Attack on Broadcast RC4
Revisited

S. Maitra\textsuperscript{1} \hspace{1cm} G. Paul\textsuperscript{2} \hspace{1cm} S. Sen Gupta\textsuperscript{1}

\textsuperscript{1}Indian Statistical Institute, Kolkata

\textsuperscript{2}Jadavpur University, Kolkata

FSE 2011, Lyngby, Denmark 15 February 2011
Outline of the Talk

Introduction
   Basics of RC4 Stream Cipher
   Motivation and Contribution

Our Result: Bias of Output Bytes
   Computing the Bias
   Exploiting the Bias
   Attack on RC4 Broadcast Scheme

Study: Non-Randomness of $j$
   Non-randomness in Different Rounds

Conclusion
   Summary of the Paper
RC4 Stream Cipher

- Designed by Ron Rivest in 1987

**Data Structure**
- $S$-array of size $N = 256$ bytes
- Key $k$ of size 5 to 16 bytes
- Final key $K$ of $N = 256$ bytes
- Two indices $i$ and $j$
- Output: Stream of bytes


RC4 Stream Cipher

**Key Scheduling Algorithm (KSA)**

\[ j = j + S[i] + K[i] \]

\[
\begin{array}{ccccccc}
0 & 1 & 2 & \cdots & i & j & 254 & 255 \\
\end{array}
\]
## RC4 Stream Cipher

### Key Scheduling Algo (KSA)

\[ j = j + S[i] + K[i] \]

\[
\begin{array}{cccccc}
0 & 1 & 2 & \cdots & i & j \\
\hline \\
\square & \square & \square & \cdots & \square & \square \\
\end{array}
\]

\[
\begin{array}{cccccc}
254 & 255 & \cdots & \square & \square \\
\hline \\
\end{array}
\]

### Pseudo-Random Gen. Algo (PRGA)

\[ j = j + S[i] \]

\[
\begin{array}{cccccc}
0 & 1 & 2 & S[i] + S[j] & \cdots & i & j \\
\hline \\
\square & \square & \square & \square & \cdots & \square & \square \\
\end{array}
\]

\[
\begin{array}{cccccc}
254 & 255 & \cdots & \square & \square \\
\hline \\
\end{array}
\]
Cryptanalysis of RC4

More than 20 years of cryptanalytic results

- Finney Cycle [1994]
- Key-Output Correlation [Roos, 1995] [Paul & Maitra, 2007, 2008]
- Key-Permutation Correlation [Roos, 1995] [Paul & Maitra, 2007]
- Non-Randomness of Permutation [Mantin, 2001]
- Fault Attacks [Hoch & Shamir, 2004] [Mantin, 2005] [Biham et al, 2005]
- Non-random event: Glimpse Bias [Jenkins, 1996]
Cryptanalysis of RC4

More than 20 years of cryptanalytic results

- Finney Cycle [1994]
- Key-Output Correlation [Roos, 1995] [Paul & Maitra, 2007, 2008]
- Key-Permutation Correlation [Roos, 1995] [Paul & Maitra, 2007]
- Non-Randomness of Permutation [Mantin, 2001]
- Fault Attacks [Hoch & Shamir, 2004] [Mantin, 2005] [Biham et al, 2005]
- Non-random event: Glimpse Bias [Jenkins, 1996]
- Distinguishing Attacks
Distinguishing Attacks

**Goal:** Find an event which occurs with different probability in RC4 than in case of a perfectly random random source.

Existing Distinguishers

- Digraph Repetition Bias (Occurrence of $ABTAB$) [Mantin, 2001]
- Biased Second Output Byte ($z_2 = 0$) [Mantin & Shamir, 2001]
- A set of new linear biases of RC4 [Sepehrdad et al, 2010]
- ... a few more in this work
Motivation for this Work

**FSE 2001.** *A Practical Attack on Broadcast RC4.*

**Main Claim:** \( \Pr(z_2 = 0) \approx \frac{2}{N} \) (bias of second byte)
Motivation for this Work

**FSE 2001.** *A Practical Attack on Broadcast RC4.*

**Main Claim:** \( \Pr(z_2 = 0) \approx \frac{2}{N} \) (bias of second byte)

Two related claims

1. \( \Pr(z_r = 0) \approx \frac{1}{N} \) at PRGA rounds \( 3 \leq r \leq 255 \).

