TESTING REPORT

Issued by:	**NMi Metrology & Gaming Ltd** Llys Helyg, Parc Menai, Bangor LL57 4EZ, United Kingdom Tel: +44 (0)1248 660550
Report to:	**Greentube Alderney Ltd** Century House, 12 Victoria Street, Alderney GY9 3UF, Channel Islands
Project name:	**Random Number Generator on the Greentube platform**
Jurisdiction:	United Kingdom
Issue date:	06 April 2018

Executive Summary

This report summarises an assessment of a Random Number Generator (RNG), deployed with the Greentube Remote Game Server (NRGS). The RNG provides random outputs to a variety of game engines (labelled as 'game servers' by Greentube).

Source code and a test harness for empirical testing were provided. The containing platform was not supplied. The assessment methods included source code review and statistical analysis of RNG outputs generated by the supplied test harness.

The RNG has been assessed for compliance with section 7A of the "Remote gambling and software technical standards" (June 2017) and the scope and methodology comply with sections 2.2 and 3.2 of the "Testing strategy for compliance with remote gambling and software technical standards" (June 2017).

The output of the RNG was determined to be acceptably random, unpredictable, and not reproducible for all of the ranges assessed. Based on the selection of tests executed, it is the assessment of NMi UK that the RNG is suitable for production of random numbers on all of the ranges tested.

No issues are raised.

Authorised by:

Aled Hughes
Quality, Compliance and Risk Manager

Table of Contents

Introduction

NMi Metrology & Gaming Ltd (NMi UK) is accredited to ISO/IEC 17025 by the United Kingdom Accreditation Service (UKAS) to undertake compliance testing of all categories of modern gaming systems and related equipment at their own and their customer's premises. NMi UK's ISO 17025 accreditation schedule is downloadable from the UKAS website.

Scope of ISO/IEC 17025

All assessments in the following sections of this report are provided under ISO/IEC 17025 except (as in the case of interpretations, opinions and suggestions) where otherwise stated.

Quality Control

The monitoring of this testing project was the responsibility of NMi UK's Quality Manager and every effort has been made to ensure the accuracy of the information contained in this report. If errors or omissions are discovered, please contact NMi UK with details as soon as possible. NMi UK reserves the right to revise and reissue this report if additional information is presented or discovered.

Caveats

The results presented in this document are a summary of the testing work undertaken. This report is subject to a number of caveats, including:

- All items provided for inspection and/or testing are declared by the customer to be configured identically to those in commercial use, with the exception of operator-configurable aspects that will not have a bearing on game fairness or player returns.
- All software and source code provided for empirical testing and/or code review is declared by the customer to behave identically to the software and code in commercial use.
- Decisions taken by the supplied software in automatic test modes / simulators are reasonable emulations of those that would be expected to be taken by real players.

All efforts have been taken to ensure that the testing undertaken was as exhaustive as necessary to demonstrate compliance or non-compliance. NMi UK takes on trust that all test items (including all hardware and software), all documentation and all communications are accurate, truthful and that there is no intention to deceive or subvert the assessment of compliance.

Test Item Details

Critical Components

SHA-1 checksum	File name
4e1dd1d07f597cc30c9f82b928ca2de3ed31d3ec	rng.jar

Testing Overview

Customer Contacts

The customer liaisons were Jacqueline Pirron and Anton Tanzer.

Dates

Testing was undertaken during the following periods:

- 06/02/2018 - 04/04/2018

Locations

Testing was undertaken at the following locations:

- Llys Helyg, Parc Menai, Bangor LL57 4EZ, United Kingdom
- 530-4445 Lougheed Highway, Burnaby, British Columbia, V5C 0E4, Canada

Regulatory Documents

Compliance with the relevant aspects of the following regulatory documents was assessed:

Document	Abbreviation Used
The Gambling Act 2005 (April 2005)	UK_TGA
Remote gambling and software technical standards (June 2017)	UK_RTS
Testing strategy for compliance with remote gambling and software technical standards (June 2017)	UK_TSC

Methods

The assessment methods included source code review and statistical analysis of RNG outputs generated by the supplied test harness.

Code Analysis

The submission consisted of Java source code files and included the RNG and wrapper classes.

The RNG is a Java implementation of the Mersenne Twister (MT19937) pseudo-random algorithm. The RNG is implemented in a remote application interface.

