Key Independent Bias in the Permutation after RC4 Key Scheduling

Goutam Paul[1], Subhamoy Maitra[2], Rohit Srivastava[3]

[1] Department of Computer Science and Engineering, Jadavpur University,
Kolkata 700 032, India.
goutam_paul@cse.jdvu.ac.in

[2] Applied Statistics Unit, Indian Statistical Institute,
203, B T Road, Kolkata 700 108, India.
subho@isical.ac.in

[3] Department of Computer Science and Engineering, Institute of Technology,
Banaras Hindu University, Varanasi 221 005 (UP), India.
rohit.engg@gmail.com

Abstract. In this paper, we present a weakness of the RC4 Key Scheduling Algorithm (KSA). Consider the RC4 permutation S of N (usually 256) bytes and denote it by S_N after the KSA. We observe for the first time and then theoretically prove that each permutation byte after the KSA is significantly biased (either positive or negative) towards many values in the range $0, \ldots, N-1$. For each byte u, $0 \leq u \leq N-2$, $P(S_N[u] = v)$ is maximum at $v = u + 1$ and this maximum probability ranges approximately between $\frac{1}{N}(1+\frac{1}{3})$ and $\frac{1}{N}(1+\frac{1}{5})$ for different values of u. Moreover, these biases are independent of the secret key and thus presents an evidence that the permutation after the KSA can be distinguished from random permutation without any assumption on the secret key.

Keywords: Bias, Cryptography, Cryptanalysis, Key Scheduling Algorithm, RC4, Stream Cipher.

1 Introduction

RC4, one of the most popular stream ciphers till date, was proposed by Rivest in 1987. The cipher gained its popularity from its extremely simple structure and substantially good strength in security, as even after lots of explored weaknesses in the literature [1–16], it could not be thoroughly cracked. Studying weaknesses of RC4 received serious attention in the literature and these studies are believed to be quite useful in further development of stream ciphers that exploit shuffle-exchange paradigm.

Before getting into our contribution, let us briefly present the Key Scheduling Algorithm (KSA) and the Pseudo Random Generation Algorithm (PRGA) of RC4. The data structure contains an array of size N (in practice 256 which is followed in this paper) with each location having an integer in the range

$[0, \ldots, N-1]$, two indices i, j and the secret key array K. Given a secret key k of l bytes (typically 5 to 32), the array K of size N is such that $K[i] = k[i \bmod l]$ for any i, $0 \leq i \leq N-1$. All additions used in the description of the algorithm are modulo N additions.

```
Algorithm KSA
Initialization:
        For $i = 0, \ldots, N-1$
            $S[i] = i$;
        $j = 0$;
Scrambling:
        For $i = 0, \ldots, N-1$
            $j = (j + S[i] + K[i])$;
            Swap($S[i], S[j]$);
```

```
Algorithm PRGA
Initialization:
        $i = j = 0$;
Output Keystream Generation Loop:
            $i = i + 1$;
            $j = j + S[i]$;
            Swap($S[i]$, $S[j]$);
            $t = S[i] + S[j]$;
            Output $z = S[t]$;
```

RC4 KSA has been analysed deeply in [15,16,3,12]. All these works discuss the relationship of the permutation bytes after the KSA with the secret key. For a proper design, the permutation S after the KSA should not have any correlation with the secret keys. However, weaknesses of RC4 in this aspect have already been reported [15,16,3,12]. These weaknesses, in turn, leak information about RC4 secret key in the initial keystream output bytes [11].

Another approach of study is to look at the permutation after the KSA in a (secret) key independent manner and try to distinguish it from random permutations. In [10], the sign of the permutation after the KSA has been studied (see [10] for the definition of the sign of a permutation). There it has been shown that, after the KSA, the sign of the permutation can be guessed with probability 56%.

In Figure 1, we present a few graphs for $P(S_N[u] = v)$, against v, $0 \leq v \leq N-1$, for a few values of u as motivating examples. These graphs are based on 10 million trials over randomly chosen keys of 32 bytes. One may clearly note that the values of $P(S_N[u] = v)$ are not that of random association $\frac{1}{N}$ (probability that any two randomly chosen integers, with replacement, in the range $[0, \ldots, N-1]$ are equal).

