Large-scale Combinatorial Testing with Adobe

Dimitris E. Simos, Manuel Leithner, Ludwig Kampel, Darryl Jarman,
Riley Smith, Jared Bellows, Rick Kuhn, Raghu Kacker

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

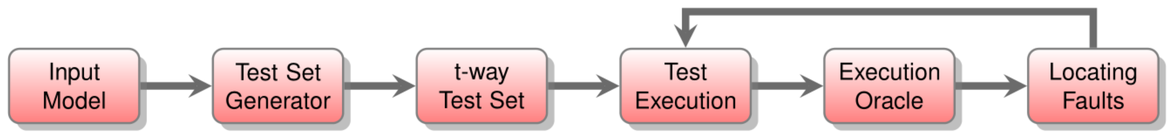
SBA Research

MATRIS

Combinatorial Testing

Combinatorial Testing

- Combinatorial Testing allows for efficient testing of large systems while maintaining certain coverage guarantees.
- In a combinatorial test set, every t-way interaction appears in at least one tests, where t is called the strength of the test set.
- Studies have shown that the majority of faults can be detected with combinatorial test sets of strength $t \leq 6$.



Challenges for Large-scale Combinatorial Testing:

- Modelling of the system under test.
- Construction of reasonably small combinatorial test sets for systems with a large number of parameters is a difficult optimization problem.
- Sufficient Input Bandwidth necessary.

Generation of Large Combinatorial Test Sets

Greedy test set generation with CAgen:

- Generate test sets with a small number of rows with the FIPOG algorithm
- Significantly faster generation than other state of the art tools

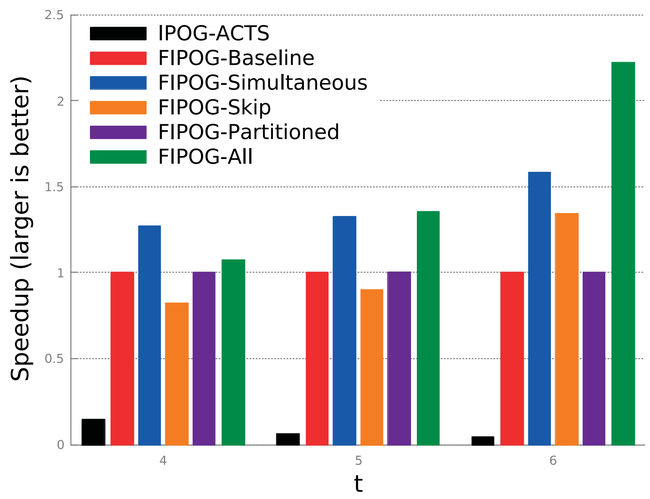


Figure 1: Speedup of CAgen compared to the state-of-the-art tool ACTS

- However: For the generation of combinatorial test sets for large systems, using greedy algorithms is no longer feasible.

Roux-type construction:

- Covering Arrays (CAs) are the mathematical structure underlying combinatorial test sets.
- Doubles the number of columns of a given CA.
- Concatenates permutations of CAs of lower strength to ensure coverage of all t-way interactions.

$$\mathsf{E}_1 = \begin{pmatrix} \mathsf{A} \\ \mathsf{A} \end{pmatrix},$$

$$\mathsf{E}_2 = \begin{pmatrix} \mathsf{B} & \mathsf{B} & \cdots & \mathsf{B} \\ \mathsf{B}^{\pi^1} & \mathsf{B}^{\pi^2} & \cdots & \mathsf{B}^{\pi^{v-1}} \end{pmatrix},$$

$$\mathsf{E}_3 = \begin{pmatrix} \mathsf{C} & \mathsf{C} & \cdots & \mathsf{C} & \mathsf{C}^{f_1} & \mathsf{C}^{f_2} & \cdots & \mathsf{C}^{f_{N_1}} \\ \mathsf{C}^{f_1} & \mathsf{C}^{f_2} & \cdots & \mathsf{C}^{f_{N_1}} & \mathsf{C} & \mathsf{C} & \cdots & \mathsf{C} \end{pmatrix},$$

$$\mathsf{E}_4 = \begin{pmatrix} \mathsf{C} & \mathsf{C} & \cdots & \mathsf{C} & \mathsf{C}^{g_{1,2}} & \mathsf{C}^{g_{1,3}} & \cdots & \mathsf{C}^{g_{v,v-1}} \\ \mathsf{C}^{g_{1,2}} & \mathsf{C}^{g_{1,3}} & \cdots & \mathsf{C}^{g_{v,v-1}} & \mathsf{C} & \mathsf{C} & \cdots & \mathsf{C} \end{pmatrix},$$

