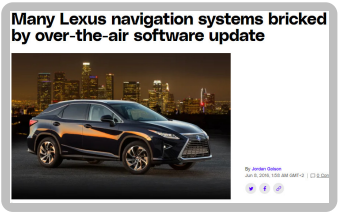
Covering T-Wise Interactions of Deployed Configurations

Rahel Sundermann, Sabrina Boehm, **Sebastian Krieter**, Malte Lochau, Thomas Thüm | VaMoS 2025 | February 05

Software Engineering
Programming Languages

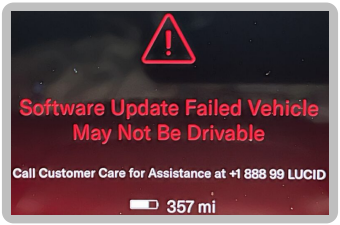
universität ulm

Motivation Software Updates



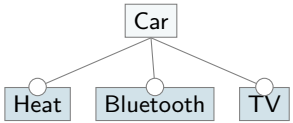
Google Fixes Pixel Update Error Leading to Overheating, Battery Drain Issues

The issue appeared to affect Pixel 6 and 7 series devices, although it seems to be resolved.



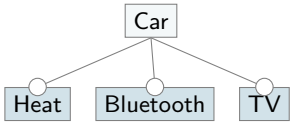
Motivation Testing a Product

Product Line	#Features	#Configurations
Automotive01	2,513	$\approx 5.4 \times 10^{217}$



Motivation Testing a Product

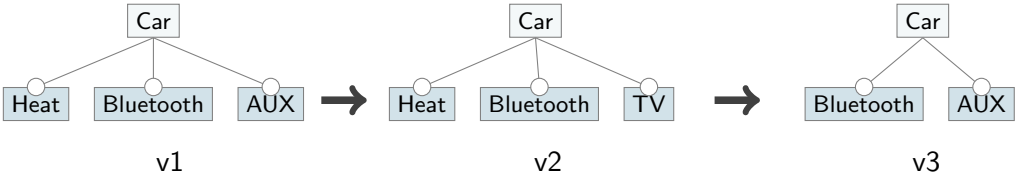
Product Line	#Features	#Configurations	Pair-wise Sample Size
Automotive01	2,513	$\approx 5.4 \times 10^{217}$	1,167



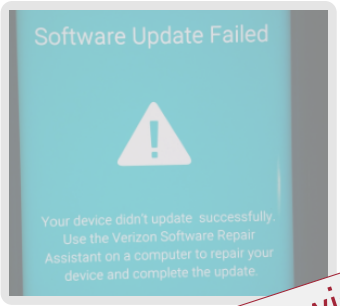
$$
\begin{aligned}
&\mathit{Bluetooth} \wedge \quad \mathit{Heat} \\
&\mathit{Bluetooth} \wedge \neg \mathit{Heat} \\
\neg &\mathit{Bluetooth} \wedge \quad \mathit{Heat} \\
\neg &\mathit{Bluetooth} \wedge \neg \mathit{Heat} \\
&\mathit{Bluetooth} \wedge \quad \mathit{TV} \\
&\ldots
\end{aligned}
$$

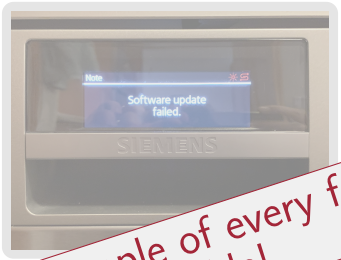
Motivation Testing a Product

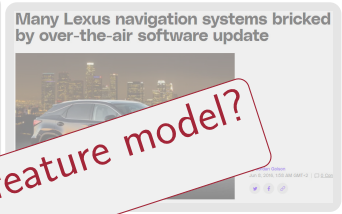
Product Line	#Features	#Configurations	Pair-wise Sample Size
Automotive01_v1	2,513	$\approx 5.4 \times 10^{217}$	1,167
Automotive01_v2	2,520	$\approx 5.4 \times 10^{217}$	1,185
Automotive01_v3	2,533	$\approx 5.4 \times 10^{217}$	1,172

