Combinatorial Coverage and Distance Measurements of Test Sets

Dimitris E. Simos, Manuel Leithner, Rick Kuhn, Raghu Kacker

COMET
Competence Centers for Excellent Technologies
www.ffg.at/comet

SBA Research

MATRIS

Key Facts

Use Cases

- Measure coverage of existing test sets.
- Verify coverage of constructed Covering Arrays.
- Qualitative comparison of test sets with identical coverage with additional distribution and distance analysis.

Architecture

- Rust implementation:
  - ⇒ Native Executable (Linux, Windows, macOS)
  - ⇒ WebAssembly
- Modular input parsing.
- Machine- or human-readable output.

Implementation Highlights of CAmetrics

- Significantly fast based on experiments:
  - ⇒ Multi-threaded
  - ⇒ Multiple algorithms
- Low memory usage.
- Web UI and command line interface.
- Sophisticated constraint support.

Future improvements

- HPC and cluster implementations.
- Faster constraint processing.
- α-balance.

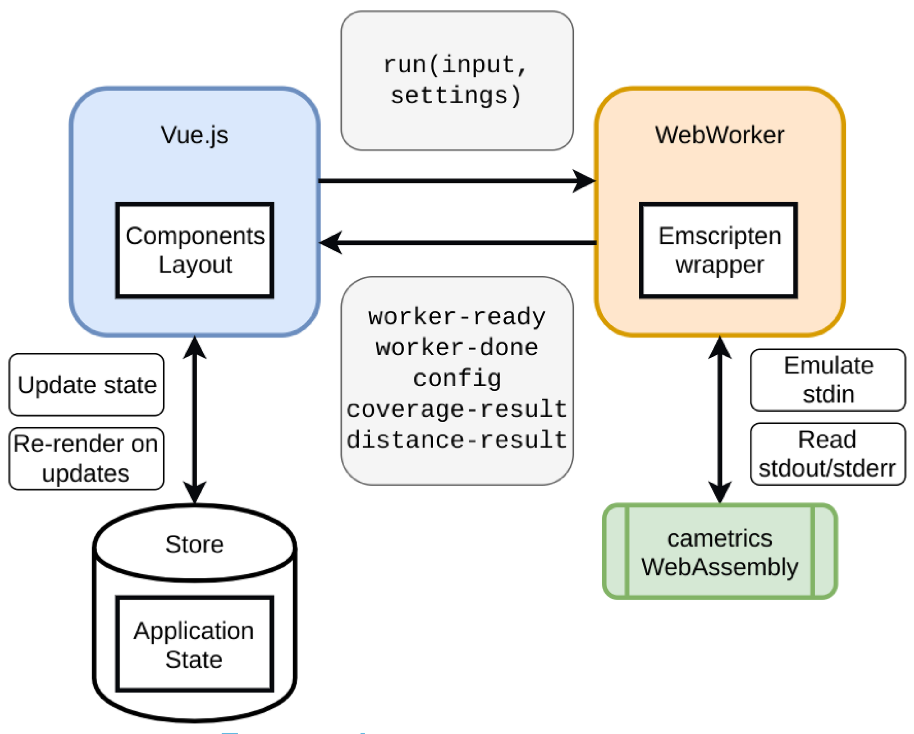

Figure 1: Web UI architecture

Coverage

- Simple t-way combination coverage: How many t-selections of parameters are fully covered?
- Simple $(t+1)$-way coverage: $t + \frac{\text{Covered } (t+i)\text{-tuples}}{\text{Total } (t+i)\text{-tuples}}$

