Predictive Models for Min-Entropy Estimation

John Kelsey Kerry A. McKay Meltem Sönmez Turan

National Institute of Standards and Technology

meltem.turan@nist.gov

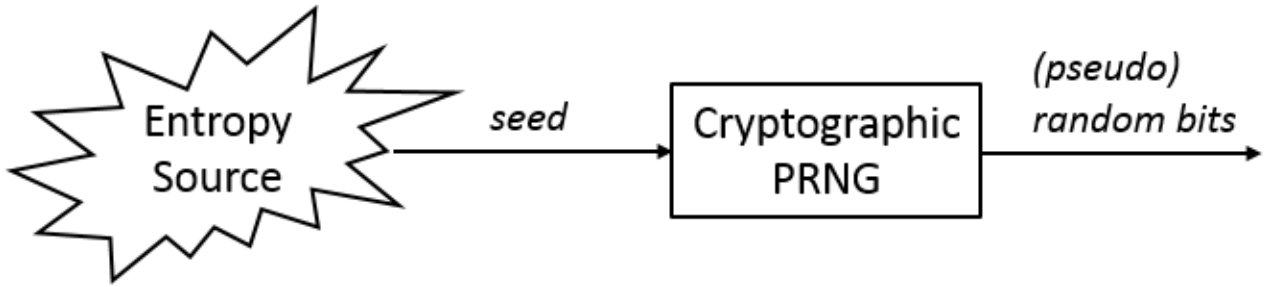
September 15, 2015

Overview

- Cryptographic random numbers and entropy
- NIST SP 800 90 Series - Recommendations on random number generation
- A framework to estimate entropy based on *predictors*
- Experimental results using simulated and real noise sources

Random Numbers in Cryptography

- We need random numbers for cryptography (e.g., secret keys, IVs, nonce, salts)
- Cryptographic PRNGs (based on hash functions, block ciphers etc.) generate strong random numbers for crypto applications.
- PRNGs need random seeds.



Entropy sources

- Entropy sources depend on physical events (e.g. thermal noise).
- The outputs of entropy sources are usually statistically weak, (e.g. correlated outputs).
- Entropy sources are fragile, sensitive to external factors (e.g. temperature).

Entropy sources

- Entropy sources depend on physical events (e.g. thermal noise).
- The outputs of entropy sources are usually statistically weak, (e.g. correlated outputs).
- Entropy sources are fragile, sensitive to external factors (e.g. temperature).

How many bits are needed to be drawn from the entropy source to produce a good seed?

Low Entropy leads to weak keys!

Heninger et al. (2012) performed the largest network survey of TLS and SSH servers.

- Collected 5.8 million unique certificates from 12.8 million TLS hosts and 6.2 million unique SSH host keys from 10.2 million hosts.
- Observed that 0.75% of TLS certificates share keys during key generation. Obtained RSA private keys for 0.50% of TLS hosts and 0.03% of SSH hosts, DSA private keys for 1.03% of SSH hosts.

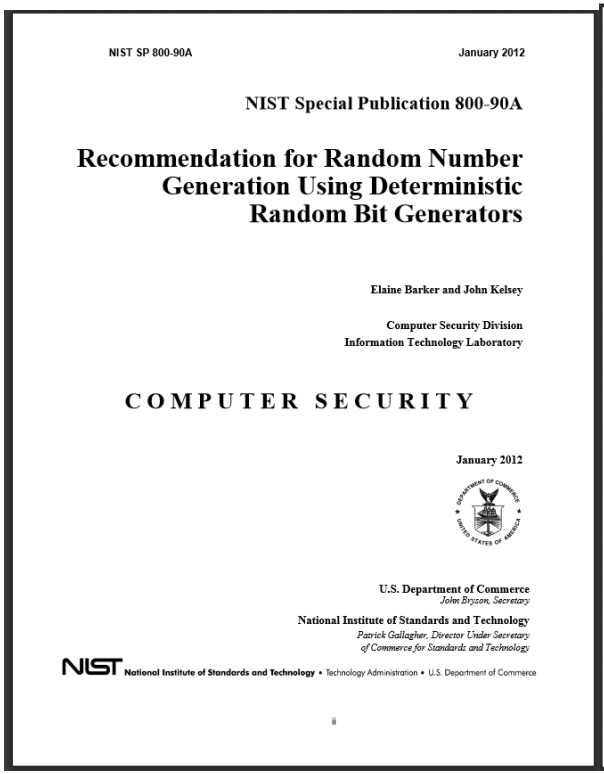
Why?

