Communications in Information and Systems
Volume 12, Number 4, 277–299, 2012

Random walk and linear switching systems

Yulei Pang, Alex Wang, Xiaozhen Xue, and Clyde F. Martin

In this paper we address the question "for a deck of cards, how many times a top-in shuffle should be performed before the top card goes back to the original position?" This problem has been studied in the literature but we are interested in the implications for linear switching systems. We simulate top-in shuffling for 6, 12, and 54 cards, and determine the underlying statistics. Finally we prove that the distribution of the stoping time is an exponential distribution, and the expect value approaches to that of the uniform distribution for large number of shuffling. We make essential use of the properties of linear, stochastic switching systems.

1. Introduction

Suppose we have a deck of n cards, labeled by integers from 1 to n. We will number the deck so that an original unshuffled deck would be written $(1, 2, 3, \ldots, n)'$. Hereafter, we will call this the natural order. From a mathematical viewpoint, shuffling a deck of n cards can be thought of as a permutation of the the numbers from 1 to n. A deck of n cards can be ordered in $n!$ ways. The outcome order is dependent on which method we choose to shuffle the cards. A description of the most popular shuffles is described in [2]. In this paper we focus on just one type, the so called top-in shuffle. Top-in Shuffle: Take the card from the top and insert it at a random position in the deck [1].

An important part of this paper is to describe the random shuffling of a deck of cards as a linear switching system. We use some results from the theory of linear discrete time stochastic switching systems to describe limiting phenomena of shuffling.

In Section 2 we examine the dynamics of a repeated shuffle and the dynamics of a pair of interlacing shuffles. In Section 3 we introduce the idea of linear stochastic switching system as a model for shuffling. In Section 4 we examine the dynamics of the expected value of random shuffling using the theory of switching systems. In Section 5 We do an indepth study of top-in shuffling and report three simulations on various size decks. In Section 6 we construct the distribution of hitting time for all top in shuffles and show that each distribution is approximately geometric and hence has limit as an exponential distribution. The rate of convergence is explicit in the proof. In Section 7 we study the limiting behavior of the expected value of the top-in shuffle. Again the result is for decks of arbitrary size. Tables and code are contained in the appendices.

2. Dynamics and control

The basis for many tricks with cards revolve around the fact that if a deck is carefully shuffled using the same exact shuffle the cards will eventually return to their original order. We look at a series of simple examples before proving general results. One standard shuffle is given by the following example.

Let a deck of six cards be ordered from 1 to 6, so we have a vector $(1,2,3,4,5,6)'$. We shuffle by interlacing the two halves of the deck to obtain the ordering $(1,4,2,5,3,6)'$. Repeating this we see that

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix} \to \begin{pmatrix} 1 \\ 4 \\ 2 \\ 5 \\ 3 \\ 6 \end{pmatrix} \to \begin{pmatrix} 1 \\ 5 \\ 4 \\ 3 \\ 2 \\ 6 \end{pmatrix} \to \begin{pmatrix} 1 \\ 3 \\ 5 \\ 2 \\ 4 \\ 6 \end{pmatrix} \to \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix}.$$

We can think of this as acting on the vector with the 6 by 6 matrix

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

and the shuffling is represented by the dynamical system

$$x_{n+1} = Ax_n.$$

The fact that it repeated is a simple consequence of the fact that $A^4 = I$.

The underlying mathematics of this card trick is that the trickster appears to be randomizing the order of the deck by repeatedly shuffling but in fact he is simply rearranging the deck in a very precise manner so that in the end it is identical to the first position. The mathematics is simply that a shuffle is a permutation and the permutation can be represented by a single matrix. Because the permutation group is finite we must have that any matrix, C, that represents a shuffle, has to have the property that $C^n = I$ for some value of n.

Now suppose we do really want to randomize the deck. As we see above a single permutation will not work. So let's consider another interlacing shuffle that is given by

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix} \to \begin{pmatrix} 4 \\ 1 \\ 5 \\ 2 \\ 6 \\ 3 \end{pmatrix} \to \begin{pmatrix} 2 \\ 4 \\ 6 \\ 1 \\ 3 \\ 5 \end{pmatrix} \to \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix}.$$

This quick calculation shows that the order of this permutation is 3. The matrix representing this shuffle is given by

$$D = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

If we simply alternate the two shuffles, we are just repeating the shuffle given by AD and we know that $(AD)^n = I$ for some $n \leq 6$. But if we randomly choose the shuffle, we get a shuffle given by

$$A^{n_k} D^{m_k} \cdots A^{n_2} D^{m_2} A^{n_1} D^{m_1},$$

where $n_i \leq 3$ and $m_i \leq 2$. Now there are 6! possible shuffles and so it is a question of algebra if we can generate all the shuffles with these two. We will come back to this question later.

With a single shuffle we were able to represent it mathematically as the dynamical system $x_{n+1} = Ax_n$, but here it is either represented as $x_{n+1} = Ax_n$ or randomly as $x_{n+1} = Dx_n$. We can rewrite these two equations as one stochastic dynamic system

$$x_{n+1} = (u(n)A + (1 - u(n))D)x_n \tag{2.1}$$

where $u(n) \in \{0, 1\}$ and $P(u_n = 1) = .5$.

