CPU Jitter Heuristic Entropy Arguments

Joshua Hill (KeyPair)
Steele Myrick (Corsec)
Lisa Rabe (cisco)

Meetings: 20240119, 20240202, 20240216, 20240304.

Agenda

~~Approach 1 – Sub-Distribution-Oblivious Presumed IID Analysis~~

~~Approach 2 – Sub-Distribution-Oblivious Empirical Analysis~~

Approach 3 – Sub-Distribution-Oblivious Essentially-IID Analysis (NIST 1 option, follow up)

Approach 4 – Multiple Sub-Distribution Empirical Analysis (NIST 2, initial)

Approach 5 – Multiple Sub-Distribution Essentially-IID Analysis (NIST 1 option, initial)

Approach 6 – Single Sub-Distribution Empirical Analysis

Approach 7 – Single Sub-Distribution Essentially-IID Analysis

~~Approach 8 – (Hypothetical) Single Sub-Distribution Stochastic Model~~

Shared Notes

Notes on Operational Conditions

- In most of these approaches, the assessed entropy will vary with the operational conditions of the entropy source.
- Raw data should be drawn from the noise source in various conditions.
- At minimum, “loaded” and “quiescent” are good targets.
  - “Loaded” may be I/O- or CPU-bound, depending on context.
- Entropy is estimated in each condition.
- The minimum entropy is reported.

Notes on Translation

- In many circumstances, the raw data will be too wide for direct analysis.
  - Internally, the raw data is 64 bits.
- The NIST tool only supports data up to 8 bits.
- Each 64-bit data value must be mapped down to 8 bits.
- There are many possible translation approaches.
  - See SP 800-90B Section 6.1 for one approach.
  - Can also group adjacent values.
  - Truncation / bit selection **is** translation.
- Some non-invasive translations can't impact any entropy estimator or IID test. Such translations are *injective* and *order-preserving*.
- Some *invasive* translations clearly interfere with the entropy estimators, and lead to artificially high entropy assessments.
  - Any non-injective translation fundamentally sums parts of the histogram.

Notes on the SP 800-90B Entropy Estimators

- These estimators are conceptually simple.
- Many estimators operate under an IID assumption.
- Many estimators essentially base their estimate on a single parameter.
- All of the estimators work better when supplied a fixed distribution.

Conclusion: Not magic boxes that output truth.

Notes on Sub-Distribution Approaches

- Oblivious approaches don't attempt to characterize sub-distributions.
  - This makes more substantial demands on the entropy estimators.
- Both the "Multiple" and "Single" approaches assess all observed sub-distributions.
  - Each identified sub-distribution is simpler, thus more likely to be accurately assessed.
- The "Single" sub-distribution approaches only treat data from a selected sub-distribution as output from the noise source.

Notes on the Sub-Distribution-Oblivious Approaches

In the “Sub-Distribution-Oblivious” approaches, no explicit attempt is made to separately assess identified sub-distributions.

- The full data sets (non-separated) are statistically assessed using the NIST tool.
- If the statistical assessment approach reasonably estimates min entropy of the distribution AND the distribution is stable, an attacker shouldn’t be able to do better.

Notes on the Multiple Sub-Distribution Approaches

In the “Multiple Sub-Distribution” approaches, each identified sub-distribution is separately assessed using the NIST tool.

- The full data sets (non-separated) are also assessed.
- The reported value is the minimum assessment from any identified sub-distribution, and the full data set.
- If the statistical assessment approach reasonably estimates min entropy of the sub-distributions AND all the sub-distributions are assessed, an attacker shouldn’t be able to do better.

Notes on the Single Sub-Distribution Approaches

In the “Single Sub-Distribution” approaches a single sub-distribution is identified, only data from that sub-distribution is used as raw data, and only data from the selected sub-distribution is assessed using the NIST tool.