- Using manufacturers-default keys,
- Low entropy during key generation.

Entropy

- Measure of the amount of information contained in a message.
- Different measurements of entropy:
  - Shannon ($-\sum p_i \log p_i$), Renyi ($\frac{1}{1-\alpha} \log(\sum p_i^{\alpha})$), Min-entropy ($-\log max_i p_i$).
- Entropy estimation
  - Plug-in estimators (maximum likelihood estimators), methods based on compression algorithms etc.
  - Challenging, when underlying distribution is unknown, and the *i.i.d.* assumption cannot be made.
  - Under/over estimation

NIST Special Publication 800-90 Series



NIST SP 800-90A January 2012

NIST Special Publication 800-90A

Recommendation for Random Number Generation Using Deterministic Random Bit Generators

Elaine Barker and John Kelsey

Computer Security Division
Information Technology Laboratory

COMPUTER SECURITY

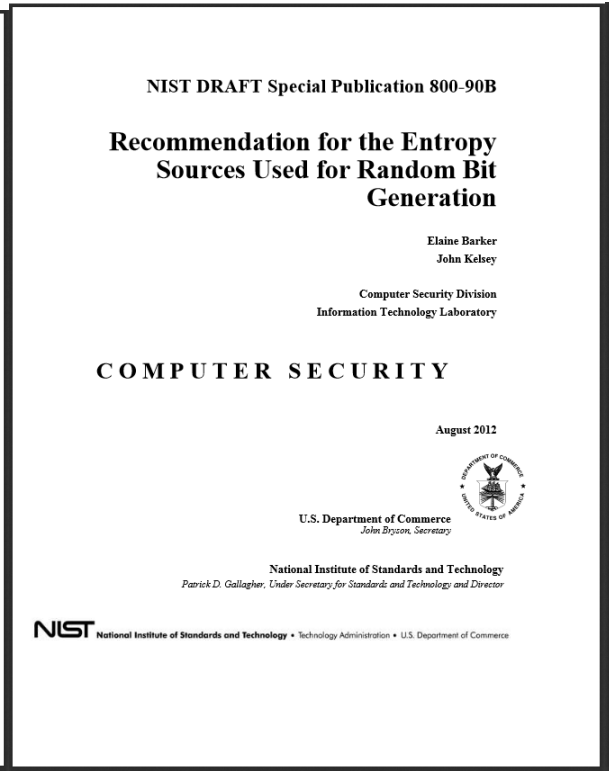
January 2012

U.S. Department of Commerce
John Bryzon, Secretary

National Institute of Standards and Technology
Patrick Gallagher, Director Under Secretary of Commerce for Standards and Technology

NIST National Institute of Standards and Technology • Technology Administration • U.S. Department of Commerce

ii



NIST DRAFT Special Publication 800-90B

Recommendation for the Entropy Sources Used for Random Bit Generation

Elaine Barker
John Kelsey

Computer Security Division
Information Technology Laboratory

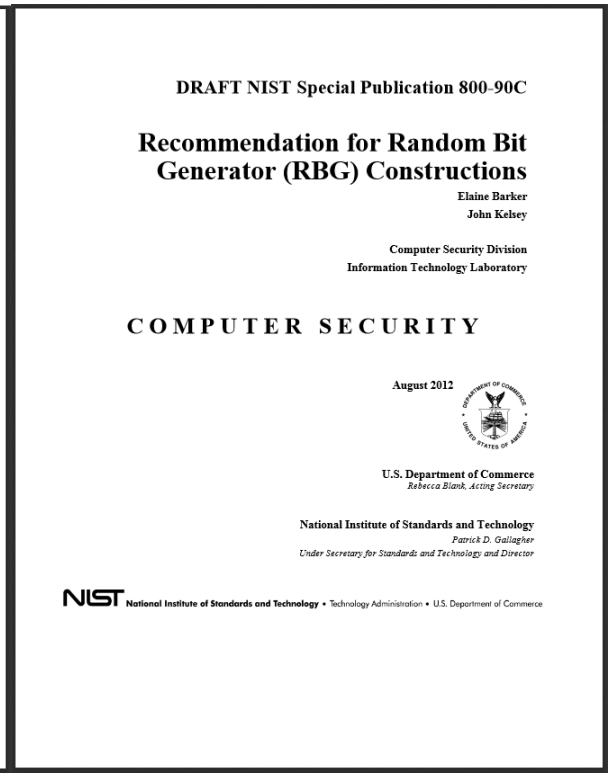
COMPUTER SECURITY

August 2012

U.S. Department of Commerce
John Bryzon, Secretary

National Institute of Standards and Technology
Patrick D. Gallagher, Under Secretary for Standards and Technology and Director

NIST National Institute of Standards and Technology • Technology Administration • U.S. Department of Commerce



DRAFT NIST Special Publication 800-90C

Recommendation for Random Bit Generator (RBG) Constructions

Elaine Barker
John Kelsey

Computer Security Division
Information Technology Laboratory

COMPUTER SECURITY

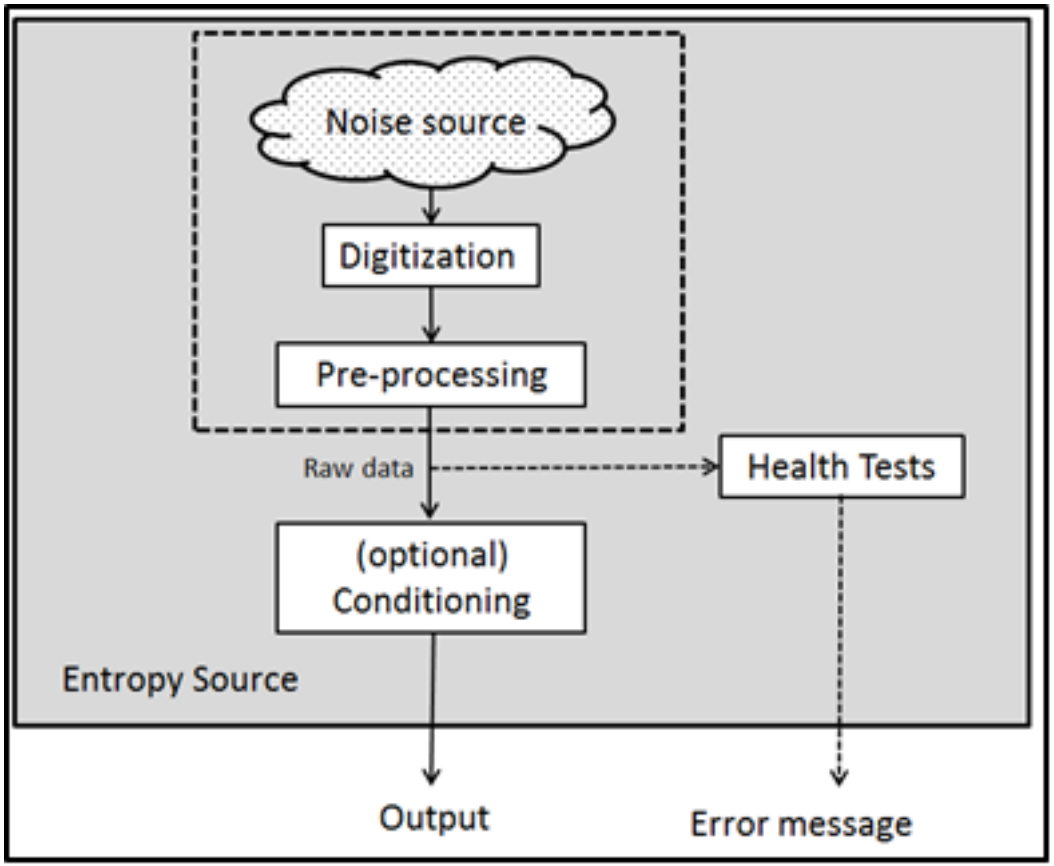
August 2012

U.S. Department of Commerce
Rebecca Blank, Acting Secretary

National Institute of Standards and Technology
Patrick D. Gallagher
Under Secretary for Standards and Technology and Director

NIST National Institute of Standards and Technology • Technology Administration • U.S. Department of Commerce

Entropy Source Model used in SP 800-90B



Entropy Estimation - 90B Perspective

- In order to comply with Federal Information Processing Standard 140-2, designers/submitters first submit analysis of their noise source.
- Labs check the noise source specification, and then generate raw data from the noise source, and estimate entropy using the 90B estimators.
- For validation purposes, estimation process cannot be too complex, due to time and cost constraints, and the process cannot require expert knowledge.
- Any two validation labs must get the same result for same outputs.

NIST's Entropy Estimation Suite

Draft SP 800-90B (Aug.2012) includes five estimators for non-i.i.d. sources:

- *Collision test* (based on the mean time for a repeated sample value.)
- *Partial collection test* (based on the number of distinct sample values observed in segments of the outputs.)
- *Markov test* (models the noise source outputs as a first-order Markov model.)
- *Compression test* (based on how much the noise source outputs can be compressed.)
- *Frequency test* (based on the number of occurrences of the most-likely value.)

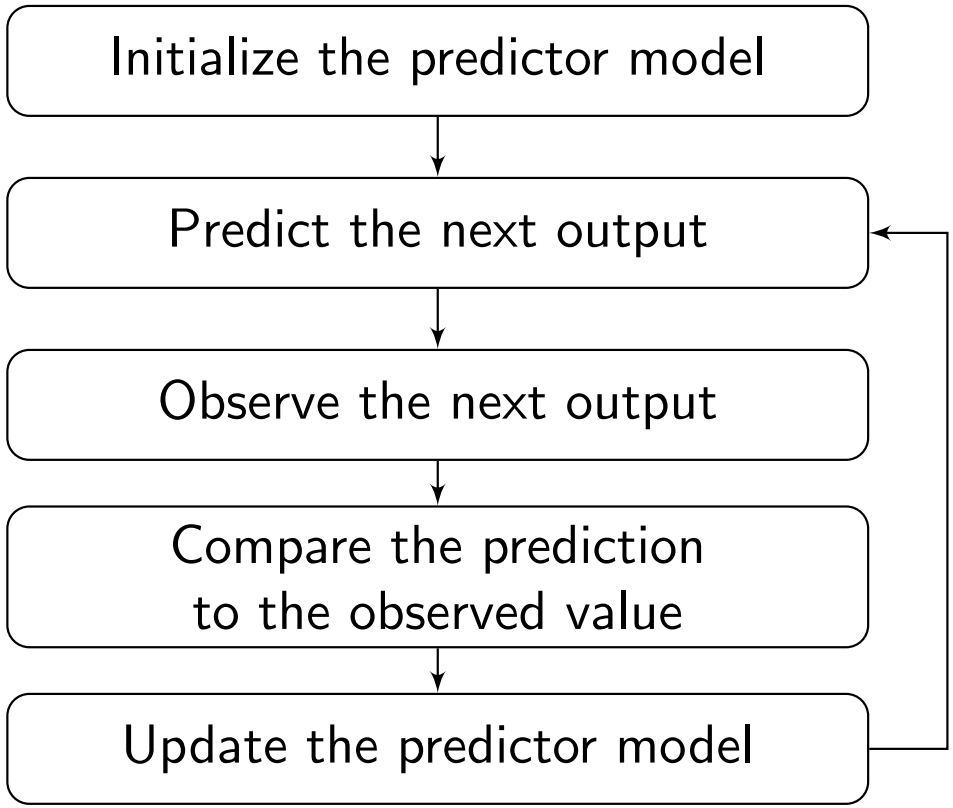
The final entropy estimate is the **minimum** of all the estimates.

Drawbacks of SP 800 90B Estimators

- Estimators do not always catch bad behavior e.g. Higher-order Markov models
- Prone to underestimate when there are many probability levels (e.g., discrete approximation of a normal distribution)
- Estimators may not work very well if the probability distribution changes over time

New Framework based on Predictors

- Predictors behave more like an attacker, who observes source outputs and makes guesses based on previous observations.
- Predictor approach complements 90B suite without significantly lowering estimates



Global vs. Local Predictability

Two ways to estimate probability of predicting the next value:

- Global predictability
  - Proportion of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, on average
- Local Predictability
  - Probability of guessing correctly given the longest run of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, should the entropy source fall into a predictable state

Estimate min-entropy using the higher probability, with a 99% confidence interval approach.