Our main results related to the biases are presented in Section 2 (see Theorem 1 and Theorem 2). Numerical values of our theoretical formulae match with the experimental results except in very few places and that we discuss in Section 3. Numerical values from our formulae in Theorem 1 show that $P(S_N[u] = v)$ varies approximately in the range $\frac{1}{N}(1 + \frac{1}{3})$ at the higher side (i.e., positive bias) to $\frac{1}{N}(1 - \frac{1}{3})$ at the lower side (i.e., negative bias). It also attains maximum at $v = u + 1$, for each u in $[0, N-2]$, and $P(S_N[u] = u + 1)$ varies approximately from $\frac{1}{N}(1 + \frac{1}{3})$ to $\frac{1}{N}(1 + \frac{1}{5})$.

Note that the result of [10] on biased sign of RC4 permutations is based on the complete permutation S_N, but not on each individual byte of the permutation. Existing [15,16,3,11,12] key dependent biases of RC4 permutation after the KSA could be identified only for the initial bytes (0 to 47 for $N = 256$) of S_N. We identify the biases for all the bytes of the permutation after the KSA and each byte is biased to many values.

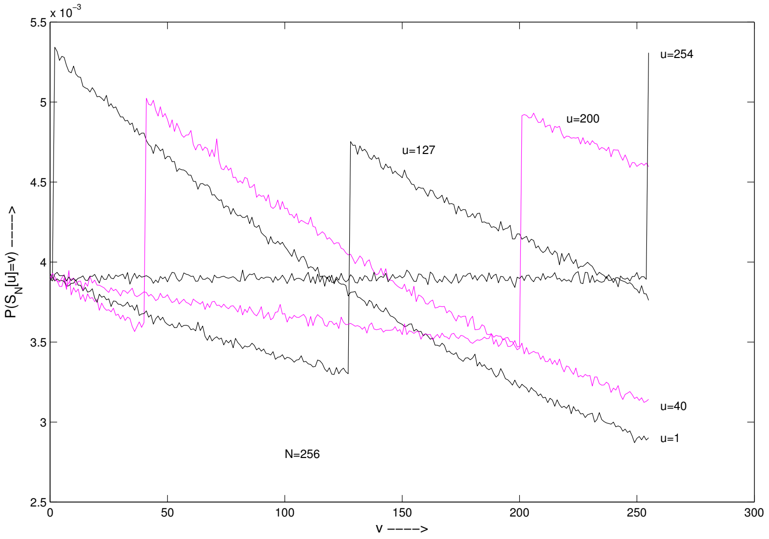

Fig. 1. $P(S_N[u] = v)$ versus v for some specific u's.

2 Bias in Each Permutation Byte

We denote the initial identity permutation by S_0 and the permutation at the end of the r-th round of KSA by S_r, $1 \leq r \leq N$ (note that round $r = i + 1$, for the deterministic index i, $0 \leq i \leq N - 1$). Thus, the permutation after the KSA will be denoted by S_N. By j_r, we denote the value of index j after it is updated in round r. In the proofs, we replace the joint probabilities with the product of the probabilities of the individual events, assuming that the events under consideration are statistically independent.

We start with the following technical result.

Lemma 1. $P(S_2[0] = 1) = \frac{2(N-1)}{N^2}$.

Proof. In the first round, we have $i = 0$, and $j_1 = 0 + S[0] + K[0] = K[0]$. In the second round, $i = 1$ and $j_2 = j_1 + S_1[1] + K[1]$. We consider two mutually exclusive and exhaustive cases, namely, $K[0] = 1$ and $K[0] \neq 1$.

1. Take $K[0] = 1$. So, after the first swap, $S_1[0] = 1$ and $S_1[1] = 0$. Now, $j_2 = K[0] + 0 + K[1] = K[0] + K[1]$. Thus, after the second swap, $S_2[0]$ will remain 1, if $K[0] + K[1] \neq 0$. Hence the contribution of this case to the event $(S_2[0] = 1)$ is $P(K[0] = 1) \cdot P(K[0] + K[1] \neq 0) = \frac{1}{N} \cdot \frac{N-1}{N} = \frac{N-1}{N^2}$.
2. Take $K[0] \neq 1$. Then after the first swap, $S_1[1]$ remains 1. Now, $j_2 = K[0] + 1 + K[1] = K[0] + K[1] + 1$. Thus, after the second swap, $S_2[0]$ will get the value 1, if $K[0] + K[1] + 1 = 0$. Hence the contribution of this case to the event $(S_2[0] = 1)$ is $P(K[0] \neq 1) \cdot P(K[0] + K[1] + 1 = 0) = \frac{N-1}{N} \cdot \frac{1}{N} = \frac{N-1}{N^2}$.