Scope

The following public functions of the RNG have been assessed:

- `getData()`

Overview

The Mersenne Twister algorithm is not cryptographically secure; however, the implementation in this RNG is multi-threaded and includes a time-based background cycling mechanism. Even given knowledge of the algorithm, implementation and seed, it would be computationally infeasible to predict the next number when the system is operating under normal conditions.

The code submitted for testing the RNG includes the use of SecureRandom as an additional source of entropy. Numbers drawn from this source are combined with the RNG output to produce random outcomes. This is in accordance with the intended use of the RNG, which states that game servers will provide this source of entropy.

Seeding

The RNG is seeded with a value based on system time and a supplied string. The RNG is not re-seeded in the submitted code.

Outputs

The RNG can produce the following outputs:

- a single chunk of 1024 random bytes (or numbers in the range [-128,127]) may be drawn from the RNG service at any time after initialisation.

Output Scaling

The output of the RNG is not scaled in the submission.

Empirical Testing

Test Methods

A number of empirical tests have been proposed to analyse frequencies of occurrence and localised correlations, patterns and intervals between generated numbers.

In this analyses the following tests are used for raw RNG output:

NIST Test Suite

The following "bitwise" tests from the NIST Test Suite were applied:

- Frequency (Monobits) Test
- Frequency Test within a Block
- Run Test
- Test for the Longest Run of Ones in a Block
- Binary Matrix Rank Test
- Discrete Fourier Transform (Spectral) Test
- Non-Overlapping Template Matching Test
- Maurer's "Universal Statistical" Test
- Linear Complexity Test
- Serial Test
- Approximate Entropy Test
- Cumulative Sums (Cumsum) Test
- Random Excursions Test
- Random Excursions Variant Test

Diehard Battery of Tests

The following "bitwise" tests from the Diehard Battery of Tests of Randomness were applied:

- Birthday Spacing
- Overlapping 5-permutations
- Binary Rank 31x31
- Binary Rank 32x32
- Binary Rank 6x8
- Bitstreams
- Overlapping Pairs Sparse Occupancy (OPSO)
- Overlapping Quadruples Sparse Occupancy (OQSO)
- DNA
- Count the 1s (Specific Bytes)
- Count the 1s (Stream of Bytes)
- Parking Lot
- Minimum Distance
- 3-D Spheres
- Squeeze
- Overlapping Sums
- Runs
- Craps

Donald Knuth's Empirical Tests for Randomness

The following tests were applied to the scaled and shuffled RNG outputs:

- Frequency test (Equidistribution test)
- Serial test (non-overlapping pairs)
- Gap test
- Poker test (Partition test)
- Permutation test
- Run test

All test results are based on the Pearson chi-squared test (also known as the chi-square "goodness of fit" test) to compare the observed results against expected outcomes and determine a level of confidence.

For the following test descriptions, assume that a number *n* of uniformly distributed random numbers on the range [0, *m*-1], with *m* being an amount of distinct outcomes, were generated.

Frequency Test

The Frequency Test is designed to ensure that the random numbers are uniformly distributed throughout a given interval. The instances of each number in the range [0, *m*-1] are counted and the counts compared to the expected populations. The probability *P* of observing any particular number *x* in a given position in the sequence is:

$$P(x) = \frac{1}{m}, \quad 0 \leq x \leq m-1.$$

The variation in observed distribution against the theoretical value is used to calculate the chi-squared statistic. The value of chi-squared statistic then maps to a probability (i.e. a p-value) that provides a measure of confidence in the observed outcomes.

Serial Test

The Serial Test checks that pairs of numbers are uniformly distributed in an independent manner. The random numbers are distributed into a number of equal bins and the frequencies of occurrence of all possible sequence pairs are checked (i.e. 0 followed by 0, 0 followed by 1, ..., *m*-1 followed by *m*-1). If the numbers are uncorrelated (i.e. no sequence pairs are favoured over any others), an equal distribution is expected and the probability of observing a sequence *(x,y)* is equal to:

$$P(x,y) = \frac{1}{m^2}, \quad 0 \leq x,y \leq m-1.$$

Similar to the frequency test, the observations and theoretical probabilities are used to compute a chi-squared statistic, which is then used to determine a probability that all serial pairs are uniformly distributed.

Gap Test

The Gap Test considers the length of "gaps" between occurrences of specific numbers (i.e. the average gap between an occurrence of the number "1" and the next occurrence of "1" should be the same as that between a "2" and the next "2").

