Bug Prediction

DV Club Europe

November 23, 2021

Verifyter is now Cadence

cadence®

Bug Prediction

- Bug Prediction = Locating bugs without running simulation
- Example:
  - 100 commits have been checked in to the revision control system (e.g. GIT, Perforce)
  - Tests started to fail
  - Which commit(s) caused the test failures?
  - Bug Prediction will point you to the most likely commit(s)
- Benefit: Bugs found/fixed faster using less simulation time

Where do we use it?

- We use it in PinDown, our automatic debugger of regression failures
- PinDown predicts bad commits before simulation starts
- PinDown then validates bug reports by repairing **faulty code** to make the **test pass** again before bug report is issued

Bug No C23 (new bug)

Test: alu_ops_seed_14829533
Build: build_y80e
Error:
runarea/test/y80/sim/alu_ops_sim.log

```
FAILED: ALU operation failed
```

Validated: true
Committer: praveen (why me?)
Commit Message:
245646. Registered data input signal h7
Changes:
checkoutarea/test/y80/rtl/datapath.v [*verilog, hdl*]

```
assign carry_daa =

(daa_l1 && (daa_h1 || daa_h2)) || (daa_l2 && daa_h2 || daa_h3)) ||
(daa_l3 && (daa_h1 daa_h4)) || (daa_l & & daa_h5( ||
(daa_l5 && daa_h7daa_h7_reg
```