Adding the two contributions, we get the total probability as $\frac{2(N-1)}{N}$. □

We here calculate $P(S_2[0] = 1)$. Note that the form of $P(S_{v+1}[u] = v)$ for $v \geq u+1$ in general (see Lemma 2 later) does not work for the case $u = 0, v = 1$ only. This will be made clear in Remark 1 after the proof of Lemma 2.

Proposition 1. *$P(S_x[x] = x) = (\frac{N-1}{N})^x$, for $x \geq 0$.*

Proof. In the rounds 1 through x, the deterministic index i touches the permutation indices $0, 1, \ldots, x-1$. Thus, after round x, $S_x[x]$ will remain the same as $S_0[x] = x$, if x has not been equal to any of the x many pseudorandom indices $j_1, j_2, \ldots, j_x$. The probability of this event is $(\frac{N-1}{N})^x$. So the result holds for $x \geq 1$. Furthermore, $P(S_0[0] = 0) = 1 = (\frac{N-1}{N})^0$. Hence, for any $x \geq 0$, we have $P(S_x[x] = x) = (\frac{N-1}{N})^x$. □

Proposition 2. *For $v \geq u+1$, $P(S_v[u] = v) = \frac{1}{N} \cdot (\frac{N-1}{N})^{v-u-1}$.*

Proof. The permutation index u is touched by the deterministic index i for the first time in round $u+1$. After round $u+1$, the probability that $S_{u+1}[u] = v$ is $\frac{1}{N}$; as in the round $u+1$, the value of i is u and the location u of the permutation will be swapped with a random location based on the secret key values and the value j_{u+1}. The probability that the index u is not touched by any of the subsequent $v-u-1$ many j values, namely, $j_{u+2}, \ldots, j_v$, is given by $(\frac{N-1}{N})^{v-u-1}$. So, after the end of round v, $P(S_v[u] = v) = \frac{1}{N} \cdot (\frac{N-1}{N})^{v-u-1}$. □

Lemma 2. *For $v \geq u+1$ (except for the case "$u = 0$ and $v = 1$"), $P(S_{v+1}[u] = v) = \frac{1}{N} \cdot (\frac{N-1}{N})^{v-u} + \frac{1}{N} \cdot (\frac{N-1}{N})^v - \frac{1}{N^2} \cdot (\frac{N-1}{N})^{2v-u-1}$.*

Proof. In round $v+1$, $i = v$ and $j_{v+1} = j_v + S_v[v] + K[v]$. The event $(S_{v+1}[u] = v)$ can occur in two ways.

1. $S_v[u]$ already had the value v and the index u is not involved in the swap in round $v+1$.
2. $S_v[u] \neq v$ and the value v comes into the index u from the index v (i.e., $S_v[v] = v$) by the swap in round $v+1$.

From Proposition 1, we have $P(S_v[v] = v) = (\frac{N-1}{N})^v$ and from Proposition 2, we have $P(S_v[u] = v) = \frac{1}{N} \cdot (\frac{N-1}{N})^{v-u-1}$. Hence, $P(S_{v+1}[u] = v)$

$$= P(S_v[u] = v) \cdot P(j_v + S_v[v] + K[v] \neq u)$$
$$+ P(S_v[u] \neq v) \cdot P(S_v[v] = v) \cdot P(j_v + S_v[v] + K[v] = u)$$
$$\text{(except for the case "}u = 0 \text{ and } v = 1\text{", see Remark 1)}$$
$$= \left(\frac{1}{N} \cdot (\frac{N-1}{N})^{v-u-1}\right) \cdot (\frac{N-1}{N}) + \left(1 - \frac{1}{N} \cdot (\frac{N-1}{N})^{v-u-1}\right) \cdot (\frac{N-1}{N})^v \cdot \frac{1}{N}$$
$$= \frac{1}{N} \cdot (\frac{N-1}{N})^{v-u} + \frac{1}{N} \cdot (\frac{N-1}{N})^v - \frac{1}{N^2} \cdot (\frac{N-1}{N})^{2v-u-1}.$$
□

Remark 1. Case 1 in the proof of Lemma 2 applies to Lemma 1 also. In case 2, i.e., when $S_v[u] \neq v$, in general we may or may not have $S_v[v] = v$. However, for $u = 0$ and $v = 1$, $(S_1[0] \neq 1) \iff (S_1[1] = 1)$, the probability of each of which is $\frac{N-1}{N}$ (note that there has been only one swap involving the indices 0 and $K[0]$ in round 1). Hence the contribution of case 2 except for "$u = 0$ and $v = 1$" would be $P(S_v[u] \neq v) \cdot P(S_v[v] = v) \cdot P(j_v + S_v[v] + K[v] = u)$, and for "$u = 0$ and $v = 1$" it would be $P(S_1[0] \neq 1) \cdot P(j_1 + S_1[1] + K[1] = 0)$ or, equivalently, $P(S_1[1] = 1) \cdot P(j_1 + S_1[1] + K[1] = 0)$.