To apply the gap test, the lengths of the gaps between occurrences of a particular number are collated and the frequencies of occurrence are compared with the expected counts for each gap size. If subsequent numbers in the sequence are random and independent, the probability of a gap of length *g*, between instances of a particular output with probability $p = 1/m$, occurring is:

$$P(g) = p(1-p)^g$$

All gaps larger than a pre-determined threshold are grouped into a single category and counted. The probability of observing a gap of length *u* or larger is:

$$\sum_{g=u}^{\infty} P(g) = (1-p)^u$$

A comparison of the observed and the expected gap sizes (via the chi-squared test) is then applied to assess if the sequence was generated by a sufficiently random source.

Poker Test

The Poker Test uses the analogy of a five-card hand in a poker game. It considers groups of five successive integers and observes which of the following 5 patterns is matched by each quintuple:

- 5 values (all different)
- 4 values (one pair)
- 3 values (two pairs or three of a kind)
- 2 values (full house or 4 of a kind)
- 1 value (five of a kind)

If each individual outcome is equally probable, the probability of achieving *v* distinct outcomes (in a group of *k* outcomes with *d* possible outcomes) is given by:

$$P(v) = S(k,v) \times \left(\frac{d(d-1)\ldots(d-v+1)}{d^k} \right)$$

where *S(k,v)* is the Stirling Number of the second kind (the number of ways to partition a set of *k* elements into *v* non-empty subsets).

To apply the Poker test, the generated random numbers are gathered into groups and categorised according the patterns listed above. The counts of each categorisation are compared with expected values via the chi-squared test.

Permutations Test

The Permutations Test divides a number sequence with a range of *m* elements into *n* groups of *t* elements. In this specific application, groups of *t* = 3 numbers were considered (denoted *a*,*b*,*c*) and counted the occurrence of each of the 6 different relative orderings:

- *a* < *b* < *c*
- *a* < *c* < *b*
- *b* < *a* < *c*
- *b* < *c* < *a*
- *c* < *a* < *b*
- *c* < *b* < *a*

The cases where two or more of the three numbers in a group are equal are also counted. The probability *P** that two or more of the instances are equal is given by:

$$P^* = \frac{1}{m} + \frac{2(m-1)}{m^2}$$

Hence, the probability of observing any of the listed permutations (*lp*) is:

$$P(lp) = \frac{(1-P^*)}{3!}$$

A chi-squared test is conducted to test whether the observed counts of the permutations (including the matching cases) is consistent with the theoretical distribution.

Run Test

A sequence of random numbers will typically contain sub-sequences in which the numbers are increasing (they "run up") and sub-sequences in which they are decreasing (they "run down"). In this test, the sequence is split into segments in which the length is determined by whether or not the next number is higher (in the case of "run up") or lower (in the case of "run down"). The number immediately following a run is discarded in order to make runs independent and make the chi-square test applicable. The observed value is then compared with the theoretical value and a level of confidence is calculated.

Consider a sequence of uniformly distributed random numbers with *m* possible individual outcomes. The expected probability of a run of *r* consecutive numbers is:

$$P(r) = \begin{cases} m^{-r}\binom{m}{r} - m^{-(r+1)}\binom{m}{r+1} & \text{if } 0 < r < m \\ m^{-m} & \text{if } r = m. \end{cases}$$

The number of independent runs of each length up to and including *m* are compiled and compared with the expected values via a chi-squared goodness-of-fit test.

References:

- Rukhin, A., Soto, J., Nechvatal, J., Smid, M., Barker, E., Leigh, S., Levenson, M., Vangel, M., Banks, D., Heckert, A., Dray, J., Vo, S. (2010) *A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications* [Online] v1a. Gaithersburg, MD, USA. National Institute of Standards & Technology. Available: http://csrc.nist.gov/groups/ST/toolkit/rng/documents/SP800-22rev1a.pdf
- Knuth, D. (1997). The Art of Computer Programming. Volume 2: Seminumerical Algorithms. 3rd ed. Boston: Addison-Wesley Longman Publishing Co, Inc.