3. Switching systems

A stochastic discrete time linear switching system is a stochastic system of the form

$$x_{n+1} = (u_1(n)A_1 + \cdots + u_k(n)A_k)x_n \tag{3.1}$$

where the u_i satisfy the following properties

$$\begin{aligned} u_i(n) &\in \{0,1\} \\ \sum_{i=1}^{k} u_i(n) &= 1. \end{aligned}$$

If the u_is are chosen randomly then we must assign a probability at each time n, i.e.

$$Prob(u_i(n) = 1) = p_{in}$$

and for the purposes of this paper we will assume that the probabilities are constant with respect to n and equal to $\frac{1}{n}$. See references [7] and [6] for examples and the treatment of such systems. An example of an application of a linear switching system is given in [5].

In this paper one of our goals is to describe a particular type of shuffle as a stochastic switching system. We use the top-in shuffle because of its simplicity but all shuffles can be represented by similar systems. Basically, any shuffle is given by a set of permutations acting on an ordered deck of cards. Thus the switching systems have as their natural state space the permutations of a single vector. Thus the matrices of the system in given by equation 3.1 are the elements of a matrix representation of the symmetric group. We can take the state space to be the set of $n!$ vectors obtained by taking all permutations of the vector $(1, 2, \ldots, n)'$.

The questions usually ask about shuffles are somewhat foreign to systems theory but are interesting never the less. In systems theory the usual questions have to with stability, controllability, reachability, etc. We will, in the next section, study the expected value of these systems using material from [7].

4. Some system theoretic questions

In [7] it was shown that the expected value and the variance of a discrete time stochastic switching system could be computed recursively with associated dynamical systems. In [3] there is an indepth discussion of when a random walk on a finite group becomes uniformly distributed on the group elements. We examine this issue with the system of equation 2.1. From [7] we see that the expected value of the system acting on the space of permutations is given by

$$E[x_{n+1}] = \frac{1}{2}(A + D)E[x_n]$$

and

$$F_6 = \frac{1}{2}(A + D) = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

First it is clear that the kernel of F_6 is three dimensional and that the vector $(1, 1, 1, 1, 1, 1)'$ is an eigenvector with eigenvalue 1. A bit of calculation yields that another eigenvector is given by $(1, 1, -2, -2, 1, 1)'$ with eigenvalue $-\frac{1}{2}$. With more calculation we find that $-\frac{1}{2}$ is an eigenvalue with multiplicity 2. The kernel of F is spanned by the three vectors $(1, 0, 0, -1, 0, 0)'$, $(0, 1, 0, 0, -1, 0)'$ and $(0, 0, 1, 0, 0, -1)'$. So it remains to find the generalized eigenvector associated with the eigenvalue $-\frac{1}{2}$. It is somewhat easier to calculate a vector that is orthogonal to the five known eigenvectors. It is easy to find $(-1, 0, 1, -1, 0, 1)'$.

We now use this basis to represent the vector $(1, 2, 3, 4, 5, 6)'$. We write

$$x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ -0 \\ 0 \\ -1 \end{pmatrix} + x_4 \begin{pmatrix} 1 \\ 1 \\ -2 \\ -2 \\ 1 \\ 1 \end{pmatrix} + x_5 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + x_6 \begin{pmatrix} -1 \\ 0 \\ 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix}. \tag{4.1}$$

We now act on both sides of this equation with the matrix F_6 to get

$$x_4 \begin{pmatrix} -.5 \\ -.5 \\ 1 \\ 1 \\ .5 \\ .5 \end{pmatrix} + x_5 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + x_6 \begin{pmatrix} -1 \\ -1 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2.5 \\ 2.5 \\ 3.5 \\ 3.5 \\ 4.5 \\ 4.5 \end{pmatrix}. \tag{4.2}$$

Some arithmetic shows that $x_4 = 0$, $x_5 = 3.5$ and $x_6 = 1$. Thus we have that

$$\frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix} = 3.5 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} -1 \\ -1 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$

After more calculation we that

$$\lim_{n\to\infty} \frac{1}{2}^n \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}^n \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix} = 3.5 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

We now calculate what the expected value of the uniform distribution would be. There are 5! vectors that have 1 in the first position. Thus the sum of all 6! vectors is given by $5!6(6+1)/2(1,1,1,1,1,1)'$ and hence the expected value is $(3.5, 3.5, 3.5, 3.5, 3.5)'$. Thus the expected value converges to the expected value of the uniform distribution. This does not guarantee that the system given by equation 2.1 results in a uniform distribution on the entire state space but it does not rule it out. So in order to determine if the system is reachable we still would have to do the algebra.

Another property of a stochastic system is the concept of hitting time. Many of the stopping time problems can be reduced to hitting times. In general what we are studying is the mixing of the deck and stopping when the deck is well mixed. An important treatment of mixing is given in [8] and [9]. Given an initial point and a large subset of the state space, we ask for the expected time for an orbit to intersect the subset. For the system we will consider in the next section, the subset is the set of all permutations with 1 as the first element. This set has cardinality $(n-1)!$ and we will show that expected time to hit this set is governed by a geometric distribution and the limiting distribution as the deck size grows is exponential. On the other hand the question that we ask here is how long do we expect to operate before position one is filled by the first entry in the deck is an example of a strong uniform time.

5. Experiment and analysis

In this section, we will discuss the following questions:

(1) For a deck of cards, how many times of top-in shuffle should be performed before the top card goes back to the original position?
(2) To what distribution does the hitting time of the top-in shuffling conform?

We simulate top-in shuffling for 6 cards, 12 cards and extend the simulation to a standard deck of 54 cards, obtaining conjectures fore the distribution of the hitting time of the top-in shuffle.

