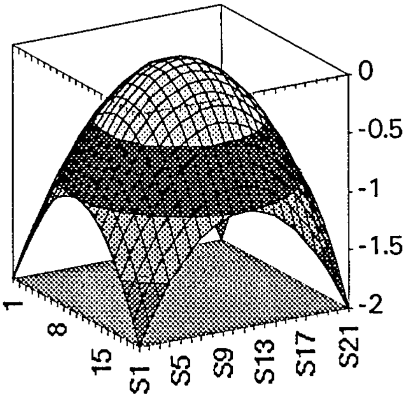

Figure 9. Three-Dimensional Shape Indicating Maximum of Two Pattern Choices

Similar extensions can be made to other orders of Markov dependence and other types of dependence. However, we will limit our discussion here to a non-random characteristic of pseudorandom patterns. In a pseudorandom pattern, the number of strings of length n ones equals (within the constraint of length-(2^m-1) sequences) the number of strings of length n zeros for 0<n<m. Improved randomness patterns can be achieved (but not with linear generators) by assuring that these matches no *not* exist.

Another interesting observation is that "mixing" two data streams using an exclusive-or function statistically increases randomness in the following sense. Whatever the probabilities of ones and zeros in the two data streams, the exclusive-or probabilities of ones and zeros will be closer to 0.5 than were either of the data streams (or equal to 0.5 if the probabilities of any entity are 0, 1, or ½. What this means at least on the average is that randomness can be enhanced by exclusive-or mixing of two (or more) data streams.

References

1. Zadeh, Lotfi "Fuzzy Sets as a Basis for a Theory of Possibility," *Fuzzy Sets and Systems*, Vol. 1, No. 1, 1978

2. Kaufmann, Arnold, and Madan Gupta, *Introduction to Fuzzy Arithmetic*, Von Nostrand Reinhold, 1991

3. Spray, Stan, and J. A. Cooper, "The Unique Signal Concept for Detonation Safety in Nuclear Weapons, Sandia National Laboratories Report SAND91-1269, June, 1993

M98000234

Report Number (14) SAND--97-2324C

CONF-971045--

Publ. Date (11) 199709

Sponsor Code (18) DOE/DP, XF

UC Category (19) UC-700, DOE/ER

DOE