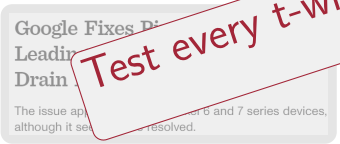


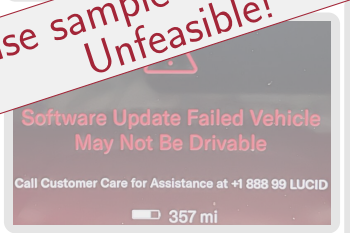
Motivation Testing Multiple Versions





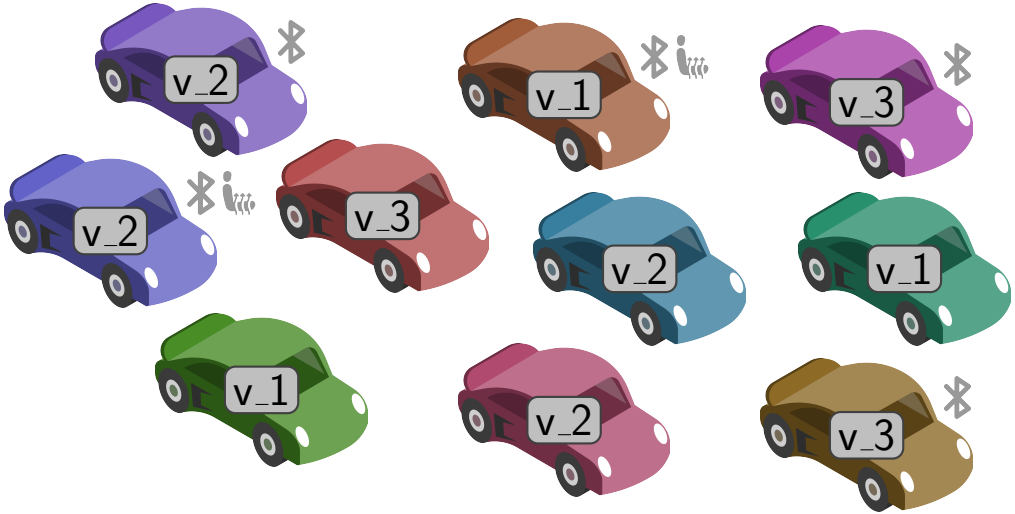




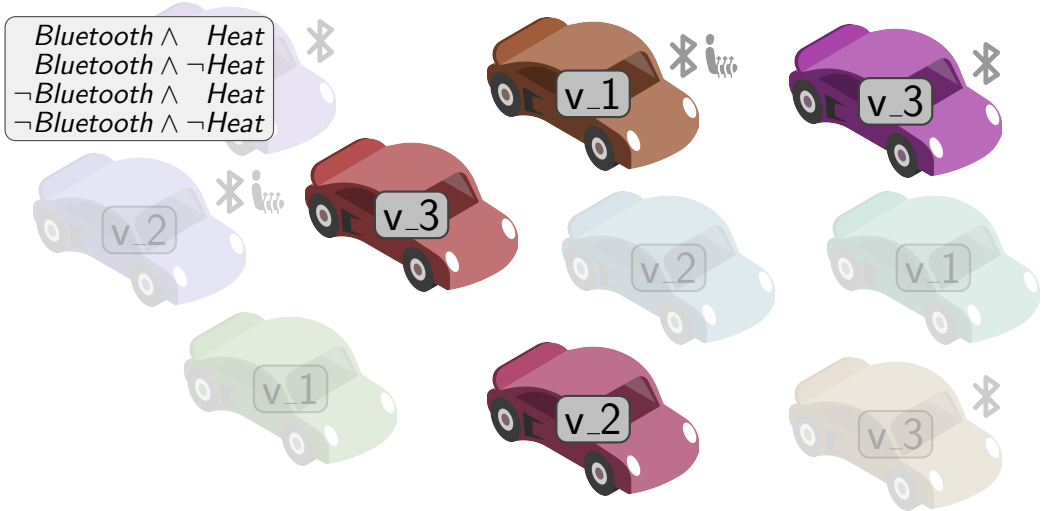


Test every t-wise sample of every feature model?
Unfeasible!