2. \( \Pr(z_r = 0 \mid j_r = 0) > \frac{1}{N} \) and \( \Pr(z_r = 0 \mid j_r \neq 0) < \frac{1}{N} \) for \( 3 \leq r \leq 255 \). These two biases, when combined, cancel each other to give no bias at \( z_r = 0 \) for rounds 3 to 255.
Contribution of this Work


1. $\Pr(z_r = 0) \approx \frac{1}{N}$ at PRGA rounds $3 \leq r \leq 255$.

   $\Pr(z_r = 0) \not\approx \frac{1}{N}$ for $3 \leq r \leq 255$

   Additional results exploiting the above bias
Contribution of this Work


1. \( \Pr(z_r = 0) \approx \frac{1}{N} \) at PRGA rounds 3 \( \leq r \leq 255 \).

   \[ \Pr(z_r = 0) \not\approx \frac{1}{N} \text{ for } 3 \leq r \leq 255 \]

   Additional results exploiting the above bias

2. \( \Pr(z_r = 0 \mid j_r = 0) > \frac{1}{N} \) and \( \Pr(z_r = 0 \mid j_r \neq 0) < \frac{1}{N} \) for 3 \( \leq r \leq 255 \). These two biases, when combined, cancel each other to give no bias at \( z_r = 0 \) for rounds 3 to 255.

   Further investigation of the events
   Careful analysis of non-randomness of \( j \)
Our Result

Bias of Output Bytes
Our Result

Output bytes 3 to 255 are also biased to Zero

Theorem
For $3 \leq r \leq 255$, the probability that the $r$-th RC4 keystream byte is equal to 0 is

$$\Pr(z_r = 0) \approx \frac{1}{N} + \frac{c_r}{N^2}.$$  

where $c_r$ is given by

$$\left[ \left( \frac{N-1}{N} \right)^r + \left( \frac{N-1}{N} \right)^{N-r-1} - \left( \frac{N-1}{N} \right)^{N-1} \right] \cdot \left[ \left( \frac{N-1}{N} \right)^{r-2} - \frac{1}{N-1} \right].$$
Proposition (Jenkins’ Correlation)

After the $r$-th ($r \geq 1$) round of the PRGA,

$$\Pr(S_r[j_r] = i_r - z_r) = \Pr(S_r[i_r] = j_r - z_r) \approx \frac{2}{N}.$$ 

Corollary

After the $r$-th ($r \geq 1$) round of the PRGA, $\Pr(z_r = r - S_{r-1}[r]) \approx \frac{2}{N}$. 
Motivation for Proof (our result)

Proposition (Jenkins’ Correlation)

After the $r$-th ($r \geq 1$) round of the PRGA,

$$\Pr(S_r[j_r] = i_r - z_r) = \Pr(S_r[i_r] = j_r - z_r) \approx \frac{2}{N}.$$

Corollary

After the $r$-th ($r \geq 1$) round of the PRGA, \( \Pr(z_r = r - S_{r-1}[r]) \approx \frac{2}{N}. \)

How about $\Pr(S_{r-1}[r] = r)$?
Mantin’s Observation

At the end of KSA, for $0 \leq u \leq N - 1$, $0 \leq v \leq N - 1$,

$$
\Pr(S_0[u] = v) = \frac{1}{N} \left[ \left( \frac{N-1}{N} \right)^v + \left( 1 - \left( \frac{N-1}{N} \right)^v \right) \left( \frac{N-1}{N} \right)^{N-u-1} \right] \quad v \leq u
$$

$$
\Pr(S_0[u] = v) = \frac{1}{N} \left[ \left( \frac{N-1}{N} \right)^{N-u-1} + \left( \frac{N-1}{N} \right)^v \right] \quad v > u
$$
Mantin’s Observation

At the end of KSA, for \(0 \leq u \leq N - 1\), \(0 \leq v \leq N - 1\),

\[
\Pr(S_0[u] = v) = \frac{1}{N} \left[ \left( \frac{N-1}{N} \right)^v + \left( 1 - \left( \frac{N-1}{N} \right)^v \right) \left( \frac{N-1}{N} \right)^{N-u-1} \right] \quad \text{if } v \leq u
\]

\[
\Pr(S_0[u] = v) = \frac{1}{N} \left[ \left( \frac{N-1}{N} \right)^{N-u-1} + \left( \frac{N-1}{N} \right)^v \right] \quad \text{if } v > u
\]

