Permutation Entropy of Card Shuffling

Guido David[*]
Institute of Mathematics and
Computational Science Research Center
University of the Philippines - Diliman
Quezon City, Philippines
gdavid@math.upd.edu.ph

ABSTRACT

The classical viewpoint is that seven shuffles are needed to sufficiently randomize a deck of playing cards. On the other hand, according to information theory, six shuffles are sufficient. The numerical shuffling model used in this paper excluded the identity as a possible shuffle, in contrast with other studies. Using Monte Carlo simulations, permutation entropies were numerically calculated for 2 to 9 riffle shuffles. The results show that in the context of entropy, six shuffles are sufficient for randomizing a deck of 52 cards, and five shuffles are acceptable for casual play.

Keywords

permutation entropy, riffle card shuffle, randomness, information theory, computational methods, Monte Carlo simulation

1. INTRODUCTION

The riffle shuffle is the most commonly used shuffling method in card rooms, home games and casinos all over, due to its ease of use. Effectively randomizing a deck of cards is necessary for the integrity of the game, and to prevent enterprising players from taking advantage of probabilistic inferences on card distribution and patterns. This also applies to computer and online games that rely on permuting decks of cards.

According to Diaconis' work on card shuffling [1][3][4], at least 7 riffle shuffles are needed to sufficiently randomize a deck of 52 cards, based on the variation distance between distributions of the shuffled deck and the theoretical uniform distribution. On the other hand, investigations by Trefethen and Trefethen [6] showed that, in the context of information theory, 99% of information has been lost after six shuffles. In this paper, permutation entropy is used as a criteria for determining the number of shuffles needed in order for a deck of cards to be considered random.

2. METHODOLOGY

2.1 Binary Representation

Performing Monte Carlo simulations of the riffle shuffle requires a model for representing shuffles. A riffle shuffle of a deck of m cards is represented by a binary string of length m, where the 0's indicate placement, in the same sequence, of the first portion of the deck, and the 1's indicate placement of the cards, also in sequence, from the second portion of the deck [1]. For example, the binary string of length 8 below represents the permutation via riffle shuffle of a deck of eight cards, labeled 1 to 8, initially unshuffled:

$$(0\,1\,0\,0\,1\,0\,0\,1) : (1\,2\,3\,4\,5\,6\,7\,8) \to (1\,6\,2\,3\,7\,4\,5\,8) \qquad (1)$$

The identity shuffles are represented by binary strings starting with k 0's followed by $(m-k)$ 1's, where k ranges from 0 to m. For example,

$$(0\,0\,0\,0\,0\,1\,1\,1) : (1\,2\,3\,4\,5\,6\,7\,8) \to (1\,2\,3\,4\,5\,6\,7\,8) \qquad (2)$$

There are a total of $m+1$ identity strings. The total number of binary strings is 2^m. Each nonidentity shuffle has a unique binary correspondence. To see this, note that dividing a deck into two portions with lengths k and $m-k$ and interleaving the cards of the second portion with the first portion results in either one or two rising sequences [5]. For example, the permuted deck in (1) has two rising sequences: (1 2 3 4 5) and (6 7 8). The only shuffle resulting in one rising sequence is the identity shuffle, as can be seen in (2). Now consider the deck that has been shuffled once. If there are exactly two rising sequences, then we know exactly the two portions of the original cut. For example, in the permuted sequence in (1), it is clear that the two portions are exactly the rising sequences, i.e. the initial cut occurred after card 5. The number of ways of mixing the two portions of the deck when the cut is at k (i.e. the two portions of the deck consist of k and $m-k$ cards) is exactly equal to ${}_mC_k$. From the Binomial Theorem, the sum of shuffles from all possible cuts is 2^m, which is exactly the number of possible binary strings of length m. For each cut k, the number of shuffles includes exactly one identity shuffle. The total number of possible cuts is $m+1$, the same as the number of identity shuffles, thus there are $2^m - m - 1$ unique nonidentity shuffles. Because an identity shuffle does not occur in practice, the algorithm used in this study discounts identity permutations. This is addressed by repeating the shuffle whenever the outcome is an identity shuffle.