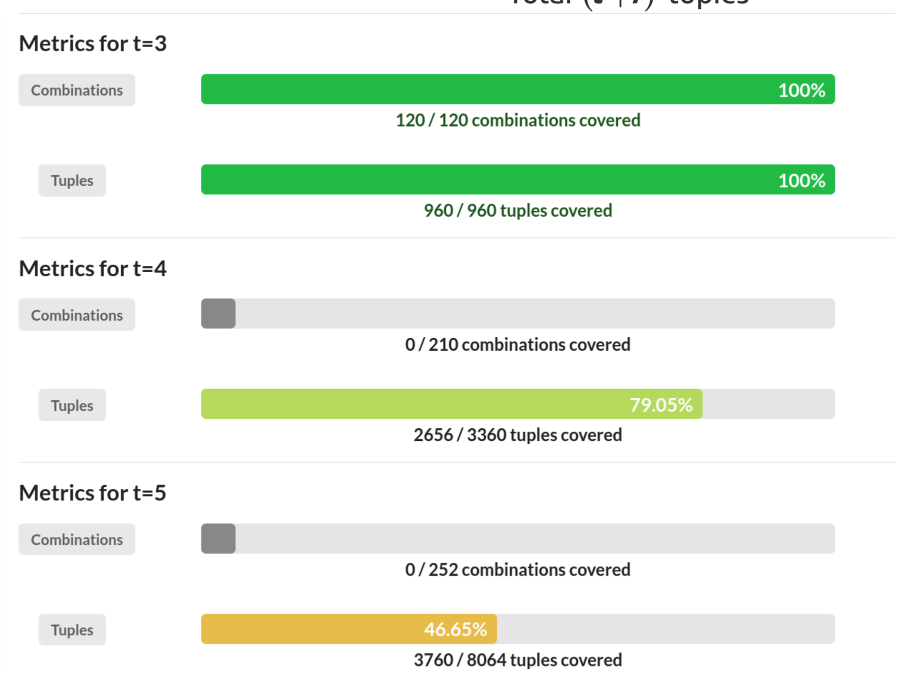

Figure 2: Simple t-way combination and tuple coverage

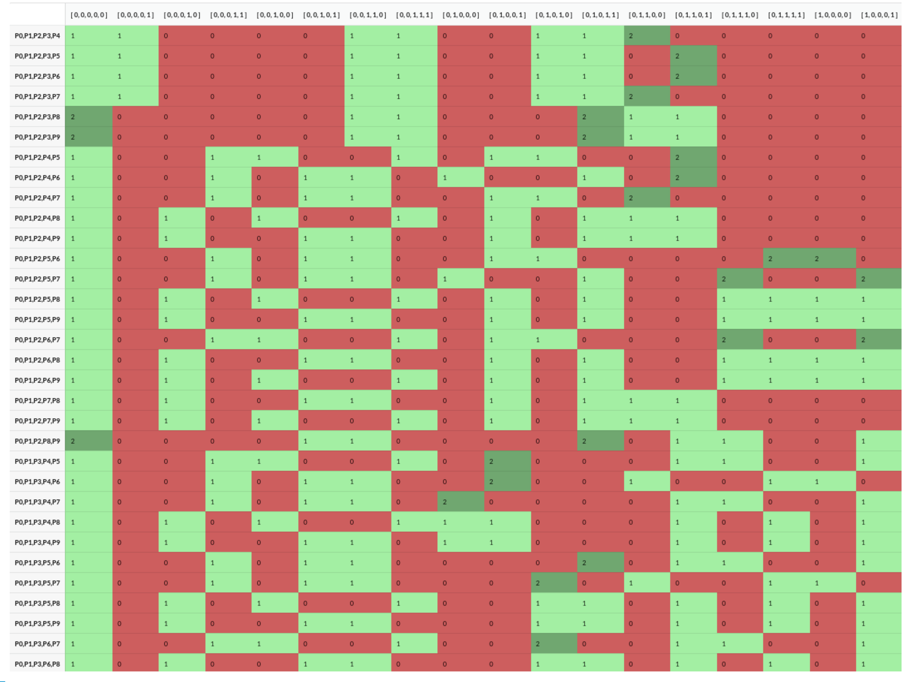

Figure 3: Coverage Map showing t-tuple coverage per parameter selection

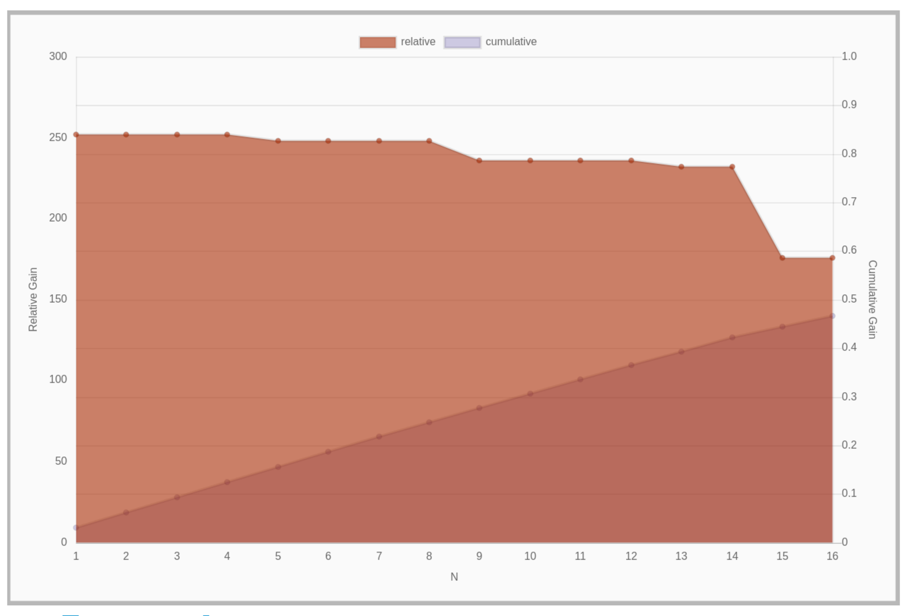

Figure 4: Per-test and cumulative coverage gain.
The coverage gain describes how many previously uncovered tuples are covered by a single test.