Global vs. Local Predictability

Two ways to estimate probability of predicting the next value:

- Global predictability
  - Proportion of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, on average
- Local Predictability
  - Probability of guessing correctly given the longest run of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, should the entropy source fall into a predictable state

Estimate min-entropy using the higher probability, with a 99% confidence interval approach.

Raw data	1
Prediction	0
Correct	

Global vs. Local Predictability

Two ways to estimate probability of predicting the next value:

- Global predictability
  - Proportion of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, on average
- Local Predictability
  - Probability of guessing correctly given the longest run of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, should the entropy source fall into a predictable state

Estimate min-entropy using the higher probability, with a 99% confidence interval approach.

Raw data	1
Prediction	0
Correct	0

Global vs. Local Predictability

Two ways to estimate probability of predicting the next value:

- Global predictability
  - Proportion of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, on average
- Local Predictability
  - Probability of guessing correctly given the longest run of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, should the entropy source fall into a predictable state

Estimate min-entropy using the higher probability, with a 99% confidence interval approach.

Raw data	1
Prediction	0 1
Correct	0

Global vs. Local Predictability

Two ways to estimate probability of predicting the next value:

- Global predictability
  - Proportion of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, on average
- Local Predictability
  - Probability of guessing correctly given the longest run of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, should the entropy source fall into a predictable state

Estimate min-entropy using the higher probability, with a 99% confidence interval approach.

Raw data	1	0
Prediction	0	1
Correct	0	

Global vs. Local Predictability

Two ways to estimate probability of predicting the next value:

- Global predictability
  - Proportion of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, on average
- Local Predictability
  - Probability of guessing correctly given the longest run of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, should the entropy source fall into a predictable state

Estimate min-entropy using the higher probability, with a 99% confidence interval approach.

Raw data	1	0
Prediction	0	1
Correct	0	0

Global vs. Local Predictability

Two ways to estimate probability of predicting the next value:

- Global predictability
  - Proportion of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, on average
- Local Predictability
  - Probability of guessing correctly given the longest run of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, should the entropy source fall into a predictable state

Estimate min-entropy using the higher probability, with a 99% confidence interval approach.

Raw data	1	0	
Prediction	0	1	1
Correct	0	0	

Global vs. Local Predictability

Two ways to estimate probability of predicting the next value:

- Global predictability
  - Proportion of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, on average
- Local Predictability
  - Probability of guessing correctly given the longest run of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, should the entropy source fall into a predictable state

Estimate min-entropy using the higher probability, with a 99% confidence interval approach.

Raw data	1	0	1
Prediction	0	1	1
Correct	0	0	

Global vs. Local Predictability

Two ways to estimate probability of predicting the next value:

- Global predictability
  - Proportion of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, on average
- Local Predictability
  - Probability of guessing correctly given the longest run of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, should the entropy source fall into a predictable state

Estimate min-entropy using the higher probability, with a 99% confidence interval approach.

Raw data	1	0	1
Prediction	0	1	1
Correct	0	0	1

Global vs. Local Predictability

Two ways to estimate probability of predicting the next value:

- Global predictability
  - Proportion of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, on average
- Local Predictability
  - Probability of guessing correctly given the longest run of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, should the entropy source fall into a predictable state

Estimate min-entropy using the higher probability, with a 99% confidence interval approach.

Raw data	1	0	1	
Prediction	0	1	1	0
Correct	0	0	1	

Global vs. Local Predictability

Two ways to estimate probability of predicting the next value:

- Global predictability
  - Proportion of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, on average
- Local Predictability
  - Probability of guessing correctly given the longest run of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, should the entropy source fall into a predictable state

Estimate min-entropy using the higher probability, with a 99% confidence interval approach.

Raw data	1	0	1	0
Prediction	0	1	1	0
Correct	0	0	1	

Global vs. Local Predictability

Two ways to estimate probability of predicting the next value:

- Global predictability
  - Proportion of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, on average
- Local Predictability
  - Probability of guessing correctly given the longest run of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, should the entropy source fall into a predictable state

Estimate min-entropy using the higher probability, with a 99% confidence interval approach.