Idea



Idea


Bluetooth ∧ Heat
Bluetooth ∧ ¬Heat
¬Bluetooth ∧ Heat
¬Bluetooth ∧ ¬Heat
v_1
v_3
v_2
v_3
v_2
v_1
v_1
v_2
v_3

T-Wise Field Coverage

How to measure coverage of **multiple model versions** and only for **deployed configurations**?

T-Wise Field Coverage

How to measure coverage of **multiple model versions** and only for **deployed configurations**?

$\Rightarrow$ Analogous to t-wise interaction sample: ratio of covered interaction to total interactions

T-Wise Field Coverage

How to measure coverage of **multiple model versions** and only for **deployed configurations**?

$\Rightarrow$ Analogous to t-wise interaction sample: ratio of covered interaction to total interactions

Field coverage for any $\mathcal{S} \subset \mathcal{S}_{Field}$

$$coverage^t(\mathcal{S}, \mathcal{S}_{Field}) = \frac{|\{i \in I^t : \exists c \in \mathcal{S}\}|}{|\{i \in I^t : \exists c \in \mathcal{S}_{Field}\}|}$$

features over all versions: $L = \bigcup_{c \in \mathcal{S}_{Field}} c$

t-wise interactions: $I^t = \{i \subset L : |i| = t\}$

Greedy Algorithm for T-Wise Field Sampling

Input: Set of field configurations $\mathcal{S}_{Field} = \{c_1, \ldots, c_n\}$

Output: Subset of $\mathcal{S}_{Field}$ with 100% field coverage

Greedy Algorithm for T-Wise Field Sampling

Input: Set of field configurations $\mathcal{S}_{Field} = \{c_1, \ldots, c_n\}$

Output: Subset of $\mathcal{S}_{Field}$ with 100% field coverage

1. Start with empty sample $\mathcal{S}'$

1. $\mathcal{S}' = \emptyset$

Greedy Algorithm for T-Wise Field Sampling

Input: Set of field configurations $\mathcal{S}_{Field} = \{c_1, \ldots, c_n\}$

Output: Subset of $\mathcal{S}_{Field}$ with 100% field coverage

1. Start with empty sample $\mathcal{S}'$
2. Count occurrences of each t-wise interaction

1. $\mathcal{S}' = \emptyset$
2. $count(i) = |\{c \in \mathcal{S}_{Field} : i \subseteq c\}|$

Greedy Algorithm for T-Wise Field Sampling

Input: Set of field configurations $\mathcal{S}_{Field} = \{c_1, \ldots, c_n\}$

Output: Subset of $\mathcal{S}_{Field}$ with 100% field coverage

1. Start with empty sample $\mathcal{S}'$
2. Count occurrences of each t-wise interaction
3. Add configurations with unique interactions to $\mathcal{S}'$

1. $\mathcal{S}' = \emptyset$
2. $count(i) = |\{c \in \mathcal{S}_{Field} : i \subseteq c\}|$
3. $\mathcal{U} = \{c \in \mathcal{S}_{Field} : \exists i \subseteq c : count(i) = 1\}$
   $\mathcal{S}' = \mathcal{S}' \cup \mathcal{U}$, $\mathcal{S}_{Field} = \mathcal{S}_{Field} \setminus \mathcal{U}$

Greedy Algorithm for T-Wise Field Sampling

Input: Set of field configurations $\mathcal{S}_{Field} = \{c_1, \ldots, c_n\}$

Output: Subset of $\mathcal{S}_{Field}$ with 100% field coverage

1. Start with empty sample $\mathcal{S}'$
2. Count occurrences of each t-wise interaction
3. Add configurations with unique interactions to $\mathcal{S}'$
    1. Update score for each configuration in $\mathcal{S}_{Field}$

1. $\mathcal{S}' = \emptyset$
2. $count(i) = |\{c \in \mathcal{S}_{Field} : i \subseteq c\}|$
3. $\mathcal{U} = \{c \in \mathcal{S}_{Field} : \exists i \subseteq c : count(i) = 1\}$
   $\mathcal{S}' = \mathcal{S}' \cup \mathcal{U}$, $\mathcal{S}_{Field} = \mathcal{S}_{Field} \setminus \mathcal{U}$
    1. $score(c) = \sum_{i \in I_t} \frac{1}{count(i)}$

Greedy Algorithm for T-Wise Field Sampling

Input: Set of field configurations $\mathcal{S}_{Field} = \{c_1, \ldots, c_n\}$

Output: Subset of $\mathcal{S}_{Field}$ with 100% field coverage

1. Start with empty sample $\mathcal{S}'$
2. Count occurrences of each t-wise interaction
3. Add configurations with unique interactions to $\mathcal{S}'$
   1. Update score for each configuration in $\mathcal{S}_{Field}$
   2. Add configuration with highest score to $\mathcal{S}'$

1. $\mathcal{S}' = \emptyset$
2. $count(i) = |\{c \in \mathcal{S}_{Field} : i \subseteq c\}|$
3. $\mathcal{U} = \{c \in \mathcal{S}_{Field} : \exists i \subseteq c : count(i) = 1\}$
   $\mathcal{S}' = \mathcal{S}' \cup \mathcal{U}$, $\mathcal{S}_{Field} = \mathcal{S}_{Field} \setminus \mathcal{U}$
   1. $score(c) = \sum_{i \in I_t} \frac{1}{count(i)}$
   2. $c_{max} = \arg\max_{c \in \mathcal{S}_{Field}} score(c)$
      $\mathcal{S}' = \mathcal{S}' \cup \{c_{max}\}$,
      $\mathcal{S}_{Field} = \mathcal{S}_{Field} \setminus \{c_{max}\}$

Greedy Algorithm for T-Wise Field Sampling

Input: Set of field configurations $\mathcal{S}_{Field} = \{c_1, \ldots, c_n\}$

Output: Subset of $\mathcal{S}_{Field}$ with 100% field coverage

1. Start with empty sample $\mathcal{S}'$
2. Count occurrences of each t-wise interaction
3. Add configurations with unique interactions to $\mathcal{S}'$
   1. Update score for each configuration in $\mathcal{S}_{Field}$
   2. Add configuration with highest score to $\mathcal{S}'$
4. Repeat until all interactions are contained in $\mathcal{S}'$

1. $\mathcal{S}' = \emptyset$
2. $count(i) = |\{c \in \mathcal{S}_{Field} : i \subseteq c\}|$
3. $\mathcal{U} = \{c \in \mathcal{S}_{Field} : \exists i \subseteq c : count(i) = 1\}$
   $\mathcal{S}' = \mathcal{S}' \cup \mathcal{U}$, $\mathcal{S}_{Field} = \mathcal{S}_{Field} \setminus \mathcal{U}$
   1. $score(c) = \sum_{i \in I_t} \frac{1}{count(i)}$
   2. $c_{max} = \arg\max_{c \in \mathcal{S}_{Field}} score(c)$
      $\mathcal{S}' = \mathcal{S}' \cup \{c_{max}\}$,
      $\mathcal{S}_{Field} = \mathcal{S}_{Field} \setminus \{c_{max}\}$

T-Wise Field Sampling: Example

T-Wise Field Sampling: Example

$Bluetooth \wedge \neg Heat : 2$ $\quad \neg Bluetooth \wedge \neg Heat : 5$

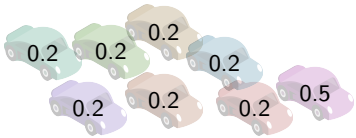
$Bluetooth \wedge \quad Heat : 1$ $\quad \neg Bluetooth \wedge \quad Heat : 0$