Training a Bug Prediction Model

Features from many sources are used to predict bugs

Revision History

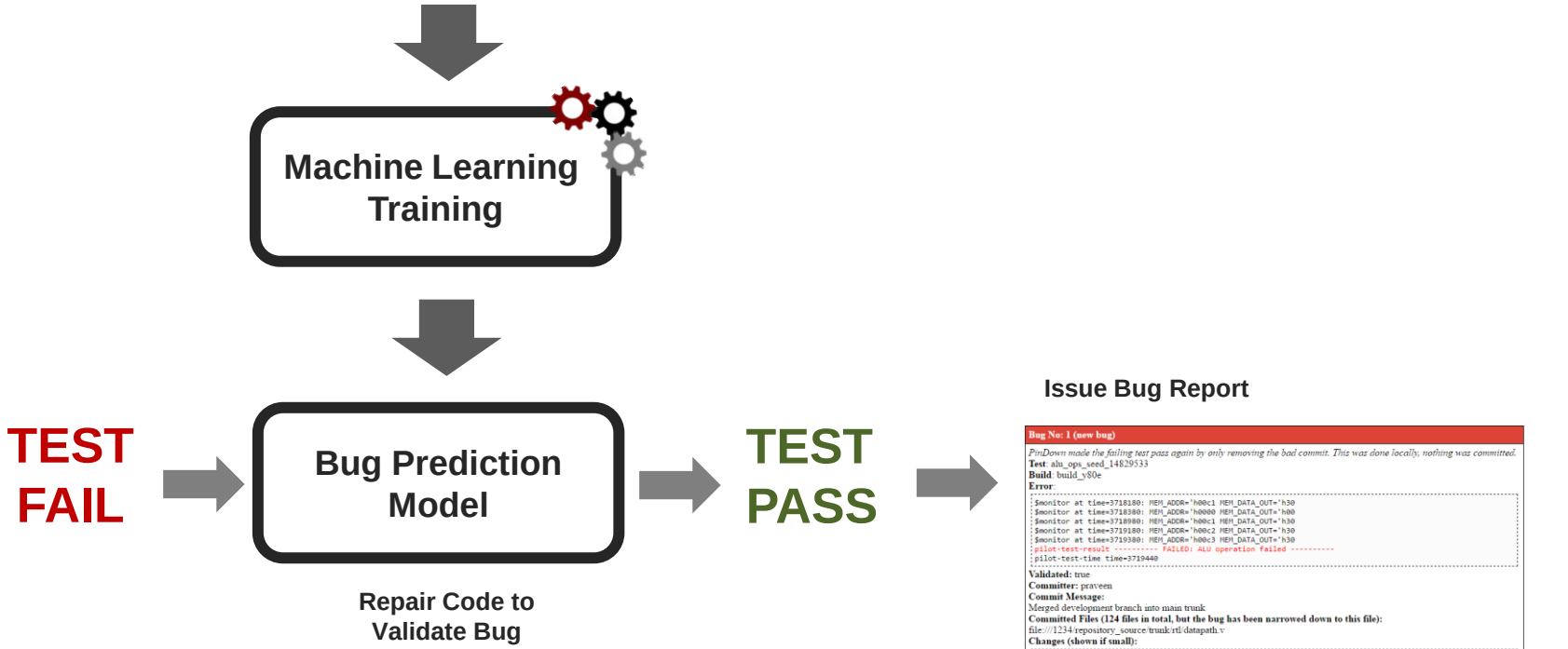
Code

Logs

Test Results

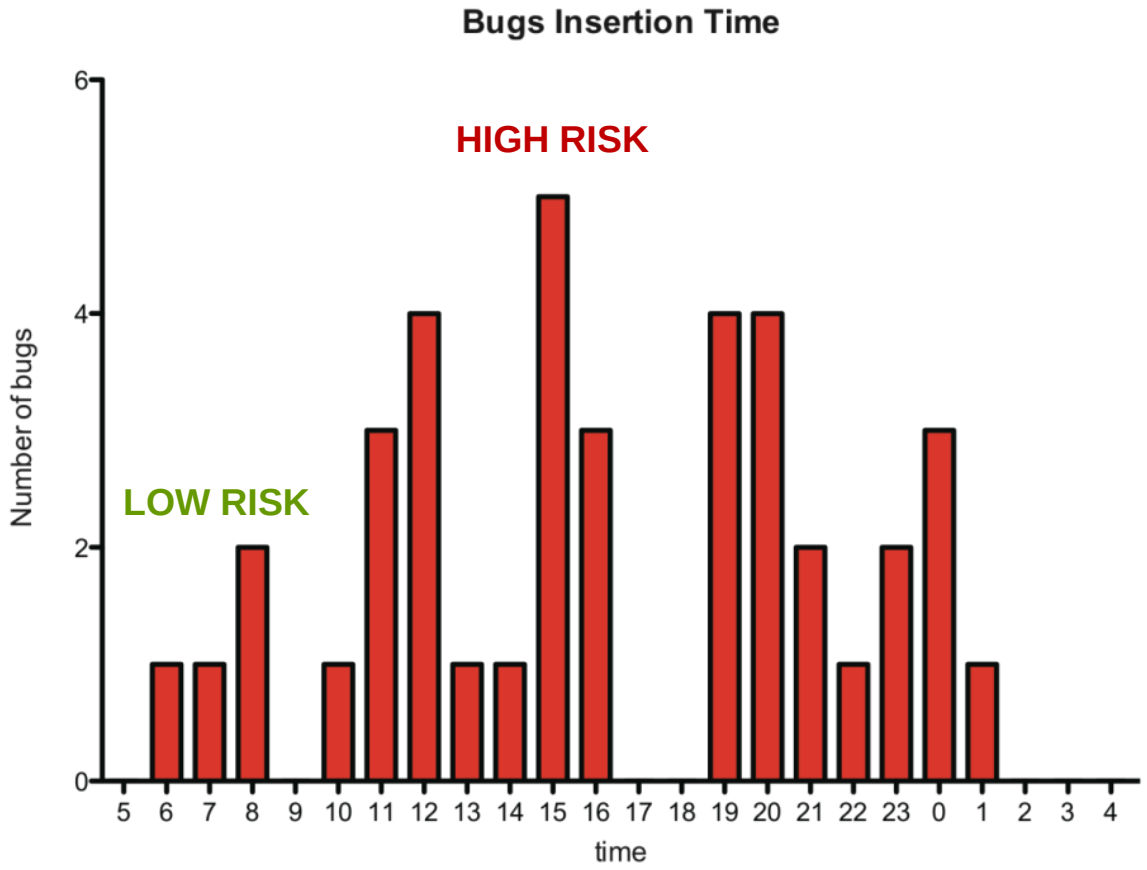
Bug Reports

Old PinDown bug reports are used to train the model on how to find bugs



cadence

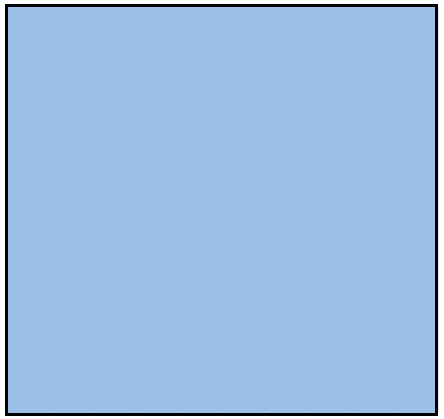
Feature: Commit Time



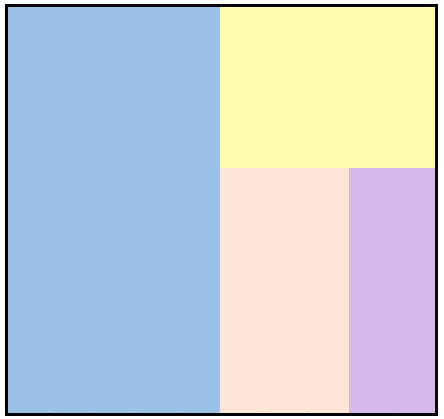
Median insertion time for bugs: 3 pm

Feature: File Ownership