- Other data may be treated as “supplemental data” (when using vetted conditioning functions).
- This yields a higher assurance assessment, as data from unexamined sub-distributions are not credited as containing entropy.
- This approach likely requires code changes to the baseline JEnt library.
- This approach is operationally less robust, as conditions may prevent the assessed sub-distribution from occurring.
  - E.g., various levels of power saving, more successful caching, better branch prediction or pipelining, etc.
- If the statistical assessment approach reasonably estimates min entropy of the identified sub-distribution, an attacker shouldn’t be able to do better.

Notes on the Essentially-IID Approaches

- Non-IID sources have statistical memory (internal state that induces relationships between the current output and some number of past outputs).
- The statistical memory "depth" is the number of symbols for which that state induces a significant interrelationship.
- If the memory depth is finite, we can decimate (throw away) enough data so that the remaining data acts like IID data.
  - "Thrown away" data can still be integrated into the conditioner as "supplemental data" (for vetted conditioning functions) and not credited as containing entropy.

Notes on the Essentially-IID Approaches

How do we know when we've thrown away enough data?

- Essentially this method involves running the SP 800-90B IID tests (section 5) a lot of times.
  - Take many samples of data.
  - Run all the IID tests on each of the data samples.
  - Check to see if each of the IID tests is passing "sufficiently often".

Repeat for each decimation level until it works. (if it works…)

Notes on the Essentially-IID Approaches

What is "sufficiently often"?

- There are 22 independent IID tests.
- We want some specific test false positive rate, say α.
- The per-test false positive rate, q, is thus

$$q \leq 1 - (1 - \alpha)^{1/22}$$

- If we've conducted n rounds of IID testing (each requiring a separate data set) then we can tolerate CDF^{-1} (BinomialDistribution(n rounds, $p = 0.001$), $1 - q$) allowed failures for each of the 22 IID tests.
- For $\alpha = 0.01$, we have $q \leq 0.00046$

Testing Rounds	Allowed Failures Per Test
32	2
147	3
348	4

Notes on the Essentially-IID Approaches

- Once you have decimated sufficiently, you can estimate entropy of the decimated data using the MCV estimator.
  - If osr is being used to cause decimation, you must divide the estimate by the decimation rate for $H_{\text{submitter}}$.
- It is probably best (and likely required) that you still not make an IID claim in the Entropy Analysis Report.
  - There is no general-purpose design-oriented reason this is an IID source.
- This approach requires that the memory is finite. This need not be the case.
- If translation causes apparent entropy to become close to 8, then the result will look IID (even if it actually isn't).
  - Mainly an issue for versions with pseudorandom behavior.
- Some systems require absurd levels of decimation.
  - osr > 20 causes problems with some interfaces.
- Some systems evidently cannot be suitably decimated.

Notes on Pseudorandom Variation

- Older versions of JEnt (by default, prior to 3.0.2) pseudorandomly vary the number of memory and conditioning rounds.
- This pseudorandom variation can't contribute entropy but does make the result pseudorandom.
- Any empirical (i.e., data-based) heuristic entropy estimation strategy used with this design must account for this pseudorandom variation.
- The most straightforward way to do this is to disable the pseudorandom "shuffle" functionality.