5.1. The top-in shuffling[1]

If we consider shuffling 6 cards with top-in shuffle, using matrices to describe this process. A permutation can be represented by its incidence matrix.

$$A_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

If we take a card from the top and insert it at the kth position in the deck, it corresponds to A_k, since in this example there are 5 other place

[1] Part of the remaining material is taken from [10].

could be chosen for the top card. Hence, we will get the following matrices.

$$A_2 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$A_3 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$A_5 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad A_6 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

For example, if we insert the top card in the 4th position, it corresponds to the matrix A_4 above.

We assume that the cards are represented by an incidence vector with initial vector

$$X_0 = (1, 2, 3, 4, 5, 6)'$$

and then the act of shuffling can be represented in the form

$$X_1 = A_i X_0$$

where the A_is are as above.

After determining the incidence matrices for top-in shuffle of 6 cards, now consider shuffling two or even more times consecutively, which can be represented as a random product of the 6 matrices, i.e. $X_2 = A_j A_i X_0$. Note that since the 6 matrices are from a group the X_i is simply a permutation of the ordering in X_0. The question that we will answer and examine in this paper is what is the expected length of the random product until a vector of the form $X = (1, i_2, \ldots, i_n)'$ is achieved, i.e. how long does it take for the top card to return to the top.

If we use X_n to denote the corresponding vector after the nth shuffle then we will get the relationship between X_{n+1} and X_n as

$$X_{n+1} = (\delta_1(n)A_1 + \delta_2(n)A_2 + \cdots + \delta_6(n)A_6)X_n \tag{5.1}$$

with

$$\begin{aligned} \delta_i(n) &\in \{0,1\} \\ \sum_{i=1}^{6} \delta_i(n) &= 1 \\ P(\delta_i(n) = 1) &= 1/6 \end{aligned}$$

That is the process can be represented as a linear stochastic switching system.

5.2. Simulation

In this section, we still use 6, 12 and 54 cards as examples, Using 6 cards and doing complete shuffles 1000 times, we get Table (a). The data sets have been moved to the appendix. The odd numbered column in the table (a) represents the number of shuffles needed to move the top card back to the top position. The even column stands for how many times this situation will happen in the 1000 complete shuffles. For example, if we look at the line (6,54|31,0) which means in the 1000 complete shuffles there are 54 times after we took 6 shuffles to bring the top card back to the top and 0 times after 31 shuffles. Meanwhile we realize with the growing first column, the second column contains more and more 0s. As a result, truncating insignificant trailing zeroes we can plot the histogram in Figure 1 and calculate the mean and standard deviation for this data set. The C++ code for 6 cards is contained in the appendix.

We repeat the above process for 12 cards to get the data in Table (b) and shuffle 10,000 times to obtain the histogram in Figure 2. Similarly, we also extend it to the entire standard deck of cards, (52 cards plus two jokers) for which we need many more shuffles to get the approximate distribution. So we do the top-in shuffling 100,000 times (see Table (c) and obtain histogram graph (see Figure 3).

The statistics from these experiments and the figures, suggest that the number of shuffles to return 1 to the top spot satisfies the exponential distribution. In the following we will prove this.

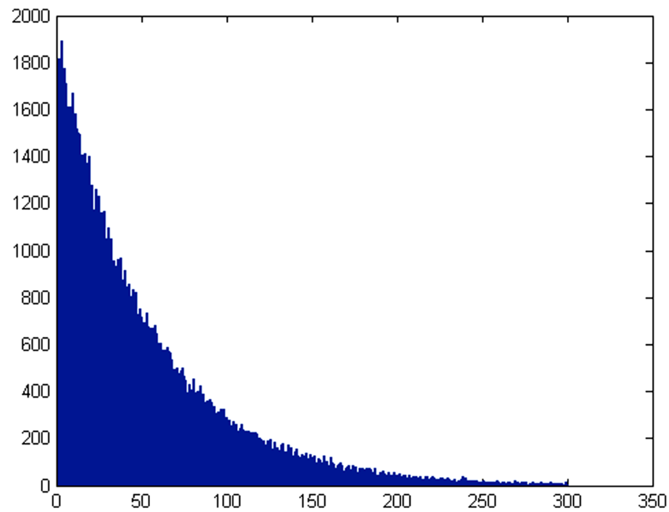

Figure 1: Steps needed for shuffling 6 cards when the top card goes back to original position. Mean = 6.02, Standard deviation: 5.55.

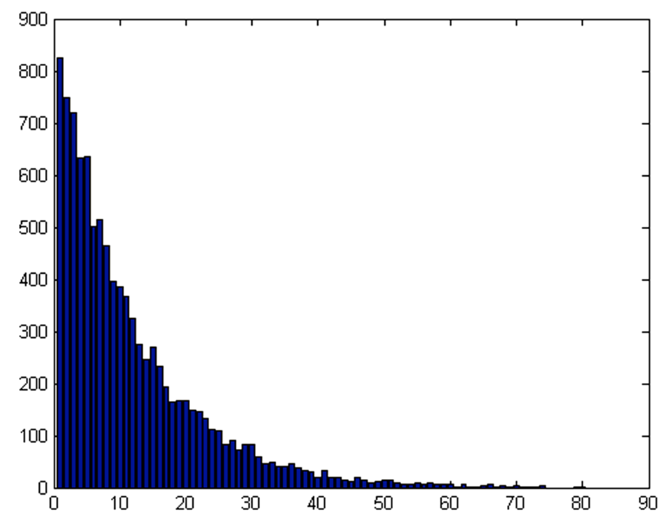

Figure 2: Steps needed for shuffling 12 cards when the top card goes back to original position. Mean; 11.97, Standard deviation; 11.48.