Lemma 3. *Let* $p_r^{u,v} = P(S_r[u] = v)$*, for* $1 \leq r \leq N$*. For any* $t > max\{u, v\}$*,* $P(S_N[u] = v) = p_t^{u,v} \cdot (\frac{N-1}{N})^{N-t} + (1 - p_t^{u,v}) \cdot \frac{1}{N}(\frac{N-1}{N})^v \cdot \Big(1 - (\frac{N-1}{N})^{N-t}\Big)$.

Proof. After round t ($> max\{u, v\}$), there may be two different cases: $S_t[u] = v$ and $S_t[u] \neq v$. Both of these can contribute to the event $(S_N[u] = v)$ in the following ways.

1. $S_t[u] = v$ and the index u is not touched by any of the subsequent $N - t$ many j values. The contribution of this part is $P(S_t[u] = v) \cdot (\frac{N-1}{N})^{N-t}$ $= p_t^{u,v} \cdot (\frac{N-1}{N})^{N-t}$.
2. $S_t[u] \neq v$ and for some i in the interval $[t, N-1]$, $S_i[i] = v$ which comes into the index u from the index i by the swap in round $i + 1$, and after that the index u is not touched by any of the subsequent $N - 1 - i$ many j values. So the contribution for this part is given by $P(S_t[u] \neq v) \cdot \Big(\sum_{i=t}^{N-1} P(S_i[i] = v) \cdot P(j_{i+1} = u) \cdot (\frac{N-1}{N})^{N-1-i}\Big)$. By Proposition 1, $P(S_v[v] = v) = (\frac{N-1}{N})^v$. Now, consider the swap in round $v + 1$. For any $x > v$, $P(S_{v+1}[x] = v) = P(S_v[v] = v) \cdot P(j_{v+1} = x) = (\frac{N-1}{N})^v \cdot \frac{1}{N}$. If $S_{v+1}[x] = v$, then during the rounds r, $v + 2 \leq r \leq x$, the value of $S_r[x]$ may be changed to some other value from v and also if $S_{v+1}[x] \neq v$, then during the rounds r, $v + 2 \leq r \leq x$, the value of $S_r[x]$ may be changed from some other value to v; we assume that these effects cancel out each other. Thus, for $x > v$, $P(S_x[x] = v) = P(S_{v+1}[x] = v) = \frac{1}{N}(\frac{N-1}{N})^v$. Since we are considering $i \geq t > max\{u, v\}$, we can write
$P(S_t[u] \neq v) \cdot \Big(\sum_{i=t}^{N-1} P(S_i[i] = v) \cdot P(j_{i+1} = u) \cdot (\frac{N-1}{N})^{N-1-i}\Big)$
$= (1 - p_t^{u,v}) \cdot \Big(\sum_{i=t}^{N-1} \frac{1}{N}(\frac{N-1}{N})^v \cdot \frac{1}{N} \cdot (\frac{N-1}{N})^{N-1-i}\Big)$
$= (1 - p_t^{u,v}) \cdot \frac{1}{N^2}(\frac{N-1}{N})^v \cdot \Big(\sum_{i=t}^{N-1} (\frac{N-1}{N})^{N-1-i}\Big)$
$= (1 - p_t^{u,v}) \cdot \frac{1}{N^2}(\frac{N-1}{N})^v \cdot \Big(\frac{1-a^{N-t}}{1-a}\Big)$, where $a = \frac{N-1}{N}$.
Substituting the value of a and simplifying, we get the above probability as $(1 - p_t^{u,v}) \cdot \frac{1}{N}(\frac{N-1}{N})^v \cdot \Big(1 - (\frac{N-1}{N})^{N-t}\Big)$.