T-Wise Field Sampling: Example

$Bluetooth \wedge \neg Heat : 2$ $\quad \neg Bluetooth \wedge \neg Heat : 5$

$Bluetooth \wedge \quad Heat : 0$ $\quad \neg Bluetooth \wedge \quad Heat : 0$

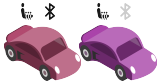
T-Wise Field Sampling: Example



$Bluetooth \wedge \neg Heat : 2$ $\quad \neg Bluetooth \wedge \neg Heat : 5$

$Bluetooth \wedge \quad Heat : 0$ $\quad \neg Bluetooth \wedge \quad Heat : 0$

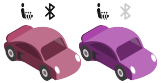
T-Wise Field Sampling: Example



$Bluetooth \wedge \neg Heat : 2$ $\quad \neg Bluetooth \wedge \neg Heat : 5$

$Bluetooth \wedge \quad Heat : 0$ $\quad \neg Bluetooth \wedge \quad Heat : 0$

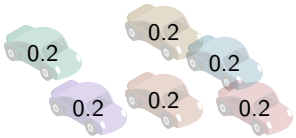
T-Wise Field Sampling: Example



$Bluetooth \wedge \neg Heat : 0$ $\quad \neg Bluetooth \wedge \neg Heat : 5$

$Bluetooth \wedge \quad Heat : 0$ $\quad \neg Bluetooth \wedge \quad Heat : 0$

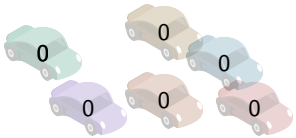
T-Wise Field Sampling: Example



$Bluetooth \wedge \neg Heat : 0$ $\quad \neg Bluetooth \wedge \neg Heat : 5$

$Bluetooth \wedge \quad Heat : 0$ $\quad \neg Bluetooth \wedge \quad Heat : 0$

T-Wise Field Sampling: Example



$Bluetooth \wedge \neg Heat : 0$ $\quad \neg Bluetooth \wedge \neg Heat : 0$

$Bluetooth \wedge \quad Heat : 0$ $\quad \neg Bluetooth \wedge \quad Heat : 0$

Evaluation Setup

Evaluation Setup

System	#Versions	#Features	#Field Config.
Soletta	173	457	11,950
Toybox	62	334	1,039
Busybox	454	1,050	31,279
Fiasco	33	258	10,872
Uclibc	177	272	110,524

Evaluation Setup

System	#Versions	#Features	#Field Config.
Soletta	173	457	11,950
Toybox	62	334	1,039
Busybox	454	1,050	31,279
Fiasco	33	258	10,872
Uclibc	177	272	110,524
ERP	1	1,920	61
Agrib	1	2,238	5,749

Evaluation Setup

System	#Versions	#Features	#Field Config.
Soletta	173	457	11,950
Toybox	62	334	1,039
Busybox	454	1,050	31,279
Fiasco	33	258	10,872
Uclibc	177	272	110,524
ERP	1	1,920	61
Agrib	1	2,238	5,749
Automotive08	unknown	221	233,309
Automotive09	unknown	279	247,780
Automotive10	unknown	279	45,387
Automotive11	unknown	218	27,808
Automotive12	unknown	147	129

Evaluation Setup

System	#Versions	#Features	#Field Config.
Soletta	173	457	11,950
Toybox	62	334	1,039
Busybox	454	1,050	31,279
Fiasco	33	258	10,872
Uclibc	177	272	110,524
ERP	1	1,920	61
Agrib	1	2,238	5,749
Automotive08	unknown	221	233,309
Automotive09	unknown	279	247,780
Automotive10	unknown	279	45,387
Automotive11	unknown	218	27,808
Automotive12	unknown	147	129

Algorithms

- YASA

Evaluation Setup

System	#Versions	#Features	#Field Config.
Soletta	173	457	11,950
Toybox	62	334	1,039
Busybox	454	1,050	31,279
Fiasco	33	258	10,872
Uclibc	177	272	110,524
ERP	1	1,920	61
Agrib	1	2,238	5,749
Automotive08	unknown	221	233,309
Automotive09	unknown	279	247,780
Automotive10	unknown	279	45,387
Automotive11	unknown	218	27,808
Automotive12	unknown	147	129

Algorithms

- YASA
- Field Sampling
  - Random
  - Set-Cover
    $score(i) = count(i)$
  - Scoring
    $score(i) = \frac{1}{count(i)}$