Approaches

Approach Summary

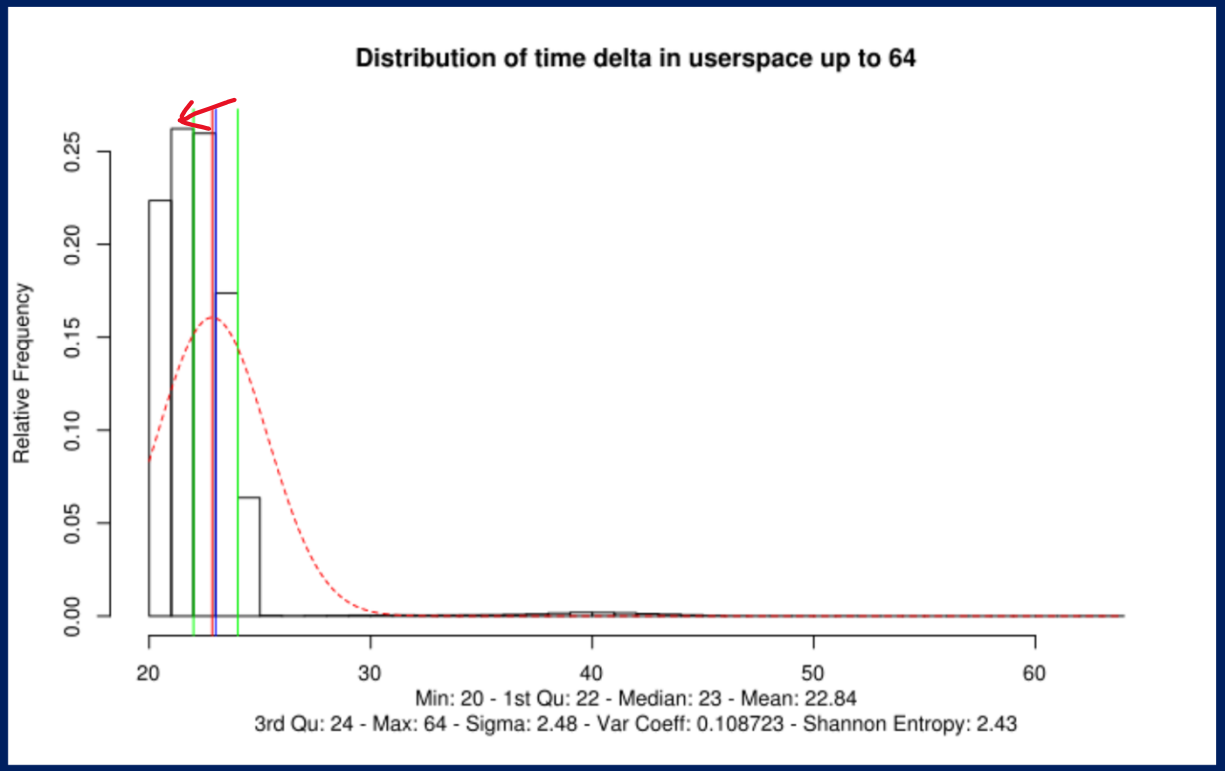
Reminder:

- The “Sub-Distribution Oblivious” approaches (1, 2, 3) are approaches that do not perform sub-distribution-based analysis.
- The “Multiple Sub-Distributions” approaches (4, 5) characterize the observed sub-distributions and establish sub-distribution-specific assessments.
- The “Single Sub-Distribution” approaches (6, 7) only credit entropy to samples from one identified sub-distribution.

	Sub-Distribution-Oblivious	**Multiple Sub-Distributions**	**Single Sub-Distribution**
Presumed IID	1		
Empirical	2	4	6
Essentially IID	3	5	7

Approach 8 is distinct.

Approach 1
Sub-Distribution-Oblivious Presumed IID Analysis



From [JEnt 2022]

Approach 1
Sub-Distribution-Oblivious Presumed IID Analysis

Steps:

- May or may not translate.
- If we make the assumption that the distribution is IID, then we find the most probable symbol, $p_{\max}$, and

$$H = -\log_2 p_{\max}$$

Approach 1
Sub-Distribution-Oblivious Presumed IID Analysis

Pros:

- It's certainly straightforward!
- This has some meaning even if the distribution is not IID.

Cons:

- For a non-IID distribution, this produces an upper bound for the min entropy, not a lower bound.
  - This is commonly a substantial overestimate for the entropy.
- The histogram may be rather complicated looking, particularly for JEnt libraries with pseudorandom variation.
- There are many ways that this could overestimate the entropy rate:
  - The attacker may force a behavior that wasn't assessed.
  - The noise source output may (and likely does) have long-term patterns that invalidate an IID assessment approach.

Approach 2
Sub-Distribution-Oblivious Empirical Analysis

Steps:

- Extract raw data from the noise source.
- Translate this data down to no more than 8 bits.
- Use the NIST SP 800-90B tool to generate an entropy assessment.