Thus, combining the above two contributions, we get
$P(S_N[u] = v) = p_t^{u,v} \cdot (\frac{N-1}{N})^{N-t} + (1 - p_t^{u,v}) \cdot \frac{1}{N}(\frac{N-1}{N})^v \cdot \Big(1 - (\frac{N-1}{N})^{N-t}\Big)$. □

Theorem 1. *For* $u+1 \le v \le N-1$*,* $P(S_N[u]=v) = p_{v+1}^{u,v} \cdot (\frac{N-1}{N})^{N-1-v}$ $+ (1-p_{v+1}^{u,v}) \cdot \frac{1}{N} \cdot \left((\frac{N-1}{N})^{v} - (\frac{N-1}{N})^{N-1}\right)$*, where*

$$p_{v+1}^{u,v} = \begin{cases} \frac{2(N-1)}{N^2} & \text{if } u=0 \text{ and } v=1; \\ \frac{1}{N} \cdot (\frac{N-1}{N})^{v-u} + \frac{1}{N} \cdot (\frac{N-1}{N})^{v} - \frac{1}{N^2} \cdot (\frac{N-1}{N})^{2v-u-1} & \text{otherwise.} \end{cases}$$

Proof. Here we have $v > u$. So for any $t > v$, we will have $t > max\{u, v\}$. Substituting $t = v+1$ in Lemma 3, we have
$P(S_N[u]=v) = p_{v+1}^{u,v} \cdot (\frac{N-1}{N})^{N-1-v} + (1-p_{v+1}^{u,v}) \cdot \frac{1}{N}(\frac{N-1}{N})^{v} \cdot \left(1-(\frac{N-1}{N})^{N-1-v}\right)$
$= p_{v+1}^{u,v} \cdot (\frac{N-1}{N})^{N-1-v} + (1-p_{v+1}^{u,v}) \cdot \frac{1}{N} \cdot \left((\frac{N-1}{N})^{v} - (\frac{N-1}{N})^{N-1}\right)$. Now, from Lemma 2, we get $p_{v+1}^{u,v} = \frac{1}{N} \cdot (\frac{N-1}{N})^{v-u} + \frac{1}{N} \cdot (\frac{N-1}{N})^{v} - \frac{1}{N^2} \cdot (\frac{N-1}{N})^{2v-u-1}$, except for "$u=0$ and $v=1$". Also, Lemma 1 gives $p_2^{0,1} = \frac{2(N-1)}{N^2}$. Substituting the value of $p_{v+1}^{u,v}$, we get the result. □

Note that for each u, $0 \le u \le N-2$, the expression for $P(S_N[u]=v)$ above (for $v \ge u+1$) is maximized at $v = u+1$. This is also supported by our experimental observations.

Theorem 2. *For* $v \le u \le N-1$*,*
$P(S_N[u]=v) = \frac{1}{N} \cdot (\frac{N-1}{N})^{N-1-u} + \frac{1}{N} \cdot (\frac{N-1}{N})^{v+1} - \frac{1}{N} \cdot (\frac{N-1}{N})^{N+v-u}$.

Proof. Here we have $u \ge v$. So for any $t > u$, we will have $t > max\{u, v\}$. Substituting $t = u+1$ in Lemma 3, we have
$P(S_N[u]=v) = p_{u+1}^{u,v} \cdot (\frac{N-1}{N})^{N-1-u} + (1-p_{u+1}^{u,v}) \cdot \frac{1}{N}(\frac{N-1}{N})^{v} \cdot \left(1-(\frac{N-1}{N})^{N-1-u}\right)$.
As discussed in the proof of Proposition 2, $p_{u+1}^{u,v} = P(S_{u+1}[u]=v) = \frac{1}{N}$. Substituting this in the above expression, we get
$P(S_N[u]=v) = \frac{1}{N} \cdot (\frac{N-1}{N})^{N-1-u} + (1-\frac{1}{N}) \cdot \frac{1}{N}(\frac{N-1}{N})^{v} \cdot \left(1-(\frac{N-1}{N})^{N-1-u}\right)$
$= \frac{1}{N} \cdot (\frac{N-1}{N})^{N-1-u} + \frac{1}{N} \cdot (\frac{N-1}{N})^{v+1} - \frac{1}{N} \cdot (\frac{N-1}{N})^{N+v-u}$. □

3 Discussion

To evaluate how closely our theoretical formulae tally with the experimental results, we use average percentage absolute error $\bar{\epsilon}$. Let $p_N^{u,v}$ and $q_N^{u,v}$ respectively denote the theoretical and the experimental value of the probability $P(S_N[u]=v)$, $0 \le u \le N-1$, $0 \le v \le N-1$. We define $\epsilon_{u,v} = \left(\frac{|p_N^{u,v} - q_N^{u,v}|}{q_N^{u,v}}\right) \cdot 100\%$ and $\bar{\epsilon} = \frac{1}{N^2} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \epsilon_{u,v}$. We ran experiments for 100 million randomly chosen secret keys of 32 bytes and found that $\bar{\epsilon} = 0.22\%$. The maximum of the $\epsilon_{u,v}$'s was 35.37% and it occured for $u = 128$ and $v = 127$. Though the maximum error is quite high, we find that out of $N^2 = 65536$ (with $N = 256$) many $\epsilon_{u,v}$'s, only 11 (< 0.02% of 65536) exceeded the 5% error margin. These cases are summarized Table 1 below. We call the pairs (u, v) for which $\epsilon_{u,v} > 5\%$ as *anomaly pairs.*

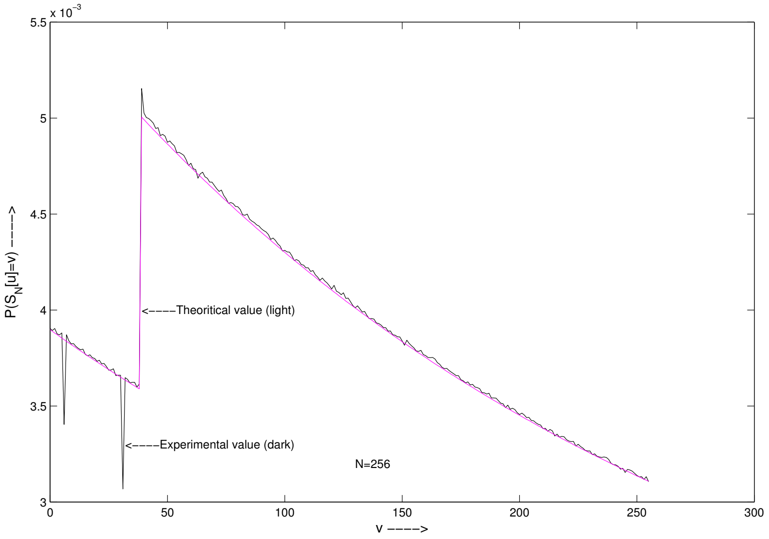

Fig. 2. Comparing Experimental and Theoretical values of $P(S_N[38] = v)$ versus v.

The experimental values of $P(S_N[u] = v)$ match with the theoretical values given by our formula except at these few anomaly pairs. As an illustration, we plot $q_N^{u,v}$ (calculated by running the KSA with 100 million random keys of length 32 bytes) and $p_N^{u,v}$ versus v for $u = 38$ in Figure 2. We see that $q_N^{38,v}$ follows the pattern predicted by $p_N^{38,v}$ for all v's, $0 \leq v \leq 255$ except at $v = 6$ and $v = 31$ as pointed out in Table 1.

We experimented with different key lengths (100 million random keys for each key length) and found that the location of the anomaly pairs and the total number of anomaly pairs vary with the key lengths in certain cases. Table 2 shows the number n_5 of anomaly pairs (when $\epsilon_{u,v} > 5\%$) for different key lengths l (in bytes) along with the average $\bar{\epsilon}$ and the maximum ϵ_{max} of the $\epsilon_{u,v}$'s. u_{max} and v_{max} are the (u, v) values which correspond to ϵ_{max}. Though for some key lengths there are more than a hundred anomaly pairs, most of them have $\epsilon_{u,v} \leq 10\%$. To illustrate this, we add the column n_{10} which shows how many of the anomaly pairs exceed the 10% error margin. The two rightmost columns show what percentage of $256^2 = 65536$ (total number of (u, v) pairs) are the numbers n_5 and n_{10}.

The results indicate that as the key length increases, the proportion of anomaly pairs tends to decrease. With 256 bytes key, we have no anomaly pair with $\epsilon_{u,v} > 5\%$, i.e., $n_5 = 0$. We are currently working on an in-depth analysis of the anomalies and the involvement of the key length.

u	v	$p_N^{u,v}$	$q_N^{u,v}$	$\lvert p_N^{u,v} - q_N^{u,v} \rvert$	$\epsilon_{u,v}$ (in %)
38	6	0.003846	0.003409	0.000437	12.82
38	31	0.003643	0.003067	0.000576	18.78
46	31	0.003649	0.003408	0.000241	7.07
47	15	0.003774	0.003991	0.000217	5.44
48	16	0.003767	0.003974	0.000207	5.21
66	2	0.003882	0.003372	0.000510	15.12
66	63	0.003454	0.002797	0.000657	23.49
70	63	0.003460	0.003237	0.000223	6.89
128	0	0.003900	0.003452	0.000448	12.98
128	127	0.003303	0.002440	0.000863	35.37
130	127	0.003311	0.003022	0.000289	9.56

