Cryptology ePrint Archive: Report 2011/448

(Non-)Random Sequences from (Non-)Random Permutations - Analysis of RC4 stream cipher

Sourav Sen Gupta and Subhamoy Maitra and Goutam Paul and Santanu Sarkar

Abstract: RC4 has been the most popular stream cipher in the history of symmetric key cryptography till date. Its internal state contains a pseudo-random permutation over all n-bit words (typically $n = 8$) and it attempts to generate a pseudo-random sequence of words by extracting elements of this permutation. Since more than last twenty years, numerous cryptanalytic results on RC4 stream cipher have been published. Many of these results are based on some non-random (biased) events involving the secret key or the state variables or the output sequence, or a combination of them.

Though biases based on the secret key is common in RC4 literature, none of the existing ones depends on the length of the secret key. In the first part of this paper, we report significant biases involving the length of the secret key, for the first time in the literature.

In the second part of the paper, theoretical proofs of some significant initial-round empirical biases observed by Sepehrdad, Vaudenay and Vuagnoux [SAC 2010] are presented. Another important result presented here is the derivation of the complete probability distribution of the first byte of RC4 output sequence, a problem left open for a decade since the observation by Mironov [CRYPTO 2002]. Further, the existence of positive biases towards zero for all the initial bytes 3 to 255 is proved and exploited towards a generalized broadcast attack on RC4 stream cipher.

The above biases discussed in this paper, like most of the existing biases in RC4 literature, are short-term and do not last after a few initial rounds. The last part of this paper investigates the long-term manifestation of short-term biases in RC4 output sequence. A careful analysis of the periodic structure of RC4 evolution proves the first long-term generalization of Mantin and Shamir's [FSE 2001] famous second-byte bias.

Category / Keywords: Bias, Cryptography, Distinguisher, Probability Distribution, Pseudo-Random Permutation, Pseudo-Random Word, Random Sequences, RC4, Sequences, Stream Ciphers.

Date: received 16 Aug 2011, last revised 2 Jan 2012

Contact author: subho at isical ac in