Test Results

A test harness was supplied, for the purposes of testing, that provided a range of scaled outputs obtained from the RNG.
A selection of ranges were used to test the compatibility of the RNG for supplying random game outcomes.
The results are summarised as follows:

Analysis of 1 set of 64 million numbers between 0 and 2^16 - 1 (inclusive)

Test Name	Sample Size	Test Result
Frequency (Monobits) Test	64,000,000	Pass
Frequency Test within a Block	64,000,000	Pass
Runs Test	64,000,000	Pass
Longest Run of Ones in a Block	64,000,000	Pass
Binary Matrix Rank Test	64,000,000	Pass
Discrete Fourier Transform (Spectral) Test	64,000,000	Pass
Non-Overlapping Template Matching Test	64,000,000	Pass
Maurer's "Universal Statistical" Test	64,000,000	Pass
Linear Complexity Test	64,000,000	Pass
Serial Test	64,000,000	Pass
Approximate Entropy Test	64,000,000	Pass
Cumulative Sums (Cumsum) Test	64,000,000	Pass
Random Excursions Test	64,000,000	Pass
Random Excursions Variant Test	64,000,000	Pass

In summary, the data set passed the NIST suite of tests at the 95% confidence level, confirming that the software RNG is functioning correctly from a bitwise randomness perspective.

Analysis of 3 sets of 6 million raw numbers between 0 and 2^16 - 1 (inclusive)

Test Name	Sample Size	Test Result
Birthday Spacing	3x 6,000,000	Pass
Overlapping 5-permutations	3x 6,000,000	Pass
Binary Rank 31x31	3x 6,000,000	Pass
Binary Rank 32x32	3x 6,000,000	Pass
Binary Rank 6x8	3x 6,000,000	Pass
Bitstreams	3x 6,000,000	Pass
OPSO	3x 6,000,000	Pass
OQSO	3x 6,000,000	Pass
DNA	3x 6,000,000	Pass
Count the 1's - Specific Bytes	3x 6,000,000	Pass
Count the 1's - Byte Stream	3x 6,000,000	Pass
Parking Lot	3x 6,000,000	Pass
Minimum Distance	3x 6,000,000	Pass
3-D Spheres	3x 6,000,000	Pass
Squeeze	3x 6,000,000	Pass
Overlapping Sums	3x 6,000,000	Pass
Runs	3x 6,000,000	Pass
Craps	3x 6,000,000	Pass

In summary, the data set passed the Diehard suite of tests at the 95% confidence level, confirming that the software RNG is functioning correctly from a bitwise randomness perspective.

Analysis of 30 sets of 1 million scaled numbers between 0 and 36 (inclusive)

Test Name	Sample Size	Test Result
Frequency test (Equidistribution test)	30,000,000	Pass
Serial test (non-overlapping pairs)	30,000,000	Pass
Gap test	30,000,000	Pass
Poker test (Partition test)	30,000,000	Pass
Permutation test	30,000,000	Pass
Run test	30,000,000	Pass

An additional assessment was applied across multiple sets of observations to ensure consistency with the expected behaviour of a high quality random number generator at the 95% confidence level. The frequency of occurrences of the possible outcomes was as expected for a random distribution and the outcomes covered the full range of possibilities.

Analysis of 30 sets of 1 million scaled numbers between 0 and 51 (inclusive)

Test Name	Sample Size	Test Result
Frequency test (Equidistribution test)	30,000,000	Pass
Serial test (non-overlapping pairs)	30,000,000	Pass
Gap test	30,000,000	Pass
Poker test (Partition test)	30,000,000	Pass
Permutation test	30,000,000	Pass
Run test	30,000,000	Pass

An additional assessment was applied across multiple sets of observations to ensure consistency with the expected behaviour of a high quality random number generator at the 95% confidence level. The frequency of occurrences of the possible outcomes was as expected for a random distribution and the outcomes covered the full range of possibilities.

Analysis of 20 sets of 1 million scaled numbers between 0 and 255 (inclusive)

Test Name	Sample Size	Test Result
Frequency test (Equidistribution test)	20,000,000	Pass
Serial test (non-overlapping pairs)	20,000,000	Pass
Gap test	20,000,000	Pass
Poker test (Partition test)	20,000,000	Pass
Permutation test	20,000,000	Pass
Run test	20,000,000	Pass

An additional assessment was applied across multiple sets of observations to ensure consistency with the expected behaviour of a high quality random number generator at the 95% confidence level. The frequency of occurrences of the possible outcomes was as expected for a random distribution and the outcomes covered the full range of possibilities.

Conclusion

The data passed the NIST and Diehard bitwise tests for randomness. The data also passed a series of Knuth tests described in the previous section.

No deviations or biases were detected from the RNG in the generated data and the RNG is deemed suitable for the intended gaming applications.

Assessment

Section		Subsection	Compliant	Observation	Potential Issue	Not Applicable
UK_RTS	RTS 7 – Generation of random outcomes	7A:1	✓			
		7A:2				✓[01]
		7A:3				✓[01]

[01] This is an RNG test only

END OF REPORT