Table 1. The anomaly pairs for key length 32 bytes

l	$\bar{\epsilon}$ (in %)	ϵ_{max} (in %)	u_{max}	v_{max}	n_5	n_{10}	n_5 (in %)	n_{10} (in %)
5	0.75	73.67	9	254	1160	763	1.770	1.164
8	0.48	42.48	15	255	548	388	0.836	0.592
12	0.30	21.09	23	183	293	198	0.447	0.302
15	0.25	11.34	44	237	241	2	0.368	0.003
16	0.24	35.15	128	127	161	7	0.246	0.011
20	0.20	5.99	30	249	3	0	0.005	0.000
24	0.19	4.91	32	247	0	0	0.000	0.000
30	0.19	6.54	45	29	1	0	0.002	0.000
32	0.22	35.37	128	127	11	6	0.017	0.009
48	0.18	4.24	194	191	0	0	0.000	0.000
64	0.26	35.26	128	127	6	4	0.009	0.006
96	0.21	4.52	194	191	0	0	0.000	0.000
128	0.34	37.00	128	127	3	2	0.005	0.003
256	0.46	2.58	15	104	0	0	0.000	0.000

Table 2. The number and percentage of anomaly pairs along with the average and maximum error for different key lengths

4 Concluding Remarks

We here theoretically prove the bias of each permutation byte of RC4 after the KSA. We run experiments for 100 million randomly chosen secret keys of 32 bytes and present three dimensional representations of $P(S_N[u] = v)$ versus $0 \leq u, v \leq N-1$ in Figure 3 (top one, see the last page). The numerical values of the theoretical results are presented in Figure 3 (bottom one, see the last page). Note that the graph from experimental data has a few downward spikes which actually correspond to the anomaly pairs as described in Table 1. If one gets a random permutation then the surface should have been flat at a height $\frac{1}{N}$. However, one may easily note that the surface is not at all flat and that identifies that the permutation after the RC4 KSA can be distinguished from random permutation with high confidence.

Acknowledgment: The authors like to thank Mr. Snehasis Mukherjee, Indian Statistical Institute, Kolkata for his support in the preparation of the graphs.

References

1. E. Biham and O. Dunkelman. Differential Cryptanalysis in Stream Ciphers. IACR Eprint Server, eprint.iacr.org, number 2007/218, June 6, 2007.

2. S. R. Fluhrer and D. A. McGrew. Statistical Analysis of the Alleged RC4 Keystream Generator. FSE 2000, pages 19-30, vol. 1978, Lecture Notes in Computer Science, Springer-Verlag.
3. S. R. Fluhrer, I. Mantin and A. Shamir. Weaknesses in the Key Scheduling Algorithm of RC4. Selected Areas in Cryptography 2001, pages 1-24, vol. 2259, Lecture Notes in Computer Science, Springer-Verlag.
4. J. Golic. Linear statistical weakness of alleged RC4 keystream generator. EUROCRYPT 1997, pages 226-238, vol. 1233, Lecture Notes in Computer Science, Springer-Verlag.
5. R. J. Jenkins. ISAAC and RC4. 1996
   Available at http://burtleburtle.net/bob/rand/isaac.html.
6. A. Klein. Attacks on the RC4 stream cipher. February 27, 2006.
   Available at http://cage.ugent.be/ klein/RC4/, [last accessed on June 27, 2007].
7. I. Mantin and A. Shamir. A Practical Attack on Broadcast RC4. FSE 2001, pages 152-164, vol. 2355, Lecture Notes in Computer Science, Springer-Verlag.
8. I. Mantin. A Practical Attack on the Fixed RC4 in the WEP Mode. ASIACRYPT 2005, pages 395-411, vol. 3788, Lecture Notes in Computer Science, Springer-Verlag.
9. I. Mantin. Predicting and Distinguishing Attacks on RC4 Keystream Generator. EUROCRYPT 2005, pages 491-506, vol. 3494, Lecture Notes in Computer Science, Springer-Verlag.
10. I. Mironov. (Not So) Random Shuffles of RC4. CRYPTO 2002, pages 304-319, vol. 2442, Lecture Notes in Computer Science, Springer-Verlag.
11. G. Paul, S. Rathi and S. Maitra. On Non-negligible Bias of the First Output Byte of RC4 towards the First Three Bytes of the Secret Key. Proceedings of the International Workshop on Coding and Cryptography 2007, pages 285-294.
12. G. Paul and S. Maitra. Permutation after RC4 Key Scheduling Reveals the Secret Key. Accepted in SAC 2007. An extended version is available as “RC4 State Information at Any Stage Reveals the Secret Key” in IACR Eprint Server, eprint.iacr.org, number 2007/208, June 1, 2007.
13. S. Paul and B. Preneel. Analysis of Non-fortuitous Predictive States of the RC4 Keystream Generator. INDOCRYPT 2003, pages 52-67, vol. 2904, Lecture Notes in Computer Science, Springer-Verlag.
14. S. Paul and B. Preneel. A New Weakness in the RC4 Keystream Generator and an Approach to Improve the Security of the Cipher. FSE 2004, pages 245-259, vol. 3017, Lecture Notes in Computer Science, Springer-Verlag.
15. A. Roos. A class of weak keys in the RC4 stream cipher. Two posts in sci.crypt, message-id 43u1eh$1j3@hermes.is.co.za and 44ebge$llf@hermes.is.co.za, 1995. Available at http://marcel.wanda.ch/Archive/WeakKeys.
16. D. Wagner. My RC4 weak keys. Post in sci.crypt, message-id 447o1l$cbj@cnn.Princeton.EDU, 26 September, 1995. Available at http://www.cs.berkeley.edu/∼daw/my-posts/my-rc4-weak-keys.