Select between Two Algorithms

Input: $N \times k$ array, alphabet size v, strength t

`cametrics-fast`

- Great general-purpose performance
- $\mathcal{O}(\binom{k}{t} v^t)$ memory
- $\mathcal{O}(\binom{k}{t} N)$ time

`cametrics-light`

- Prevents memory explosion, great for binary/small arrays
- $\mathcal{O}(\binom{k}{t})$ memory
- $\mathcal{O}(\binom{k}{t} N v^t)$ time

Distance Metrics

Inter-test distance: baseline for success

- (Generalized) Hamming Distance: How many parameter values differ between tests?
- Total Cartesian/Euclidian Distance of test T to array A:

$$CD(A,T) = \sum_{n=1}^{|A|} \sqrt{\sum_{i=1}^{k} (A_{ni} - T_i)^2}$$

Balance is everything!

- *Balanced* array: Each parameter value (or t-tuple) appears roughly the same number of times.

Modified χ^2 Distance:

- How close to ideal distribution of parameter values?

Ideal distribution: $D'_{ij} = \frac{1}{v_i} | i \in \{1, \ldots, k\}, j \in V_i$

Actual distribution:

$$D_{ij} = \frac{1}{N} \sum_{n=1}^{N} \begin{cases} 1 & \text{if } A_{ni} = V_{ij} \\ 0 & \text{otherwise.} \end{cases}$$

Modified χ^2 Distance: $\left(\sum_{i=1}^{k} \frac{v_i - 1}{v_i + 1}\right) - \frac{\sum_{i=1}^{k} \sum_{j=1}^{v_i} \frac{(D_{ij} - D'_{ij})^2}{D_{ij} + D'_{ij}}}{2}$

```
960 tuples to cover in total, 120 max coverage gain per test
--------------------------- snip ---------------------------
[1, 1, 1, 1, 0, 0, 0, 0, 0, 0] i = 10 H_min = 4 H_max = 8 H_sum = 48 CD_sum = 20.6717 X^2 = 3.31313131 848 covered (88.33%), 40 gain
[0, 1, 0, 1, 1, 0, 1, 0, 1, 1] i = 11 H_min = 4 H_max = 6 H_sum = 52 CD_sum = 22.7422 X^2 = 3.31262940 880 covered (91.67%), 32 gain
[1, 0, 1, 0, 0, 1, 0, 1, 1, 1] i = 12 H_min = 4 H_max = 8 H_sum = 60 CD_sum = 25.5707 X^2 = 3.33333333 912 covered (95.00%), 32 gain
[0, 0, 1, 1, 1, 1, 0, 0, 1, 0] i = 13 H_min = 4 H_max = 6 H_sum = 60 CD_sum = 26.7875 X^2 = 3.31851852 928 covered (96.67%), 16 gain
[1, 1, 0, 0, 0, 0, 1, 1, 1, 0] i = 14 H_min = 4 H_max = 8 H_sum = 68 CD_sum = 29.6160 X^2 = 3.32307692 944 covered (98.33%), 16 gain
[0, 1, 1, 0, 0, 1, 1, 0, 0, 1] i = 15 H_min = 2 H_max = 10 H_sum = 72 CD_sum = 31.3640 X^2 = 3.32220986 952 covered (99.17%), 8 gain
[1, 0, 0, 1, 1, 0, 0, 1, 0, 1] i = 16 H_min = 2 H_max = 10 H_sum = 80 CD_sum = 34.1924 X^2 = 3.33333333 960 covered (100.00%), 8 gain
Optimal X^2 value: 3.3333333333333335
```

Figure 5: Distance measurements: Minimum/Maximum/Total Hamming Distance, Total Cartesian Distance, χ^2, and coverage

D Richard Kuhn, Itzel Dominguez Mendoza, Raghu N Kacker, and Yu Lei. Combinatorial coverage measurement concepts and applications. In *2013 IEEE Sixth International Conference on Software Testing, Verification and Validation Workshops*, pages 352–361. IEEE, 2013.
Manuel Leithner, Kristoffer Kleine, and Dimitris E Simos. CAmetrics: A tool for advanced combinatorial analysis and measurement of test sets. In *2018 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, pages 318–327. IEEE, 2018.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung wirkt.
wirtschafts agentur wien Ein Fonds der Stadt Wien
FWF Der Wissenschaftsfonds.
netidee OPEN INNOVATIONS

SBA Research (SBA-K1) is a COMET Centre within the framework of COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG.