New results on reduced-round Tiny Encryption Algorithm using genetic programming

Karel Kubíček

karel.kubicek@mail.muni.cz

Faculty of Informatics, Masaryk University, Brno, Czech Republic

Abstract

Analysing cryptographic functions usually requires extensive work of a skilled cryptanalyst. However, some automation is possible, e.g., by using randomness testing batteries such as NIST STS or Dieharder. Yet such tests are limited to predefined test patterns. However, there is a new approach – EACirc is a novel randomness testing framework based on finding a distinguisher for a given cipher output. In this work, we use EACirc to analyse the outputs of Tiny Encryption Algorithm (TEA). TEA was previously used with genetic algorithms for evolution of bit masks used for restriction of cipher input. Instead of evolving bit masks, we create a software circuit applicable as a distinguisher for limited-round TEA (up to 4 rounds). Results of EACirc are also compared to the standard statistical batteries.

Keywords: randomness, statistical testing, TEA, genetic algorithms, software circuit

1 Introduction

Automatized randomness testing is useful for checking one of the expected cipher properties – output ciphertext should be indistinguishable from a stream of random data. This property alone is not sufficient for cipher to be secure, but the ability to distinguish ciphertexts from random data constitutes an important hint on cipher weakness.

The common way to automate testing of randomness is using statistical batteries. NIST STS [1], Dieharder [2] are standard batteries of tests, that are commonly used for this purpose. The batteries contain sets of fixed tests checking expected statistics of tested output stream (TEA ciphertext in our case) in comparison to truly random data.

The limitation of the standard batteries for randomness testing is the fact they implement a fixed set of tests and can detect only a limited set of patterns and statistical irregularities. In this work[1] we use EACirc [3], a novel framework for constructing empirical tests of randomness that can succeed in finding such a test (at least hypothetically). Our goal is to find an empirical test of randomness that indicates if a given sequence is either non-random (with a high probability) or sufficiently indistinguishable from truly random data stream. In the framework, randomness tests are computed iteratively, adapting to the processed sequence. The construction is stochastic and uses genetic programming. Tests are constructed from a predefined pool of operations (building blocks). Set of operations, together with a limit for the number of operations, allows us to control the complexity of the tests. The framework theoretically allows us to construct an arbitrary randomness test over a set of chosen operations (in practice, however, the total number of operations used is limited). Therefore, it can be viewed as a general framework for the test construction and should (hypothetically) provide a better detection ability than standard tests.

The approach of genetic algorithms is different then running predefined sets of tests from a statistical battery. Firstly, a set of individuals is created with each individual representing a candidate distinguisher function. Secondly, every individual decides if the provided block of input data is random or non-random. Thirdly, better individuals are randomly mutated or cross-bred to create better descendants (the ratio of correct guesses constitutes a fitness function). Therefore, if ciphertexts have a common property expressible as a distinguisher function an individual representing this function can be potentially evolved.

We performed several experiments on Tiny Encryption Algorithm. Nowadays, the cipher is not considered to be secure for regular use as it suffers from multiple weaknesses. However, it was used as a benchmark for randomness testing using genetic algorithms starting with a paper in 2002 by Julio C. Hernández, José M. Sierra, Pedro Isasi and Arturo Ribagorda [4], who were successful with TEA limited to 1 and 2

[1] For full version of this text, follow http://crcs.cz/papers/mkb2015

rounds. 2 years later, a similar team published new results with improved settings [5], which resulted in 3 to 4 rounds. The newest results from 2010 by Wei Hu [6], using quantum inspired genetic algorithm, succeeded with 4 to 5 rounds TEA.

2 Results

Although previous papers provide exact number of rounds, where finding distinguisher was possible, comparison with our approach is not straightforward.

All previous works published weights of constructed masks, which were used to change both the input data and key for the encryption. Maximum weight of mask is $64+128 = 192$ bits, such mask would allow to use any input data possible. They published the mask weights (denoted MW in table) together with average χ^2 statistics of maximal deviation of the ciphertext from random distribution.

Rounds	[4]		[5]		[6]		NIST	Dieharder	EACirc
	χ^2	MW	χ^2	MW	χ^2	MW	$(x/162)$	$(x/20)$	%
1	8380416	72	522240	153	522.240	153	162	20	100
2	1900	77	736.05	155	602	171	162	18.5	100
3	(untested)		393.6	116	530.756	117.8	162	16.5	100
4	(untested)		294.86	50	742.632	67.6	6	11.0	100
5	(untested)		(untested)		631.74	76	(0)	(0)	(5)

Table 1: Comparison of selected results for reduced rounds TEA.

The next part of the table provides results from statistical test batteries NIST STS (version 2.1.1) [1] and Dieharder (version 3.31.1) [2]. Dieharder provides three levels of evaluation (pass, weak, fail), we assigned to these levels values 1, 0.5 and 0 respectively. The result in the cell is sum of the 20 tests. From NIST STS we were using 162 tests (with pass or fail level). Result with 0 passed tests is in parenthesis as mark of indistinguishable from random result. The column for EACirc represents the best achieved results from our measurements. The values express the percentage of runs, where statistically significant distinguisher was found. For the reference random-random distinguishing experiment, the value of 5% is expected (and also measured), so we mark this value with parenthesis as indistinguishable from random.

3 Summary

EACirc is a different approach for randomness testing, where we allow for the creation of new distinguishers. We were able to find distinguisher for TEA limited to 4 rounds, which is a comparable result to the statistical batteries and previous works.

Acknowledgement

I would like to thank Petr Švenda and Martin Ukrop for notes and corrections during work on this paper.

References

[1] A. Rukhin, "A Statistical Test Suite for the Validation of Random Number Generators and Pseudo Random Number Generators for Cryptographic Applications, Version STS-2.1" in *NIST Special Publication 800-22rev1a*, 2010.

[2] R. G. Brown, D. Eddelbuettel and D. Bauer, "Dieharder: A random number test suite", *Duke University Physics Department*, 2009.

[3] P. Švenda, M. Ukrop and others, "EACirc project", [Online]. Available: https://github.com/crocs-muni/EACirc.

[4] J. C. Hernández, J. M. Sierra, P. Isasi and A. Ribagorda, "Genetic Cryptoanalysis of Two Rounds TEA", in *Computational Science—ICCS*, 2002, pp. 1024–1031.

[5] J. C. Hernández, J. M. Sierra and P. Isasi, "Finding Efficient Distinguishers for Cryptographic Mappings, with an Application to the Block Cipher TEA", *Computational Intelligence*, vol. 20, no. 3, pp. 517–525, 2004.

[6] W. Hu *et al.*, "Cryptanalysis of TEA Using Quantum-Inspired Genetic Algorithms", *Journal of Software Engineering and Applications*, vol. 3, no. 01, pp. 50-57, 2010.