Does this propagate to PRGA?
Sketch of Proof (our result)

- Mantin’s Observation: Bias for $S_0[r] = r$

- $S_{r-1}[r] = r$ may happen in two ways:
  1. $S_0[r] = r$ and $i, j$ never touches this cell
  2. $S_0[r] \neq r$ but $S_{r-1}[r] = r$ occurs at random
Our Result: Bias of Output Bytes

Sketch of Proof (our result)

- Mantin’s Observation: Bias for $S_0[r] = r$

- $S_{r-1}[r] = r$ may happen in two ways:
  1. $S_0[r] = r$ and $i, j$ never touches this cell
  2. $S_0[r] \neq r$ but $S_{r-1}[r] = r$ occurs at random

Lemma

*For $r \geq 3$, the probability that $S_{r-1}[r] = r$ is*

$$
\Pr(S_{r-1}[r] = r) \approx \Pr(S_0[r] = r) \cdot \left[ \left( \frac{N-1}{N} \right)^{r-1} - \frac{1}{N} \right] + \frac{1}{N}.
$$
Sketch of the Proof (our result)

$z_r = 0$ can be branched as follows:

- $S_{r-1}[r] = r$ (lemma) and $z_r = r - S_{r-1}[r]$ (Jenkin)
- $S_{r-1}[r] \neq r$ (lemma) and $z_r = 0$ (random)
Sketch of the Proof (our result)

$z_r = 0$ can be branched as follows:

- $S_{r-1}[r] = r$ (*lemma*) and $z_r = r - S_{r-1}[r]$ (*Jenkin*)
- $S_{r-1}[r] \neq r$ (*lemma*) and $z_r = 0$ (*random*)

Hence the result: $\Pr(z_r = 0) \approx \frac{1}{N} + \frac{c_r}{N^2}$

with $c_r = \left[ \left( \frac{N-1}{N} \right)^r + \left( \frac{N-1}{N} \right)^{N-r-1} - \left( \frac{N-1}{N} \right)^{N-1} \right] \left[ \left( \frac{N-1}{N} \right)^{r-2} - \frac{1}{N-1} \right]$. 

Our Result: Bias of Output Bytes

Numerical Bound on $c_r$

\[
\max_{3 \leq r \leq 255} c_r = c_3 = 0.98490994 \quad \text{and} \quad \min_{3 \leq r \leq 255} c_r = c_{255} = 0.36757467
\]

\[
\frac{1}{N} + \frac{0.98490994}{N^2} \geq \Pr(z_r = 0) \geq \frac{1}{N} + \frac{0.36757467}{N^2}
\]
Our Result: Bias of Output Bytes

Experimental Verification

- Number of trials = 1 Billion
- Key size = 16 Bytes

[Note: Sepehrdad et al (2010) do not cover these biases]
Applications

Of the Biases Discovered
Appl. 1: A Class of New Distinguishers

$E$ occurs in $X$ with probability $p$ and in $Y$ with probability $p(1 + \epsilon)$ implies a possible distinguisher with $O(p^{-1}\epsilon^{-2})$ required samples.

In case of our $E$: $z_r = 0$ for $3 \leq r \leq 255$,

- Random source: $p = \frac{1}{N}$
- RC4 Keystream: $p(1 + \epsilon) = \frac{1}{N} \left(1 + \frac{c_r}{N}\right)$
Appl. 1: A Class of New Distinguishers

$E$ occurs in $X$ with probability $p$ and in $Y$ with probability $p(1 + \epsilon)$ implies a possible distinguisher with $O(p^{-1}\epsilon^{-2})$ required samples.

In case of our $E$: $z_r = 0$ for $3 \leq r \leq 255$,

- Random source: $p = \frac{1}{N}$
- RC4 Keystream: $p(1 + \epsilon) = \frac{1}{N} \left(1 + \frac{c_r}{N}\right)$

We get 253 new distinguishers, each requiring $O(N^3)$ samples!