Approach 2
Sub-Distribution-Oblivious Empirical Analysis

Always produces an entropy estimate less than or equal to the one created using Approach 1.

Pros:

- Very straightforward!

Cons:

- Invasive translation is likely required.
  - Such translation likely reduces the meaningfulness of the analysis.
- Multiple sub-distributions are likely to occur.
  - The resulting composite distribution is very complicated.
  - The entropy assessment likely isn't very meaningful (for most translations).
- It isn't clear that any of the (non-prediction) entropy estimates directly apply.
- There are some ways that this could overestimate the entropy rate:
  - The attacker may force a behavior that wasn't assessed.
  - Any invasive translation may obscure patterns that ought to have reduced the entropy assessment.
  - The entropy estimators may not be adequate to assess the (likely quite complicated) distribution.

Approach 3
Sub-Distribution-Oblivious Essentially-IID Analysis

Steps:

- Extract raw data from the noise source.
- Translate this data down to no more than 8 bits.
- Establish an effective decimation rate.
- Set `osr` or decimation based on this rate.
- Use the NIST SP 800-90B MCV test tool to generate an entropy assessment.

Approach 3
Sub-Distribution-Oblivious Essentially-IID Analysis

Always produces an entropy rate less than or equal to the bound created by Approach 1.

- Usually results in a lower per-sample entropy bound than the bound created by Approach 2.

Pros:

- Inducing IID-like behavior makes the entropy level easier to reliably assess.

Cons:

- Presumes that the SP 800-90B IID tests are sensitive to the particular non-IID behavior of the source.
- Translation likely reduces the meaningfulness of the IID analysis.
- When multiple sub-distributions occur
  - The resulting distribution is very complicated.
  - It is easy to "saturate" apparent entropy, thus getting an artificial IID result.
- There are some ways that this could overestimate the entropy rate:
  - The attacker may force a behavior that wasn't assessed.
  - IID testing may not be sensitive to all non-IID behavior present, so decimation may not be adequate.

Approach 4
Multiple Sub-Distribution Empirical Analysis

Steps:

- Extract raw data from the noise source.
- Identify the sub-distributions.
- Generate data-subsets separating the identified sub-distributions.
- Independently translate all the data sets down to no more than 8 bits.
  - The sub-distributions necessarily have a subset of the symbols, so may not require invasive translation.
- Use the NIST SP 800-90B tool to generate entropy assessments for each of the sub-distributions and the whole data set.

Approach 4
Multiple Sub-Distribution Empirical Analysis

Always less than or equal to the bound created by Approach 2.

Pros:

- Sub-distributions can be easier to assess.
- Non-invasive translation may be sufficient for the sub-distributions.

Cons:

- Identification of sub-distributions is manual.
- It is important to identify all of the possible sub-distributions that an attacker could induce.
- It isn't clear that any of the (non-prediction) entropy estimates directly apply.
- There are some ways that this could overestimate the entropy rate:
  - The attacker may force a behavior that wasn't assessed.
  - Any invasive translation may obscure patterns that ought to have reduced the entropy assessment.
  - The entropy estimators may not be adequate to assess the (simplified, but possibly still complicated) distribution.

Approach 5
Multiple Sub-Distribution Essentially-IID Analysis

Steps:

- Extract raw data from the noise source.
- Translate the data sets down to no more than 8 bits.
- Establish an effective decimation rate.
- Set `osr` or decimation based on this rate.
- Identify the sub-distributions.
- Independently translate all the data sets data down to no more than 8 bits.
  - The sub-distributions necessarily have a subset of the symbols, so may not require invasive translation.
- Generate data-subsets separating the identified sub-distributions.
- Use the NIST SP 800-90B MCV test tool to generate an entropy assessment for each sub-distribution and for the overall data set.

Approach 5
Multiple Sub-Distribution Essentially-IID Analysis

Always less than or equal to the bound created by Approach 3.