6. Exponential distribution

In this section, we prove some results about the distribution of the number of shuffles to make the top card back to top again. These results are indicated strongly by our experiments.

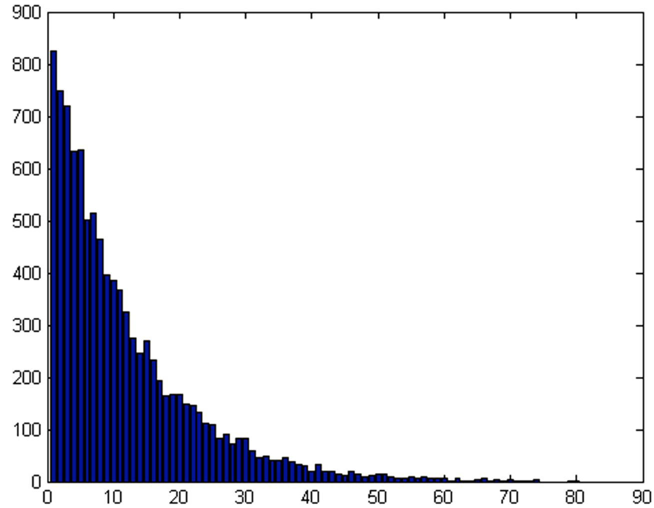

Figure 3: Steps needed for shuffling 54 cards when the top card goes back to original position. Mean = 53.93, Standard deviation = 53.47.

Theorem 6.1. *Let x be the number of shuffles to make the top card back to top again the first time. For a deck of k cards, the distribution of x is*

$$p(x) = \begin{cases} \frac{1}{k}(\frac{k-1}{k})^{x-1}, & x \geq 1 \\ 0, & x < 0. \end{cases} \tag{6.1}$$

Proof. Clearly $x \geq 1$, so the distribution is 0 for $x < 1$.

Let x_i be the position of the original top card after i shuffles before it returns back to the top. We claim that

$$P(x_i = j) = \frac{1}{k}\left(\frac{k-1}{k}\right)^{i-1} \quad \text{for all } j = 2, 3, \ldots, k.$$

We use induction on i. It is true for $i = 1$ because the probability of top card going to any position after one shuffle is $1/k$. Assume it is true for i. Then

$$P(x_{i+1} = j) = \sum_{m=1}^{k} P(x_{i+1} = j | x_i = m) P(x_i = m).$$

Notice that

$$P(x_{i+1} = j | x_i = m) = 0 \quad \text{for all} \quad m \neq j, j+1,$$

and since there are $j-1$ positions above the jth card and $k-j+1$ positions below the jth card, we have

$$P(x_{i+1}=j|x_i=j)=\frac{j-1}{k} \quad \text{and}$$
$$P(x_{i+1}=j|x_i=j+1)=\frac{k-(j+1)+1}{k}=\frac{k-j}{k}.$$

So

$$\begin{aligned}
P(x_{i+1}=j) &= P(x_{i+1}=j|x_i=j)P(x_i=j) \\
&\quad + P(x_{i+1}=j|x_i=j+1)P(x_i=j+1) \\
&= \frac{j-1}{k}\left(\frac{1}{k}\right)\left(\frac{k-1}{k}\right)^{i-1}+\frac{k-j}{k}\left(\frac{1}{k}\right)\left(\frac{k-1}{k}\right)^{i-1} \\
&= \frac{1}{k}\left(\frac{k-1}{k}\right)^{i}, \quad \text{for all } j=2,\ldots,k-1
\end{aligned}$$

and

$$\begin{aligned}
P(x_{i+1}=k) &= P(x_{i+1}=k|x_i=k)P(x_i=k) \\
&= \frac{k-1}{k}\left(\frac{1}{k}\right)\left(\frac{k-1}{k}\right)^{i-1}=\frac{1}{k}\left(\frac{k-1}{k}\right)^{i}.
\end{aligned}$$

Now we are ready to prove the result. Clearly for $j=1$

$$p(1)=P(x=1)=\frac{1}{k}.$$

For every $j\geq 2$

$$\begin{aligned}
p(j)=P(x=j) &= P(x=j|x_{j-1}=2)P(x_{j-1}=2) \\
&= \left(\frac{k-1}{k}\right)\left(\frac{1}{k}\left(\frac{k-1}{k}\right)^{j-2}\right) \\
&= \frac{1}{k}\left(\frac{k-1}{k}\right)^{j-1} \\
&= \frac{1}{k}\left(\frac{k-1}{k}\right)^{x-1}.
\end{aligned}$$

□

Theorem 6.2. *Let x be the number of shuffles to make the top card returns to the top again the first time. Then for a deck of k cards,*

i) the mean of x is $E(x) = k$;
ii) the m-th moment $E(x^m)$ satisfies

$$E(x^m) = k^2 \frac{d}{dk}\left(\frac{k-1}{k}E(x^{m-1})\right) + (k-1)E(x^{m-1}) \tag{6.2}$$

and

$$E(x^m) = m!k^m + O(k^{m-1}); \tag{6.3}$$

iii) the standard deviation of x is $\sigma = \sqrt{k^2 - k}$.

Proof. i) The mean of x is given by

$$\begin{aligned} E(x) &= \sum_{i=1}^{\infty} ip(i) = \sum_{i=1}^{\infty} \frac{i}{k}\left(\frac{k-1}{k}\right)^{i-1} \\ &= k\frac{d}{dk}\sum_{i=1}^{\infty}\left(\frac{k-1}{k}\right)^{i} = k\frac{d}{dk}(k-1) = k. \end{aligned}$$

ii)

$$\begin{aligned} & k^2\frac{d}{dk}\left(\frac{k-1}{k}E(x^{m-1})\right) + (k-1)E(x^{m-1}) \\ = {} & k^2\frac{d}{dk}\sum_{i=1}^{\infty}\frac{i^{m-1}}{k}\left(1-\frac{1}{k}\right)^{i} + (k-1)\sum_{i=1}^{\infty}\frac{i^{m-1}}{k}\left(1-\frac{1}{k}\right)^{i-1} \\ = {} & k^2\left(\sum_{i=1}^{\infty}\frac{-i^{m-1}}{k^2}\left(1-\frac{1}{k}\right)^{i} + \sum_{i=1}^{\infty}\frac{i^m}{k^3}\left(1-\frac{1}{k}\right)^{i-1}\right) \\ & + (k-1)\sum_{i=1}^{\infty}\frac{i^{m-1}}{k}\left(1-\frac{1}{k}\right)^{i-1} \\ = {} & \sum_{i=1}^{\infty}\frac{-i^{m-1}(k-1)}{k}\left(1-\frac{1}{k}\right)^{i-1} + \sum_{i=1}^{\infty}\frac{i^m}{k}\left(1-\frac{1}{k}\right)^{i-1} \\ & + (k-1)\sum_{i=1}^{\infty}\frac{i^{m-1}}{k}\left(1-\frac{1}{k}\right)^{i-1} \\ = {} & \sum_{i=1}^{\infty}\frac{i^m}{k}\left(1-\frac{1}{k}\right)^{i-1} = E(x^m) \quad \text{for } m = 2, 3, \ldots. \end{aligned}$$

To prove $E(x^m) = m!k^m + O(k^{m-1})$, we use induction on m. It is true for $m = 1$. Assume it is true for $m-1$, $E(x^{m-1}) = (m-1)!k^{m-1} + O(k^{m-2})$, then

$$\begin{aligned} E(x^m) &= k^2\frac{d}{dk}((m-1)!k^{m-1} + O(k^{m-2})) \\ &\quad + (k-1)((m-1)!k^{m-1} + O(k^{m-2})) \\ &= k^2((m-1)(m-1)!k^{m-2} + O(k^{m-3})) \\ &\quad + (m-1)!k^m + O(k^{m-1}) \\ &= (m-1)(m-1)!k^m + (m-1)!k^m + O(k^{m-1}) \\ &= m!k^m + O(k^{m-1}). \end{aligned}$$

iii) By ii)

$$E(x^2) = k^2\frac{d}{dk}\left(\frac{k-1}{k}k\right) + (k-1)k = 2k^2 - k,$$

and

$$\sigma = \sqrt{E(x^2) - E(x)^2} = \sqrt{k^2 - k}.$$

□

Theorem 6.3. *As the number of cards k becomes large, the distribution $p(x)$ asymptotically approaches to exponential distribution that*

$$f(x) = \begin{cases} \frac{1}{k}\exp\left(-\frac{1}{k}(x-1)\right), & x \geq 1 \\ 0, & x < 1. \end{cases} \tag{6.4}$$

for each fixed x.

Proof. Since

$$p(x) = \begin{cases} \frac{1}{k}(\frac{k-1}{k})^{x-1}, & x \geq 1 \\ 0, & x < 0. \end{cases} \tag{6.5}$$

Note that

$$\frac{1}{k}\left(\frac{k-1}{k}\right)^{x-1} = \frac{1}{k}\exp\left((x-1)\log\left(1-\frac{1}{k}\right)\right).$$

Therefore

$$p(x) = \begin{cases} \frac{1}{k}\exp\left((x-1)\log\left(1-\frac{1}{k}\right)\right), & x \geq 1 \\ 0, & x < 0. \end{cases}$$

Since

$$\log\left(1-\frac{1}{k}\right) = -\frac{1}{k} + O\left(\frac{1}{k^2}\right),$$

So for each fixed x, $p(x)$ approaches to

$$f(x) = \begin{cases} \frac{1}{k}\exp\left(-\frac{1}{k}(x-1)\right), & x \geq 1 \\ 0, & x < 1, \end{cases}$$

as k becomes large. □

7. Distribution

We begin by calculating the expected value for shuffling of the 6 card deck in the example. The expected value is given by the

$$E[x_{n+1}] = FE[x_n]$$

where F is given by the weighted sum of the matrices.

$$F = \frac{1}{6}\begin{pmatrix} 1 & 5 & 0 & 0 & 0 & 0 \\ 1 & 1 & 4 & 0 & 0 & 0 \\ 1 & 0 & 2 & 3 & 0 & 0 \\ 1 & 0 & 0 & 3 & 2 & 0 \\ 1 & 0 & 0 & 0 & 4 & 1 \\ 1 & 0 & 0 & 0 & 0 & 5 \end{pmatrix}. \tag{7.1}$$

It is easy to see that there are eigenvalues $1, \frac{2}{3}$ and 0 but it is not so obvious that there are eigenvalues $\frac{1}{6}$, $\frac{1}{3}$ and $\frac{1}{2}$. These are discovered by taking advantage of the structure of F which make the characteristic polynomial fairly easy to evaluate. Since there is a dominate eigenvalue the system will converge to the final distribution. It is best to start with the unit vector $(1,2,3,4,5,6)'$. If we had the eigenvectors of F then the we could expand the initial vector in terms of the eigenvectors and the limit would be given by the coefficient of the eigenvector $(1,1,1,1,1,1)$. The structure of the eigenvectors is not immediately obvious.

$$F_6^{10}\begin{bmatrix}1\\2\\3\\4\\5\\6\end{bmatrix}=\begin{bmatrix}3.4566\\3.4740\\3.4913\\3.5087\\3.5260\\3.5434\end{bmatrix},\quad F_6^{20}\begin{bmatrix}1\\2\\3\\4\\5\\6\end{bmatrix}=\begin{bmatrix}3.4992\\3.4995\\3.4998\\3.5002\\3.5005\\3.5008\end{bmatrix},\quad \text{and}$$

$$F_6^{30}\begin{bmatrix}1\\2\\3\\4\\5\\6\end{bmatrix}=\begin{bmatrix}3.5000\\3.5000\\3.5000\\3.5000\\3.5000\\3.5000\end{bmatrix}.$$

Generally for a deck of n cards we have the following results:

Theorem 7.1. *The eigenpairs of F are*

$$(1,\beta_n),\quad \left(0,\begin{bmatrix}-n+1\\\beta_{n-1}\end{bmatrix}\right),\quad \left(\frac{k}{n},\begin{bmatrix}k-n+1\\\alpha_k\\\beta_{n-k-1}\end{bmatrix}\right)\qquad \text{for } k=1,\dots,n-2$$

where β_i is the i-vector with all entries equal 1 and

$$\alpha_k=\begin{bmatrix}a_{k1}\\\vdots\\a_{kk}\end{bmatrix}\quad \text{with}\quad a_{k1}=\frac{(1-k)(n-k-1)}{n-1}\quad \text{and}$$

$$a_{ki}=\frac{n-k-1+(k+1-i)a_{i-1}}{n-i},\quad \text{for}\quad i=2,\dots,k.$$

Proof. The eigenvectors associated with eigenvalues 1 and 0 are obvious. For each $k=1,\dots,n-2$, write

$$F=\frac{1}{n}\begin{bmatrix}1 & n-1 & 0 & \cdots & \cdots & 0\\ 1 & 1 & n-2 & \ddots & & \vdots\\ \vdots & 0 & 2 & \ddots & \ddots & \vdots\\ \vdots & \vdots & \ddots & \ddots & 2 & 0\\ \vdots & \vdots & & \ddots & n-2 & 1\\ 1 & 0 & \cdots & \cdots & 0 & n-1\end{bmatrix}=\frac{1}{n}\begin{bmatrix}\beta_k & A_k & 0\\ 1 & \gamma_k & \delta_k\\ \beta_{n-k-1} & 0 & B_k\end{bmatrix}$$

where $A_k, B_k, \gamma_k, \delta_k$ are $k \times k, (n-k-1)\times(n-k-1), 1\times k, 1\times(n-k-1)$ matrices, respectively, given by

$$A_k = \begin{bmatrix} n-1 & 0 & \cdots & 0 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & k-1 & n-k \end{bmatrix}, \; B_k = \begin{bmatrix} k+1 & n-k-2 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & n-1 \end{bmatrix},$$

$\gamma_k = [0, \ldots, 0, k], \quad \delta_k = [n-k-1, 0, \ldots, 0].$

Let

$$\frac{1}{n} \begin{bmatrix} \beta_k & A_k & 0 \\ 1 & \gamma_k & \delta_k \\ \beta_{n-k-1} & 0 & B_k \end{bmatrix} \begin{bmatrix} k-n+1 \\ \alpha_k \\ \beta_{n-k-1} \end{bmatrix} = \frac{k}{n} \begin{bmatrix} k-n+1 \\ \alpha_k \\ \beta_{n-k-1} \end{bmatrix}$$

we then have

$$(k-n+1)\beta_k + A_k\alpha_k = k\left((k-n+1)e_1 + E_k\alpha_k\right), \tag{7.2}$$

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad E_k = \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix}$$

$$k-n+1+\gamma_k\alpha_k + \delta_k\beta_{n-k-1} = ka_{kk}, \tag{7.3}$$

$$(k-n+1)\beta_{n-k-1} + B_k\beta_{n-k-1} = k\beta_{n-k-1}. \tag{7.4}$$

(7.4) is clearly satisfied because $(n-1, \beta_{n-k-1})$ is an eigenpair of B_k.
(7.3) is an identity

$$k-n+1+ka_{kk}+n-k-1 = ka_{kk}.$$

(7.2) is equivalent to $(A_k - kE_k)\alpha_k = -k(n-k-1)e_1 + (n-k-1)\beta_k$, or

$$\begin{bmatrix} n-1 & 0 & \cdots & 0 \\ -k+1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & -1 & n-k \end{bmatrix} \alpha_k = \begin{bmatrix} (1-k)(n-k-1) \\ n-k-1 \\ \vdots \\ n-k-1 \end{bmatrix}.$$

The iteration formula for the entries of α_k follows immediately. $\square$

Theorem 7.2. *The expected value $E[x_m]$ approaches to that of the uniform distribution as $m \to \infty$, i.e.*

$$\lim_{m\to\infty} F^m \begin{bmatrix} 1 \\ \vdots \\ n \end{bmatrix} = \frac{n+1}{2} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}.$$

Proof. Let v_k be the eigenvector of F associated with the eigenvalue k/n, $k = 0, 1, \ldots, n-2, n$, and let $x = [1, 2, \ldots, n]'$. Then $x = c_0 v_0 + \cdots + c_{n-2} v_{n-2} + c_n v_n$ for some $c_0, \ldots, c_{n-2}, c_n$ and

$$\lim_{m\to\infty} F^n x = \lim_{m\to\infty} \left(c_1 \left(\frac{1}{n}\right)^m v_1 + \cdots c_{n-2} \left(\frac{n-2}{n}\right)^m v_{n-1} + c_n v_n \right) = c_n v_n.$$

So we only need to show that $c_n = (n+1)/2$. Note that $(1, v_n)$ is also an eigenpair for F'. So $\langle v_n, v_k \rangle = \langle F' v_n, v_k \rangle = \langle v_n, F v_k \rangle = \frac{k}{n} \langle v_n, v_k \rangle$, and it implies that $\langle v_n, v_k \rangle = 0$ for $k = 0, 1, \ldots, n-2$. Therefore $\langle v_n, x \rangle = c_0 \langle v_n, v_0 \rangle + \cdots c_{n-2} \langle v_n, v_{n-2} \rangle + c_n \langle v_n, v_n \rangle = c_n \langle v_n, v_n \rangle$, and

$$c_n = \frac{\langle x, v_n \rangle}{\langle v_n, v_n \rangle} = \frac{1 + 2 + \cdots + n}{n} = \frac{n+1}{2}. \qquad \square$$

It is interesting to note that the largest eigenvalue less than 1 is given by $\frac{n-2}{n}$ and so we expect the system to converge at the rate

$$\left(\frac{n-2}{n}\right)^n.$$

Note that the explicit calculation for F_6 confirms this.

8. Conclusion

In this paper we have begun the process of using discrete time stochastic linear switching systems to study the mathematics of card shuffling. We have shown that any shuffling scheme can be represented by such a system. In future work we will extend this to other more complicated shuffles and will explore some of the tree based schemes proposed in [3] and [4].

Appendix 1. Data sets

(a) The top-in shuffle with 6 cards

1	162	26	1
2	141	27	2
3	123	28	1
4	97	29	1
5	81	30	0
6	54	31	0
7	57	32	1
8	58	33	1
9	35	34	1
10	27	35	1
11	22	36	0
12	23	37	1
13	27	38	0
14	13	39	0
15	17	40	0
16	15	41	0
17	3	42	0
18	11	43	0
19	4	44	1
20	3	45	0
21	4	46	0
22	1	47	0
23	6	48	
24	3	49	
25	2	...	...

(b) The top-in shuffle with 12 cards

1	824	29	84	57	8	85	0	113	0	141	0
2	748	30	83	58	7	86	0	114	0	142	0
3	721	31	58	59	7	87	0	115	0	143	0
4	634	32	47	60	7	88	0	116	0	144	0
5	635	33	49	61	1	89	0	117	0	145	0
6	501	34	41	62	6	90	0	118	0	146	0
7	515	35	40	63	2	91	0	119	0	147	0
8	465	36	45	64	1	92	0	120	0	148	0
9	397	37	38	65	5	93	0	121	0	149	0
10	386	38	33	66	7	94	0	122	0	150	0
11	366	39	31	67	2	95	0	123	0	...	...
12	325	40	21	68	3	96	0	124	0		
13	276	41	32	69	2	97	1	125	0		
14	246	42	19	70	4	98	0	126	0		
15	270	43	20	71	1	99	1	127	0		
16	233	44	15	72	1	100	0	128	0		
17	193	45	11	73	2	101	0	129	1		
18	164	46	21	74	3	102	0	130	0		
19	166	47	15	75	0	103	0	131	0		
20	166	48	10	76	0	104	0	132	0		
21	150	49	13	77	0	105	1	133	0		
22	145	50	15	78	0	106	1	134	0		
23	132	51	15	79	2	107	0	135	0		
24	113	52	10	80	1	108	0	136	0		
25	110	53	7	81	1	109	0	137	0		
26	82	54	6	82	2	110	0	138	0		
27	90	55	9	83	0	111	0	139	0		
28	72	56	7	84	0	112	0	140	0		

(c) The top-in shuffle with 54 cards

1	1773	29	1051	57	672	85	425	113	232	141	153	169	59
2	1816	30	1025	58	679	86	388	114	225	142	128	170	75
3	1889	31	1094	59	645	87	348	115	226	143	118	171	80
4	1737	32	1052	60	607	88	355	116	218	144	131	172	85
5	1775	33	954	61	606	89	359	117	226	145	122	173	66
6	1712	34	957	62	576	90	365	118	220	146	124	174	85
7	1610	35	933	63	557	91	353	119	201	147	140	175	76
8	1599	36	961	64	575	92	335	120	196	148	116	176	78
9	1613	37	964	65	588	93	334	121	190	149	132	177	56
10	1668	38	968	66	572	94	308	122	191	150	118	178	71
11	1584	39	873	67	563	95	313	123	170	151	118	179	72
12	1515	40	913	68	533	96	284	124	192	152	124	180	71
13	1501	41	845	69	494	97	324	125	191	153	97	181	69
14	1493	42	842	70	458	98	326	126	197	154	113	182	71
15	1404	43	854	71	499	99	248	127	154	155	107	183	71
16	1370	44	807	72	474	100	290	128	186	156	94	184	72
17	1412	45	833	73	487	101	259	129	161	157	124	185	67
18	1372	46	731	74	500	102	279	130	163	158	104	186	56
19	1402	47	820	75	462	103	255	131	149	159	105	187	74
20	1265	48	727	76	450	104	272	132	174	160	85	188	41
21	1280	49	752	77	397	105	261	133	181	161	120	189	45
22	1170	50	715	78	427	106	259	134	145	162	98	190	55
23	1260	51	686	79	409	107	229	135	144	163	89	...	...
24	1179	52	694	80	409	108	243	136	175	164	74		
25	1232	53	731	81	454	109	263	137	139	165	83		
26	1132	54	675	82	392	110	238	138	159	166	90		
27	1162	55	668	83	401	111	229	139	124	167	97		
28	1166	56	667	84	387	112	220	140	144	168	85		