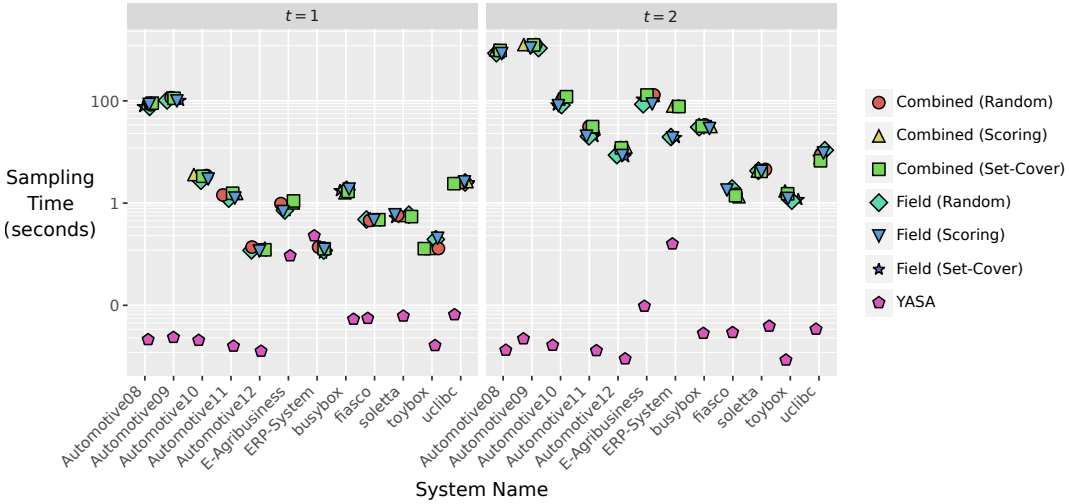
Evaluation Setup

System	#Versions	#Features	#Field Config.
Soletta	173	457	11,950
Toybox	62	334	1,039
Busybox	454	1,050	31,279
Fiasco	33	258	10,872
Uclibc	177	272	110,524
ERP	1	1,920	61
Agrib	1	2,238	5,749
Automotive08	unknown	221	233,309
Automotive09	unknown	279	247,780
Automotive10	unknown	279	45,387
Automotive11	unknown	218	27,808
Automotive12	unknown	147	129

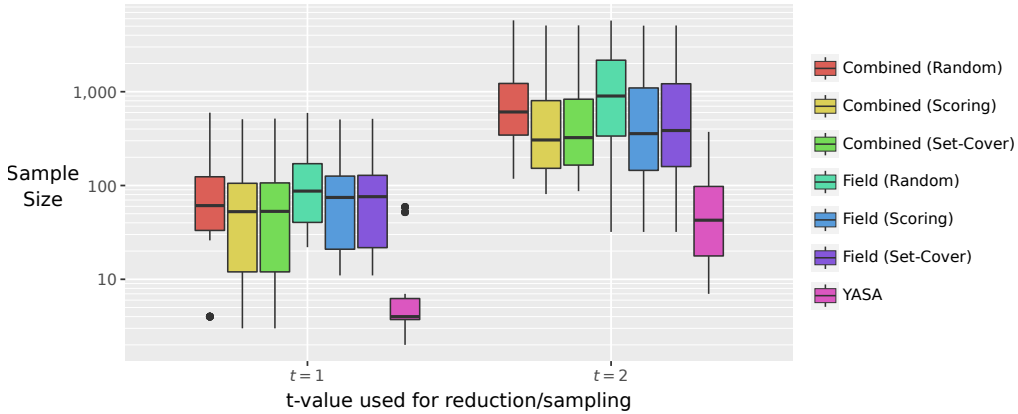
Algorithms

- YASA
- Field Sampling
  - Random
  - Set-Cover
    $score(i) = count(i)$
  - Scoring
    $score(i) = \frac{1}{count(i)}$
- Combined
  - Random
  - Set-Cover
  - Scoring