Pros:

- Sub-distributions can be easier to assess.
- Non-invasive translation may be sufficient for the sub-distributions.

Cons:

- Identification of sub-distributions is manual.
- It is important to identify all of the possible sub-distributions that an attacker could induce.
- There are some ways that this could overestimate the entropy rate:
  - The attacker may force a behavior that wasn't assessed.
  - IID testing may not be sensitive to all non-IID behavior present, so decimation may not be adequate.

Approach 6
Single Sub-Distribution Empirical Analysis

Steps:

- Extract raw data from the noise source.
- Identify the sub-distributions. Designate one of them the sub-distribution of interest.
- Separate out the designated sub-distribution.
- Translate the designated sub-distribution down to no more than 8 bits.
  - The designated sub-distribution necessarily has a subset of the symbols, so may not require invasive translation.
- Use the NIST SP 800-90B tool to generate entropy assessments for the designated sub-distribution.

Approach 6
Single Sub-Distribution Empirical Analysis

Produces a per-symbol entropy assessment greater than or equal to Approach 4.

Pros:

- The designated sub-distribution can be easier to assess, so the tool output is more likely to be meaningful.
- Non-invasive translation may be sufficient for the designated sub-distribution.
- The attacker may be able to reduce the data rate, but (if the designated sub-distribution is well chosen) they should not be able to reduce the entropy-per-symbol rate.
  - This is a “fail secure” style design.

Cons:

- Identification of sub-distributions is manual.
- If conditions shift, the output data rate may fall precipitously.
- There are some ways that this could overestimate the entropy rate:
  - Any invasive translation may obscure patterns that ought to have reduced the entropy assessment.
  - The entropy estimators may not be adequate to assess the (simplified, but possibly still complicated) sub-distribution.

Approach 7
Single Sub-Distribution Essentially-IID Analysis

Steps:

- Extract raw data from the noise source.
- Identify the sub-distributions. Designate one of them the sub-distribution of interest.
- Separate out the designated sub-distribution.
- Translate the designated sub-distribution down to no more than 8 bits.
  - The designated sub-distribution necessarily has a subset of the symbols, so it may not require invasive translation.
- Establish an effective decimation rate.
- Set `osr` or decimation based on this rate.
- Use the NIST SP 800-90B MCV test tool to generate an entropy assessment.

Approach 7
Single Sub-Distribution Essentially-IID Analysis

Produces a per-symbol entropy assessment greater than or equal to Approach 5.

Pros:

- IID behavior makes the entropy level easier to reliably assess.
- Non-invasive translation may be sufficient for the designated sub-distribution.
- The attacker may be able to reduce the data rate, but if the designated sub-distribution is well chosen they should not be able to reduce the entropy-per-symbol rate.
  - This a “fail secure” style design.

Cons:

- Identification of sub-distributions is manual.
- If conditions shift, the output data rate may fall precipitously.
- Presumes that the SP 800-90B IID tests are sensitive to the particular non-IID behavior of the source in the designated sub-distribution.
- There are some ways that this could overestimate the entropy rate:
  - IID testing may not be sensitive to all non-IID behavior present, so decimation may not be adequate.

Approach 8
(Hypothetical) Single Sub-Distribution Stochastic Model

Steps:

- For a particular (hyper-specific) piece of hardware, develop an abstracted stochastic model for an identified source of entropy in the system.
  - e.g., relative jitter between different clocks in a clock tree.
- Model the impact of only the identified phenomena, and use the stochastic model to produce an $H_{\text{submitter}}$ value based on this model.
- This $H_{\text{submitter}}$ value could only apply to a specific hardware instance (fixed architecture and configuration).
- This would require a substantial amount of effort for each individual configuration.

Approach 8
(Hypothetical) Single Sub-Distribution Stochastic Model

Pros:

- A high level of assurance for the claimed min entropy.

Cons:

- Very labor intensive, and the result is profoundly fragile.
  - e.g., Changing the particular memory part could completely undermine the stochastic model.
- Presently hypothetical.