[Note: Mantin & Shamir (2001) distinguisher is much stronger]
Appl. 2: Guessing State Information

**Idea:** Guess $S_{r-1}[r] = r$ using output information $z_r = 0$

$$
\Pr(S_{r-1}[r] = r \mid z_r = 0) = \frac{\Pr(S_{r-1}[r]=r)}{\Pr(z_r=0)} \cdot \Pr(z_r = 0 \mid S_{r-1}[r] = r)
$$

$$
\approx 2 \cdot \left( \frac{1}{N} + \frac{c_r}{N} - \frac{c_r}{N^2} \right) \cdot \left(1 + \frac{c_r}{N}\right)^{-1} \geq \frac{2}{N}
$$
Appl. 3: Attack on RC4 Broadcast Scheme

**Situation:** Message $M$ is broadcast to $k$ parties (random keys)

**Attack:** Reliably extract byte(s) of $M$ from the $k$ ciphertexts
Appl. 3: Attack on RC4 Broadcast Scheme

Situation: Message $M$ is broadcast to $k$ parties (random keys)

Attack: Reliably extract byte(s) of $M$ from the $k$ ciphertexts

Mantin & Shamir (FSE 2001): Extract 2nd byte of $M$ given $k = \Omega(N)$

We can extract bytes 3 to 255 of $M$ given $k = \Omega(N^3)$

Idea: $r$-th byte of $M$ gets XOR-ed with $z_r$, which is 0 most often.
Study

NON-RANDOMNESS OF $j$
Non-Randomness of $j_1$

Note that $j_1 = j_0 + S_0[i_1] = 0 + S_0[1] = S_0[1]$, where $S_0$ is the state array right after KSA is over.

$$
\Pr(j_1 = v) = \Pr(S_0[1] = v) = \begin{cases} 
\frac{1}{N}, & v = 0 \\
\frac{1}{N} \left( \frac{N-1}{N} + \frac{1}{N} \left( \frac{N-1}{N} \right)^{N-2} \right), & v = 1 \\
\frac{1}{N} \left( \left( \frac{N-1}{N} \right)^{N-2} + \left( \frac{N-1}{N} \right)^v \right), & v > 1 
\end{cases}
$$

Clearly not random!
Non-Randomness of $j_1$

Note that $j_1 = j_0 + S_0[i_1] = 0 + S_0[1] = S_0[1]$, where $S_0$ is the state array right after KSA is over.

$$\Pr(j_1 = \nu) = \Pr(S_0[1] = \nu) = \begin{cases} 
\frac{1}{N}, & \nu = 0 \\
\frac{1}{N} \left( \frac{N-1}{N} + \frac{1}{N} \left( \frac{N-1}{N} \right)^{N-2} \right), & \nu = 1 \\
\frac{1}{N} \left( \left( \frac{N-1}{N} \right)^{N-2} + \left( \frac{N-1}{N} \right)^{\nu} \right), & \nu > 1
\end{cases}$$

Clearly not random!
Non-Randomness of $j_2$

Note that $j_2 = j_1 + S_1[i_2] = S_0[1] + S_1[2]$

$$\Pr(j_2 = \nu) = \sum_{w=0}^{N-1} \Pr(S_0[1] = w) \cdot \Pr((S_1[2] = \nu - w) \mid (S_0[1] = w))$$

**Case I.** $S_0[1] = 2 \Rightarrow S_1[2] = 2$.

$$\Pr((S_1[2] = \nu - 2) \mid (S_0[1] = 2)) = \begin{cases} 1 & \text{if } \nu = 4, \\ 0 & \text{otherwise.} \end{cases}$$

**Case II.** $S_0[1] \neq 2 \Rightarrow S_1[2] = S_0[2]$.

$$\Pr((S_1[2] = \nu - w) \mid (S_0[1] \neq 2)) = \Pr(S_0[2] = \nu - w).$$
Non-Randomness of $j_2$

Study: Non-Randomness of $j_n$
Non-Randomness of $j_2$

**Appl:** Combine Jenkin’s bias $\Pr(S_r[i_r] = j_r - z_r) = \frac{2}{N}$ to get

$$\Pr(S_2[i_2] = 4 - z_2) \approx \frac{1}{N} + \frac{4/3}{N^2}$$

[Note: Sepehrdad et al (2010) do not cover this bias]

[Note: $j$ behaves almost random round 3 onwards]
Summary

This paper: Revisiting Mantin–Shamir paper from FSE 2001

1. Bias of Keystream bytes 3–255 towards Zero  NEW
   - A new class of distinguishers for RC4
   - Attack on RC4 broadcast scheme along this line
   - Guessing related state information from keystream

2. Strong bias of $j_2$ towards 4  NEW
   - Guessing related state information from keystream
THANK YOU
FOR YOUR KIND ATTENTION