For a standard 52 card deck, the total number of possible riffle shuffles is less than 2^{52} or approximately 4.5×10^{15}. In

[*] Corresponding author

a 64-bit numerical software such as Matlab®, the smallest number, $eps = 2.2 \times 10^{-16}$. This implies that there is sufficient resolution for a standard random number generator to produce every shuffle of a deck of 52 cards. By applying several shuffles, every permutation of the deck may be achieved. In contrast, simply mixing the cards requires 52! possibilities, a number too large for most random number generators. As a result, most numerical card shufflers only achieve a subset of all possible shuffles, or use several randomly generated numbers to mix the cards.

2.2 Permutation Entropy

The permutation entropy of order n is defined as

$$H(n) = -\sum_{k \in S_n} \left(\frac{p_k}{m-n+1}\right) \log_2 \left(\frac{p_k}{m-n+1}\right) \qquad (3)$$

where m is the number of cards in the deck and p_k is the number of occurrences of the permutation k, an element of the permutation group S_n [2].

The maximum of the permutation entropy of order n is $\max H(n) = \log_2(n!)$. However, for a finite deck of cards, the theoretical maximum is usually not achieved. In order to compare entropies of various orders, we calculate the percentage of entropy from the maximum as

$$R(n) = \frac{H(n)}{\log_2 n!}. \qquad (4)$$

As an illustration, two successive random riffle shuffles of a deck of $m = 12$ cards produce the permutation

$$(5\,9\,10\,1\,2\,3\,6\,4\,11\,7\,8\,12) \qquad (5)$$

The total number of strings of length 2 is $m-1 = 11$. There are 8 strings of length 2 which are increasing, i.e. of (01)-type, and 3 strings of length 2 that are decreasing, i.e. of (10)-type. Thus for this example, the permutation entropy of order 2 is

$$H(2) = -\left(\frac{8}{11}\right)\log_2\left(\frac{8}{11}\right) - \left(\frac{3}{11}\right)\log_2\left(\frac{3}{11}\right) = 0.8454 \qquad (6)$$

Note that $\log_2 2! = 1$, thus the percentage of entropy is $R(2) = 0.8454$. Similarly, to compute $H(3)$, we count the strings of length 3 of type (012), (021), (102), (120), (201) and (210). The calculated value of $H(3)$ is 2.1219. Since $\log_2 3! = 2.5850$, then $R(3) = 0.8209$. For a deck of 52 cards, the permutation entropy up to order 4 is calculated. The permutation entropy of order 5 has $5! = 120$ possible strings of length 5, but only 48 of such strings can be counted from 52 cards, hence permutation entropies of order 5 and higher are not considered.

3. RESULTS AND DISCUSSION

Monte Carlo method is used to calculate permutation entropy by numerically shuffling a 52 card deck 2 to 9 times, and repeating over 100,000 simulations. To reduce correlation between the entropy results, only one entropy value is calculated per run, with 100,000 simulations per run. The shuffle entropies of orders 2, 3 and 4 as a percentage of the maximum plotted against the number of shuffles are presented in Figure 1.

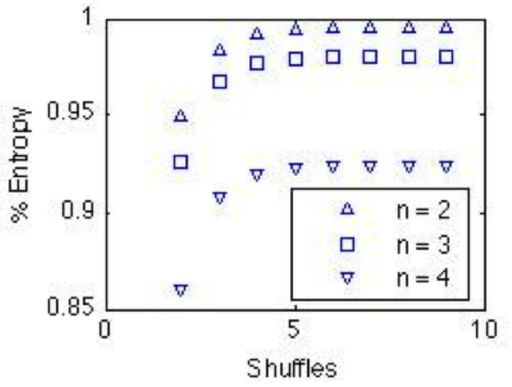

Figure 1: $R(n)$ or percentage of permutation entropy vs. the number of shuffles, obtained by Monte Carlo method with 100,000 simulations per entropy calculation.