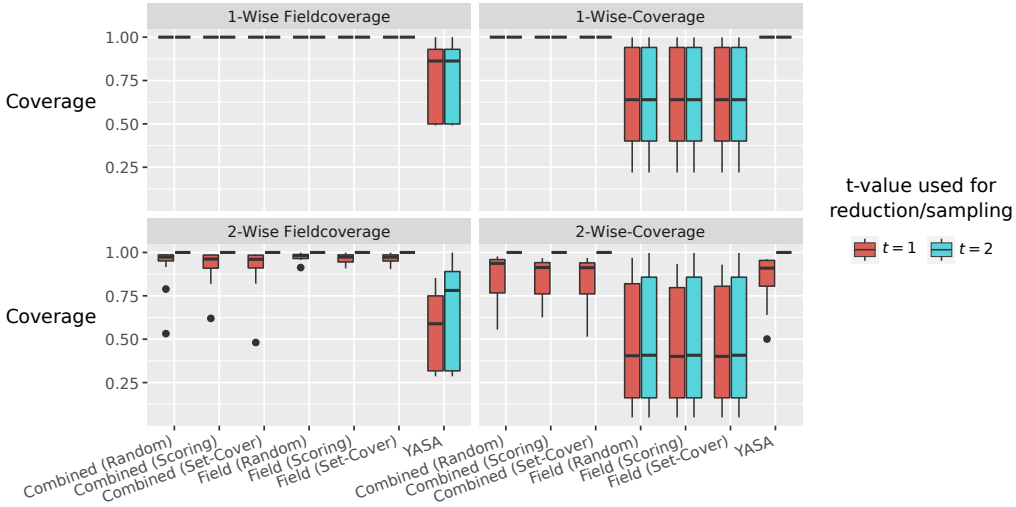
Evaluation Results Sampling Time



Evaluation Results Sample Size



Evaluation Results Coverage

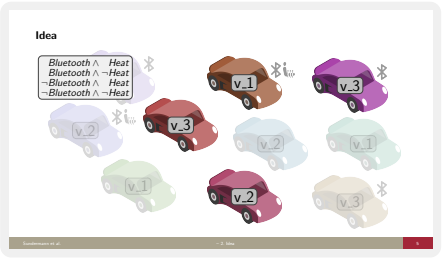


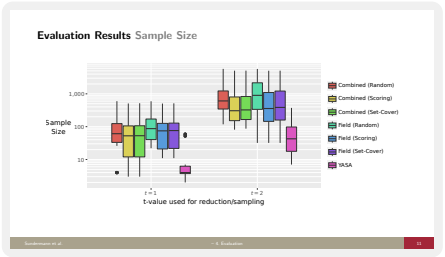
Future Work

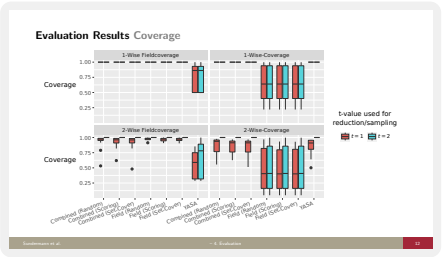
- Investigate alternative algorithms and scoring functions
- Take planned configurations into account

Conclusion


Motivation Testing Multiple Versions
Software Update Failed
Many Lexus navigation systems bricked by over-the-air software update
Test every t-wise sample of every feature model?
Unfeasible!
Software Update Failed Vehicle May Not Be Drivable
Sundermann et al.
1. Motivation
4


Idea
Bluetooth ∧ Heat
Bluetooth ∧ ¬Heat
¬Bluetooth ∧ Heat
¬Bluetooth ∧ ¬Heat
v_1
v_2
v_3
Sundermann et al.
2. Idea
5


Evaluation Results Sample Size
Sample Size
1,000
100
10
t = 1
t = 2
t-value used for reduction/sampling
Combined (Random)
Combined (Scoring)
Combined (Set-Cover)
Field (Random)
Field (Scoring)
Field (Set-Cover)
VASA
Sundermann et al.
6. Evaluation
11


Evaluation Results Coverage
1-Wise Fieldcoverage
1-Wise-Coverage
2-Wise Fieldcoverage
2-Wise-Coverage
Coverage
1.00
0.75
0.50
0.25
t-value used for reduction/sampling
t = 1
t = 2
Combined (Random)
Combined (Scoring)
Combined (Set-Cover)
Field (Random)
Field (Scoring)
Field (Set-Cover)
VASA
Sundermann et al.
6. Evaluation
12