$$\mathsf{E}_5 = \begin{pmatrix} \mathsf{C} & \mathsf{C} & \cdots & \mathsf{C} & \mathsf{C}^{\tilde{g}_{1,2}} & \mathsf{C}^{\tilde{g}_{1,3}} & \cdots & \mathsf{C}^{\tilde{g}_{v,v-1}} \\ \mathsf{C}^{\tilde{g}_{1,2}} & \mathsf{C}^{\tilde{g}_{1,3}} & \cdots & \mathsf{C}^{\tilde{g}_{v,v-1}} & \mathsf{C} & \mathsf{C} & \cdots & \mathsf{C} \end{pmatrix},$$

$$\mathsf{E}_6 = \begin{pmatrix} \mathsf{C} & \mathsf{C} & \cdots & \mathsf{C} & \mathsf{C}^{h_{1,2}} & \mathsf{C}^{h_{1,3}} & \cdots & \mathsf{C}^{h_{v,v-1}} \\ \mathsf{C}^{h_{1,2}} & \mathsf{C}^{h_{1,3}} & \cdots & \mathsf{C}^{h_{v,v-1}} & \mathsf{C} & \mathsf{C} & \cdots & \mathsf{C} \end{pmatrix}.$$

Two-stage approach:

- Combines greedy methods with a mathematical doubling construction.
- CAgen generates seed arrays with as many columns as possible.
- Roux-type construction extends the seed arrays to the desired combinatorial test sets.

Large-scale Combinatorial Testing at Adobe

- Collaboration with Adobe.
- Application of largest combinatorial test sets documented in research.
- We generated test sets for models with more than 2000 parameters and 10 values:

N	t	k	v
6337	4	2127	3
107514	5	2127	3
87669	5	2127	3
322	2	2127	7
7439	3	2127	7
688	2	2127	10
23422	3	2127	10

Adobe

Adobe Analytics

Testing Results

- New faults found in each subject system.

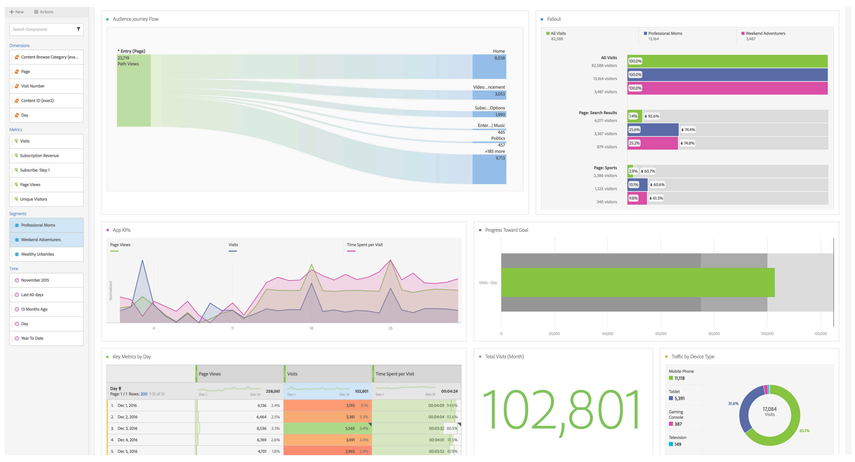


Table I. REZZER FAULTS DETECTED BY COMBINATORIAL TESTING

Fault Descriptions, Causes, and Resolutions		
Description	***Cause***	***Resolution***
Flag-type fields throw error	Undocumented value constraint	Update input space model
Event-type fields throw error	Undocumented format constraint	Update input space model
Parser throws error (CDS)	Undocumented value constraint	Update input space model
Parser throws error (JSON)	Undocumented format constraint	Add input validation
Invalid date fields interaction	Undocumented value constraint	Update input space model

Darryl Jarman, Riley Smith, Gabe Gosney, Ludwig Kampel, Manuel Leithner, Dimitris E. Simos, Raghu Kacker, and Rick Kuhn. Applying combinatorial testing to large-scale data processing at adobe. In *2019 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, pages 190–193, 2019.
Kristoffer Kleine and Dimitris E. Simos. An efficient design and implementation of the in-parameter-order algorithm. *Mathematics in Computer Science*, 12(1):51–67, 2018.
Riley Smith, Darryl Jarman, Jared Bellows, Richard Kuhn, Raghu Kacker, and Dimitris Simos. Measuring combinatorial coverage at adobe. In *2019 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, pages 194–197, 2019.
Michael Wagner, Kristoffer Kleine, Dimiris E. Simos, Richard Kuhn, and Raghu Kacker. Cagen : A fast combinatorial test generation tool with support for constraints and higher-index arrays. In *To be published in Proceedings of IWCT 2020*. Springer International Publishing, 2020.

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie
Bundesministerium Digitalisierung und Wirtschaftsstandort
FFG Forschung wirkt.
wirtschaftsagentur wien Ein Fonds der Stadt Wien
Europäische Kommission
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.