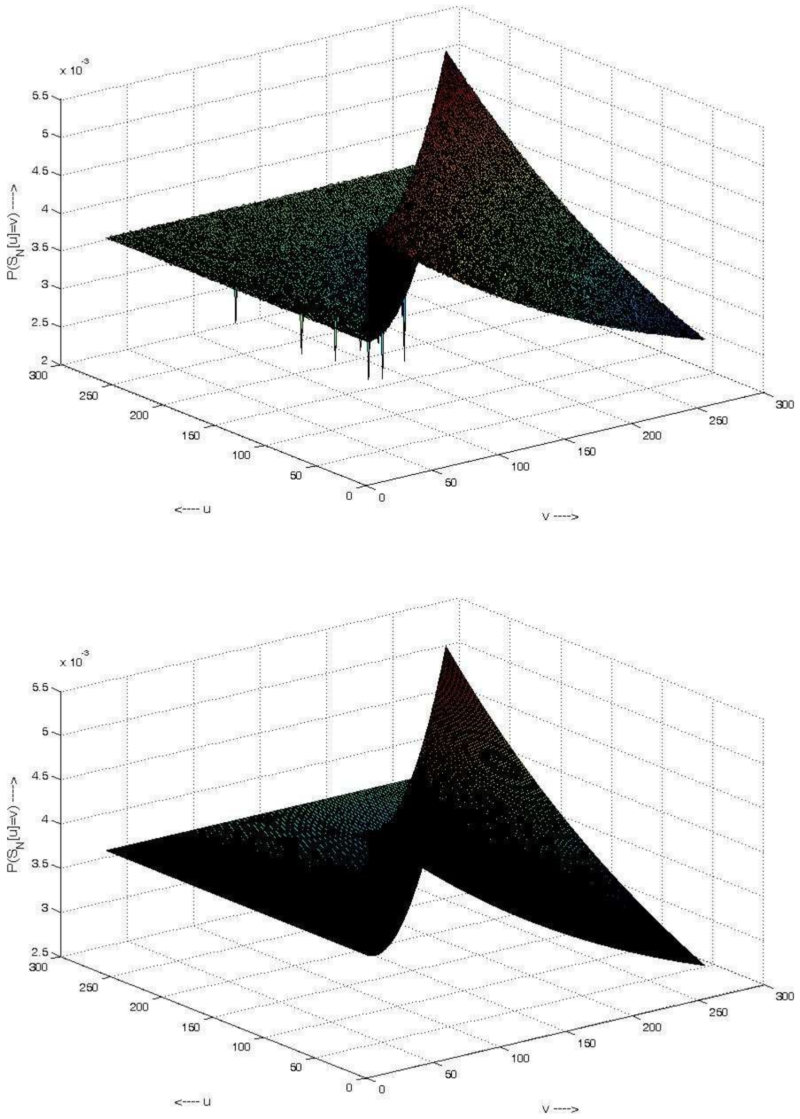

Fig. 3. $P(S_N[u] = v)$ versus $0 \leq u, v \leq N - 1$. Top: Experimental data, Bottom: Theoretical data.