Appendix 2. C++ code for 6 cards

```
#include<iostream.h>
#include<string.h>
#include <time.h>
#include <stdlib.h>
#include <iomanip.h>
#include <fstream.h>
class CardList{
    public:
        int cardNode[6] ;
    public:
        void initial(){
```

```
        for(int i=0;i<6;i++){
                cardNode[i]=i+1;
            }
        }
    void nreshuffle(int n){
                int temp = cardNode[n-1];
                if (n>1){
                    for(int j=n-1;j>=1;j--){
                        cardNode[j]=cardNode[j-1];
                    }
                    cardNode[0]=temp;
                }
        }
        void allPrint(){
                for(int k=0;k<54;k++){
                        cout<<cardNode[k];
                        cout<<" ";
                }
        }
};
void main(){
    CardList cardList;
    cardList.initial();
    cout<<"the original order is:  ";
    cardList.allPrint();
    cout<<"\n";
    int steps[100002];
    for(int k=0;k<100002;k++){
        steps[k]=0;
    }
        // compute all the steps
    int sum=1000001;
    srand(unsigned (time(0)));
    for (int i=1;i<sum;i++){
            int tempstep=0;
        while(true){
                int auton = rand()%12+1;
            cardList.nreshuffle(auton);
            tempstep++;
      if (cardList.cardNode[11]==1){
                if (tempstep<=100000){steps[tempstep]++;}
                else {steps[100001]++;}
            cardList.initial();
                break;
                }
        }
    }
    // open an file to record
    ofstream f1("test.txt");
```

```
    if(!f1) cout<<"no test.txt file";
    double average=0;
    for(int j=1;j<100002;j++){
        cout<<j<<" steps occur "<<steps[j]<<" times";
        f1<<j<<" steps occur "<<steps[j]<<" times"<<endl;
        cout<<"\n";
        average+=j*steps[j];
    }
    for( j=1;j<100002;j++){
        f1<<j<<endl;
    }
    for( j=1;j<100002;j++){
        f1<<steps[j]<<endl;
    }
    average=average/sum;
    cout<<"the average step is "<<average<<"  steps";
    cout<<"\n";
    f1<<"the average step is "<<average<<" steps"<<endl;
    f1.close();
}
```

References

[1] D. Aldous and P. Diaconis, Strong uniform times and finite random walks, *Advances in Applied Mathematics* **8**(1) (1987), 69–97. MR0876954

[2] P. Diaconis, Group representations in probability and statistics. Institute of Mathematical Statistics, Lecture Notes–Monograph Series, 11. Institute of Mathematical Statistics, Hayward, CA, 1988. MR0964069

[3] P. Diaconis and L. Saloff-Conte, Comparison techniques for random walk on finite groups, *The annals of Probability* **21** (1993), 2131–2156. MR1245303

[4] P. Diaconis and M. Shahshanani, Generating a random permutation with random transpositions, *Z. Wahrsch. Verw. Gebiete* **57** (1981), 159–179. MR0626813

[5] Y. Du and C. Martin, Switched systems as models for dynamic clinical trials, to appear, *Communications in Information and Systems.*

[6] B. Hanlon, V. Tyuryaev, and C. Martin, Stability of switched linear systems with Poisson switching, *Communications in Information and Systems* **11**(4) (2011), 307–326. MR2950833

[7] B. Hanlon, N. Wang, M. Egerstedt, and C. Martin, Switched linear systems: Stability and the convergence of random products, *Communications in Information and Systems* **11**(4) (2011), 327–342. MR2950834

[8] L. Lovasz and P. Winkler, Mixing of random walks and other diffusions on a graph, Surveys in Combinatorics, 1995, P. Rowlinson ed., London Math. Soc. Lecture, Note Series, 218, 1995, pp. 119–154. MR1358634

[9] L. Lovasz and P. Winkler, Mixing times, Microsurveys in Discrete Probability, D. Aldous and J. Propp, eds, DIMACS Series in Discrete Math. and Theoretical Computer Science 41, Amer. Math. Soc., Providence RI, 1998, pp. 85–134. MR1630411

[10] Y. Pang, Random walk on a finite group, Master's Thesis, Texas Tech University, 2012.

Yulei Pang
Department of Mathematics and Statistics
Texas Tech University
USA
E-mail address: yulei.pang@ttu.edu

Alex Wang
Department of Mathematics and Statistics
Texas Tech University
USA
E-mail address: alex.wang@ttu.edu

Xiaozhen Xue
Department of Computer Science
Texas Tech University
USA
E-mail address: xiaozhen.xue@ttu.edu

Clyde F. Martin
Department of Mathematics and Statistics
Texas Tech University
USA
E-mail address: clyde.f.martin@ttu.edu

Received February 18, 2013