June 1, 2010

The Best Known State Recovery Attacks on RC4

| JOVAN GOLIĆ | Security Innovation |

| GUGLIELMO MORGARI | Telsy Elettronica e Telecomunicazioni |

Rump Session of Eurocrypt 2010, Monaco, May 30 – June 3, 2010

Outline

1. Description of RC4 Keystream Generator

- RC4-N works over $\mathbb{Z}_N$ (additions modulo N)
- Internal state: permutation S and pointers i, j
- Initial state: S_0 and $i, j = 0$
- Next-state function:

$$i \leftarrow i + 1$$
$$j \leftarrow j + S[i]$$
$$\text{Swap } S[i], S[j]$$

- Output function:

$$z \leftarrow S[S[i] + S[j]]$$

2. Objective and Previous Results

- **Objective:** *Recover S_0 or any S from a keystream segment*
  - Data complexity: segment length D
  - Time complexity: T computational steps
- **Best previous results, for $N=256$:**
  - **Knudsen et al. Asiacrypt '98:** $D=2^8$, $T=2^{779}$
    - ***State recovery by systematic state search with backtracking,*** in consistency with keystream
  - **Maximov & Khovratovich Crypto '08:** $^*D=2^{248}$, $T=2^{242}$ (hypothetical)
    - Same as above, but using special state patterns that uniquely determine consecutive values of j pointer
    - Such a state pattern needs to be found along keystream in a D&C manner, without running state recovery

3. Iterative Probabilistic Algorithms-1

- Knudsen et al. Asiacrypt '98:
  - **Recursive forward computation** of approximate *a posteriori* probabilities for state components, given keystream
  - *A priori* distribution of S_0 consists of d known consecutive entries and the remaining uniform probabilities (estimate for $D=N=256$: $d=155$ suffices for full state recovery; too large for G&D attack)
- Golić ACISP '00:
  - **Improved forward computation** of approximate *a posteriori* probabilities for state components, given keystream (joint effects ‘change of state’ and ‘observation of output symbol’)
  - **Backward computation** of these probabilities, given keystream
  - **Iterative algorithm** composed of rounds consisting of one Forward and one Backward pass

3. Iterative Probabilistic Algorithms-2

- Algorithm IPA, improvement of Golić ACISP '00:
  - *A priori* distribution of S_0 or any S adapted to deal with:
    - *Consecutive state patterns:* $i, j; S[k+i], 1 \le k \le d$; it follows that (at least) d subsequent values of j pointer are uniquely determined – *d patterns*
    - *Maximum state patterns:* $i, j; S[p_k]=v_k, 1 \le k \le d$, such that w subsequent values of j pointer are uniquely determined and w is (close to) maximal – *(d,w) patterns*
  - **Other improvements include:** hard preprocessing, hard reset of Backward, soft preprocessing, modifying initial probability matrix of Forward, soft zero row reset, and soft inconsistent column reset
- Round complexities: $T=2N^6$, $D=N$, memory $M=2(N^2+N)$

4. G&D Attack for Consecutive State Patterns

- In a G&D attack, value of *(shifted) d pattern* is guessed and, for each guess, **IPA** is run on $D=N$ keystream
- If guess is correct, then states are fully recovered with a success probability p, depending on d and N
- **On the basis of systematic experiments** for $16 \le N \le 80$ and various d and about 10 experiments for $N=128$ (about one month per 5+1 rounds of **IPA**), we make

  Conjecture 1: If $d/N \cong 1/3$, then $p \cong 0.5$, for $N \ge 48$.
- Attack complexities, for $N=256$:
  - Basic version ($p \cong 0.5$): $D=2^9$, $T=2^{724}$
  - Optimized version (optimized p): $D=2^{57}$, $T=2^{676}$

5. G&D Attack for Maximum State Patterns

- In a G&D attack, position of *shifted (d,w) pattern* is guessed and, for each guess, **IPA** is run on $D = N$ subsequent keystream
- If guess is correct, then states are fully recovered with a success probability p, depending on w, d, and N
- **On the basis of systematic experiments** for $4 \leq d \leq 9$ and various N and about 10 experiments for $N = 128$, we make

  Conjecture 2: If $^{*}d \cong N/10 + 1$, then $p \cong 0.5$, for $N \geq 256$.
  (*using conjecture $w/(d-1) \cong 6$ of Maximov & Khovratovich)
- Attack complexities, for $N = 256$:
  - Basic version ($p \cong 0.5$): $D = 2^{223}$, $T = 2^{275}$
  - Optimized version (optimized p): $D = 2^{208}$, $T = 2^{260}$

6. D&C Attack for Maximum State Patterns

- In a G&D attack, **IPA**, with time complexity $T \approx N^6$, is run for each guessed position of shifted maximum state pattern
- In a D&C attack, correct position of *unshifted pattern* is found in a D&C manner, without running **IPA** for state recovery, by the method from Maximov & Khovratovich Crypto '08
- Required additional properties of state pattern are not shift invariant and hence the value of i pointer has to be matched, so that D increases N times; $w/(d-1)$ is somewhat reduced and hence D further increases, but $T \cong D/N$
- Attack complexities, for $N = 256$:
  - Basic version ($p \cong 0.5$): $D \approx 2^{231}$, $T \approx 2^{223}$ (hypothetical)
  - Optimized version (optimized p): $D \approx 2^{216}$, $T \approx 2^{208}$ (hypothetical)