The percentages of entropy after 5 shuffles are approximately $R(2) = 99.4\%$, $R(3) = 97.9\%$ and $R(4) = 92.3\%$. From 6 to 9 shuffles, the changes in entropy are very small, less than our tolerance of 0.001%. The approximate permutation entropies are $R(2) = 99.5\%$, $R(3) = 98.0\%$ and $R(4) = 92.4\%$. In other words, shuffling more than 6 times will not cause any noticeable increase in entropy or randomness. Hence, using entropy as the criteria, 6 shuffles are sufficient for randomizing a deck of 52 cards.

Note that the change in entropy from 5 to 6 shuffles is 0.01%, which is almost negligible. It may be argued that 5 shuffles are sufficient for general purposes. In other words, for most games where the participants will not be able to draw probabilistic inferences based on available information, then 5 shuffles may suffice. However, when erring on the side of caution, at least 6 shuffles are recommended. The results also show that there is very strong correlation between the permutation entropies of order 2, 3 and 4. This is evident from Figure 1, which shows similar slopes for the different orders of entropies when the number of shuffles is the same. However, calculating all three orders of entropy is still necessary. For example, a large permutation entropy of order 2 may be obtained after one perfect shuffle [4], which is clearly not well shuffled.

4. CONCLUSIONS

Numerical results showed that in the context of entropy, 6 riffle shuffles are sufficient to randomize a deck of 52 cards. This agrees with results from information theory [6], and suggests that the classic rule of using 7 shuffles [1][3][4] may be a little conservative. Note that other models included the identity shuffle (i.e. no shuffle) when counting all possible riffle shuffles, whereas in this study, the identity shuffle was not counted.

If time is not a factor, then shuffling more times than recommended cannot hurt. For numerical shufflers used in computers and online games, additional shuffles will hardly add computational time. For example, the numerical algorithm used in this paper took 0.0034 seconds to shuffle 10 times. However, when time is critical such as in casinos and card rooms, the negligible change in randomness after the 6th shuffle does not justify additional shuffles. In fact, for most play purposes, 5 shuffles may be sufficient because additional shuffles would only increase entropy by less than 0.01%. Diaconis himself stated that 5 shuffles are sufficient for 'casual play randomization' [4]. One common argument against 5 shuffles is that not every permutation of the deck is represented. For example, it can be shown using rising sequences that at least 6 shuffles are needed to reverse the order of the deck. Rigorous requirements are not essential for most games where casual players are unable to take advantage of flaws in the shuffling method. Note that most of the time, the starting state of the deck is already somewhat random from previous play. Thus a dealer may opt to shuffle 6 or more times when using a new (unshuffled) deck of cards, and less for subsequent shuffles.

5. ACKNOWLEDGMENTS

This project was supported by the CHED Center of Excellence Small Scale Grants for Cutting Edge Math Research, Project Reference No. 2008.149.

6. REFERENCES

[1] D. Aldous and P. Diaconis. Shuffling cards and stopping times. *The American Mathematical Monthly*, 93(5):333–348, 1986.

[2] C. Bandt and B. Pompe. Permutation entropy - a natural complexity measure for time series. *Physical Review Letters*, 88(174102):1–4, 2002.

[3] D. Bayer and P. Diaconis. Trailing the dovetail shuffle to its lair. *Annals of Applied Probability*, 2(2):294–313, 1992.

[4] P. Diaconis, R. L. Graham, and W. M. Kantor. The mathematics of perfect shuffles. *Advances in Applied Mathematics*, 4(2):175–196, 1983.

[5] B. Mann. How many times should you shuffle a deck of cards? *Undergraduate Mathematics and Its Applications Journal*, 15(4):303–332, 1994.

[6] L. N. Trefethen and L. M. Trefethen. How many shuffles to randomize a deck of cards? *Proceedings: Mathematical, Physical and Engineering Sciences*, 456:2561–2568, 2002.