Randomness test suite for evaluation of cryptographic primitives

Ph.D. Thesis - abstract

mgr. inż. Krzysztof MAŃK

thesis supervisor:

dr hab. Marek KOJDECKI

secondary supervisor:

dr inż. Michał WROŃSKI

Statistical analysis is one of the most basic and powerful tools in science. In particular, it is one of the tools used in the work of a cryptologist, where in the form of the so-called randomness tests, it is used to detect all kinds of irregularities that may appear in the tested bit sequences, indicating imperfections of the tested cryptographic primitives.

However, the enormous computing power and memory capacity of modern computers is limited and puts an effective barrier on the way to checking everything; hence, it is important to use it for calculations that give reliable results and do not have redundancies that can be removed, and at the same time cover the maximum area of properties that random sequence should have.

Subjecting a set of sequences produced using a tested cryptographic primitive (e.g. a random or pseudorandom generator, a block cipher, or a hash function) to a series of randomness tests is one of the important elements of assessing its quality, sometimes deciding on its acceptance or rejection. This procedure was used, for example, during the AES competition.

In the world literature, one can find references to at least seven sets of test procedures, such as: Chapter 3.2 The Art of Programming by Donald Knuth, DIEHARD, NIST Statistical Test Suite, Dieharder, Test U01, ENT, Crypt-X. Some of them are literature items or are available on the Internet, but all of them, in our opinion, do not meet the above requirements. These are very diverse sets, often containing each other or completely different sets of basic tests. In the presented work, we directly or indirectly demonstrate that the test procedures included in them have not been properly verified in terms of compliance of the distribution or correlation of statistics. In such conditions, it is very difficult to make a decision about using a specific package or to establish evaluation criteria based on many partial results.

This is how the need arose to create our own, properly functioning test suite, which will be a non-trivial compilation of the knowledge available in this field. The presented work and, above all, the results of the research are intended to provide premises and lead to the formulation of an assessment of all recognized randomness tests in two levels. This will be the individual

correctness of the results in each test, including determining the scope of its applicability. The second aspect will be to analyze the dependencies between various tests, with the aim of narrowing down the set of tests necessary to perform as much as possible. This has a double meaning: the first is purely economic, allowing the testing to be carried out in a shorter time, and the second is more scientific – it makes us aware of the relationships between various tests that are not obvious at first glance.

The process of creating the dissertation itself was divided into several stages, which were reflected in its structure. The work is divided into six main, numbered chapters, preceded by an introduction, and concluded with a summary. Chapter I is intended to introduce the reader to general concepts related to the study of randomness and to present the objectives of the work against. Chapters II and III constitute a compendium of test and tests packages. The tests described there have been grouped according to their purpose and the way they interpret the tested sequence. The primary goal was the reliability of the results obtained, so, already at the stage of collecting the specifications of individual tests, we tried to obtain the most up-to-date results regarding them, repeatedly determining distribution parameters and other constants given in the literature, in order to achieve high precision of the determined *p-values*. In several cases, theoretical results were also obtained.

We consider the most important results obtained while building the package to be:

- determining the exact values of the parameters of the test statistics distribution in first- and second-order moment tests for a bit sequence,
- generalization of the autocorrelation test for a binary sequence and determination of distribution parameters,
- co-development of a method to determine distribution parameters in Marsaglia monkey tests,
- speeding up and simplifying the calculation of the test statistics in the frequency test for overlapping vectors,
- acceleration of the implementation of the minimum distance, *m* nearest pairs and Bickel-Breiman tests,
- proposal and determination of the test distribution for 2 and 3 numbers,
- determining the exact parameter values for the Maurer test,
- determining the parameters of the distribution of the test statistics in the derivative test of a sequence of numbers,
- development of a method for determining distributions in binary tree filling tests based on a sequence of bits and a sequence of numbers.

In Chapter IV, we presented and justified the basic tool used in this work, the three-level testing procedure. This prepared the ground for the main part of the work contained in chapters V and VI. The first one was originally titled "Elimination of procedures", which perfectly reflected what we devoted this part to – we removed from the package being built all procedures that did not pass the three-level testing procedure. We did it in three steps, gradually narrowing down the set of considered procedures, which allowed us to define three sets of procedures with different scopes of applicability and at the same time reduced the computational effort.

The content of Chapter VI can be described as an attempt to further reduce the number of procedures using two-dimensional linear and nonlinear correlation, and finally multivariate linear correlation. Again, we used the gradation of the tools used. After obtaining confirmation of the existence of correlation for most of the procedures, we computationally used the simplest linear correlation as a tool for eliminating redundant procedures. We then generalized this tool to test 11 nonlinear models. The analysis of classes into which the set of procedures was divided by a correlation relationship led us to use another tool, the form of multivariate linear regression.

In many cases, we used only selected block or shift sizes, such as prime numbers and powers of two, with a simple assumption - if the test works correctly for parameter values m and $m + 2$, it will function correctly for $m + 1$. This was done to save computational time and contributed to the less spectacular results of the correlation part. However, the results obtained in Chapter VI for such a reduced set showed that in many cases the parameters were too dense, and our proposed method can be successfully used in the process of building a test suite.

In two Appendices attached to presented work, we have compiled lists of test procedures that, in our opinion, may constitute proposals for packages useful for various variants of use. In particular, lists contained in paragraph 5 of both annexes are an important practical result of our work.

The package built as part of the work contains implementations of 80 randomness tests, for which 36,584 combinations of parameters were considered for sequences of 1 kiB and 6,879,437 combinations for 1 Mb. For use in the evaluation for large numbers of repetitions, 13,249 procedures for 1 kiB sequences and 2,034,402 procedures for 1 Mb sequences were recommended. The package was implemented in C++ in the form of 20 main libraries, the volume of the source code is over 79 Mi characters, contained in over 157 thousand lines of code.

Krzysztof Mark