Raw data	1	0	1	0
Prediction	0	1	1	0
Correct	0	0	1	1

Global vs. Local Predictability

Two ways to estimate probability of predicting the next value:

- Global predictability
  - Proportion of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, on average
- Local Predictability
  - Probability of guessing correctly given the longest run of correct guesses over a long sequence
  - Captures attacker's ability to guess next output, should the entropy source fall into a predictable state

Estimate min-entropy using the higher probability, with a 99% confidence interval approach.

Raw data	1	0	1	0	0	1	1	1	0	0	1
Prediction	0	1	1	0	0	1	0	1	1	0	0
Correct	0	0	1	1	1	1	0	1	0	1	0

Set of Predictors

Single predictors

- LZ78Y – most common sample that follows value according to dictionary
- Most Common in Window (MCW) – most common sample in last w observations
- Markov Model with Counting (MMC) – most common sample that follows last N values
- Single Lag – sample that appeared N observations back
- Moving average – prediction based on moving average of last w samples
- D1 – prediction based on difference equation of two most recent observations

Set of Predictors

Ensemble predictors

- Contain multiple instances of single predictors of the same type with different parameter values
- Scoreboard keeps track of instance that is performing best
- Use best prediction from best-performing instance at the time

Experimental Results - Simulated Sources

- *Discrete uniform distribution:* an i.i.d. source in which the samples are equally-likely.
- *Discrete near-uniform distribution:* an i.i.d source where all samples but one are equally-likely; the remaining sample has a higher probability than the rest.
- *Normal distribution rounded to integers:* an i.i.d. source where samples are drawn from a normal distribution and rounded to integer values.
- *Time-varying normal distribution rounded to integers:* a non-i.i.d. source where samples are drawn from a normal distribution and rounded to integer values, but the mean of the distribution moves along a sine curve to simulate a time-varying signal.
- *Markov Model:* a non-i.i.d. source where samples are generated using a kth-order Markov model.

Result Summary: Simulated Sources

Error measures for the lowest 90B and predictor estimates by simulated source class

Simulated data class	90B MSE	Predictor MSE	90B MPE	Predictor MPE
Uniform	2.4196	0.5031	37.9762	17.4796
Near-uniform	1.4136	0.1544	26.6566	6.4899
Normal	4.9680	0.4686	62.6330	14.1492
Time-varying normal	3.0706	0.2564	54.1453	3.1706
Markov	0.9973	0.8294	6.4339	-11.7939

MSE = mean squared error, MPE = mean percentage error

Result Summary: Real-World Sources

Entropy estimates for real world sources. The lowest entropy estimate for each source is shown in bold font

Estimator	RDTSC1	RDTSC4	RDTSC8	RANDOM.ORG	Ubld.it1	Ubld.it8
Collision	0.9125	3.8052	5.3240	**5.1830**	0.9447	**5.2771**
Compression	0.9178	3.6601	**5.3134**	5.1926	0.9285	5.5081
Frequency	0.9952	3.9577	5.8666	5.6662	0.8068	5.8660
Markov	0.9983	3.9582	5.7858	5.3829	0.8291	5.7229
Partial Collection	0.9258	3.7505	5.3574	5.5250	0.9407	5.8238
D1	0.9616	3.9986	7.9619	7.9126	0.8734	7.9489
Lag	**0.7075**	3.9883	7.9546	7.9237	**0.7997**	7.9862
LZ78Y	0.9079	3.9989	11.9615	11.5924	**0.7997**	11.8375
MultiMA	0.9079	3.6458	7.9594	7.8508	0.8073	7.9441
MultiMCW	0.9079	3.9888	7.9381	7.9744	0.8072	7.9544
MultiMMC	0.9079	**3.6457**	7.9663	7.9237	0.8072	7.9880

Conclusion

- *For the designer:* The best efforts of the designer to understand the behavior of his noise source may not be fully successful. An independent test of the unpredictability of the source can help the designer recognize these errors.
- *For an evaluator:* A testing lab or independent evaluator trying to decide how much entropy per sample a source provides will have limited time and expertise to understand and verify the designer's analysis of his design. Entropy tests are very useful as a way for the evaluator to double-check the claims of the designer.
- *For the user:* A developer making use of one or more noise sources can sensibly use an entropy estimation tool to verify any assumptions made by the designer.

Thanks for you attention!

meltem.turan@nist.gov