LOW RISK

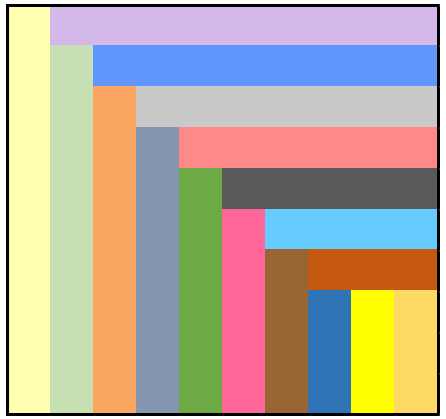


Single Committer



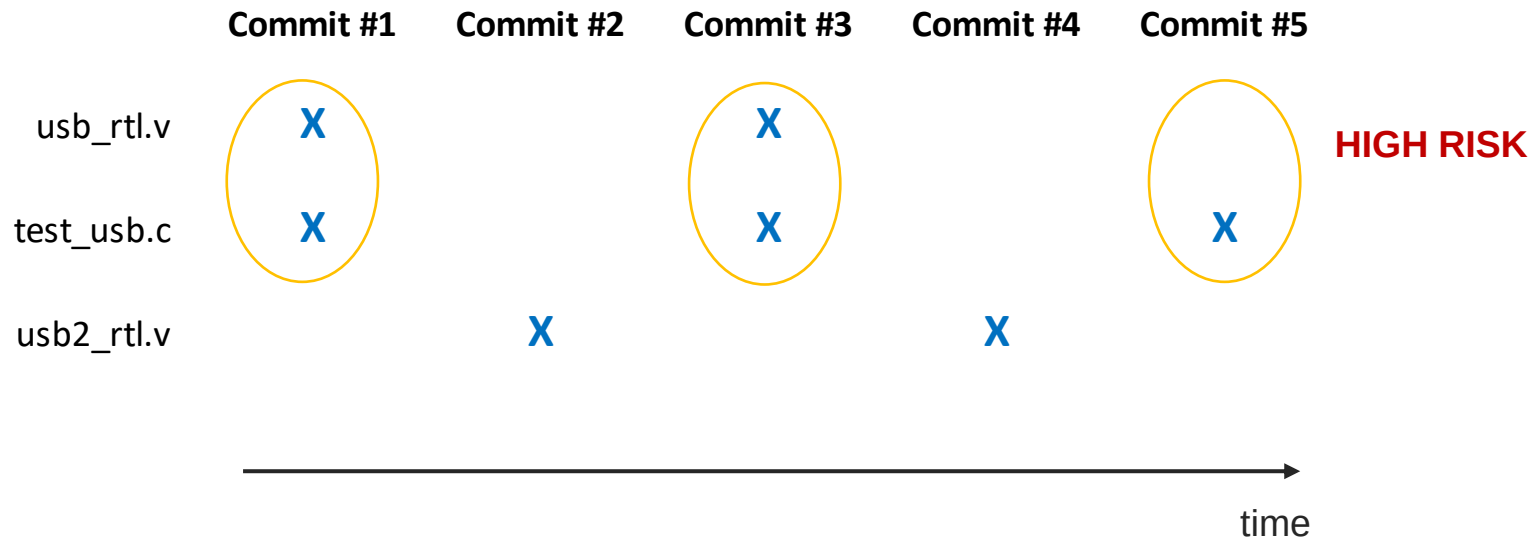
Many Committers

HIGH RISK



Everyone

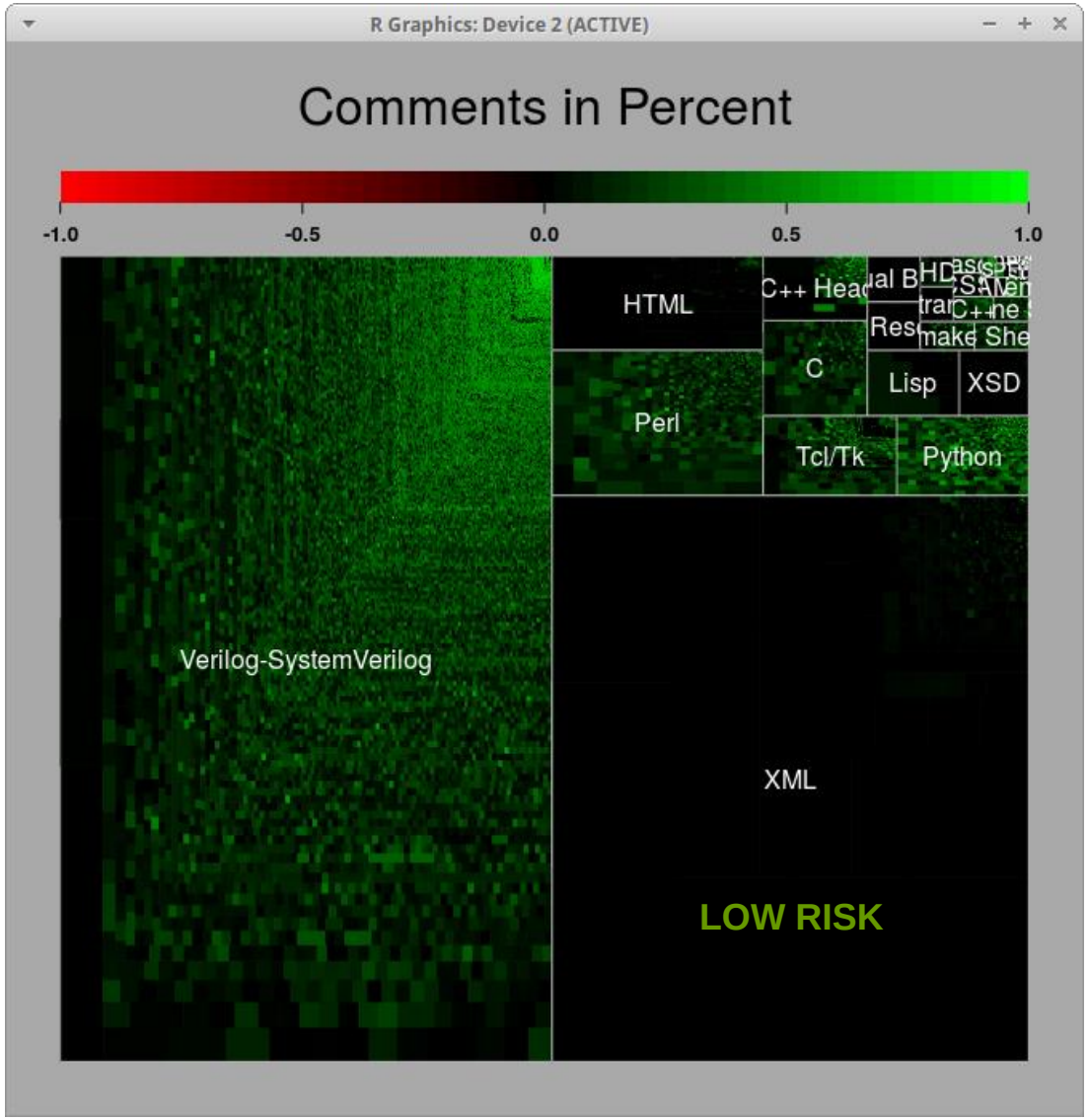
Feature: Change Coupling



Change Coupling = How often are files committed together?

This relationship can sometimes only be seen in the revision control system

Feature: Comment Ratio

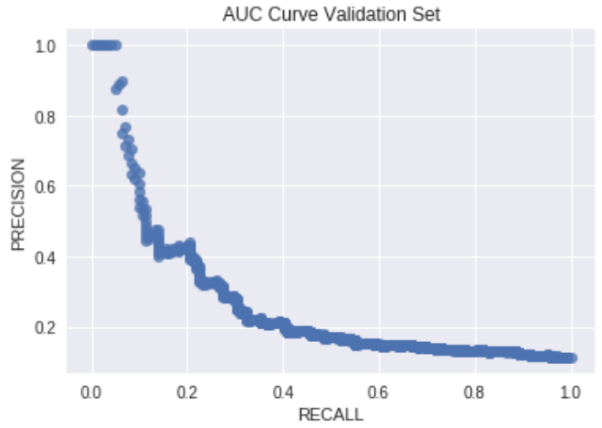


- Ratio Comments/Lines
- Detects Automation
  - Less Commented
  - Less Error-Prone
- Requires Language Awareness

Result: 41 % Precision

41% Precision

Metric	Value
Precision Mean	0.415
Precision Standard Deviation	0.0456
Recall Mean	0.186
Recall Standard Deviation	0.0255



=> 96% chance bug is in top 6 commits

Bug Predictions

Prediction	Revision	Date	Committer	Commit message
0.999403	...stbed0.git:7c643333a5	Feb 12 2011 5:19 AM CET	carlos	repo1@7 change c2t3: c2t6: c2t5:
0.996919	...stbed0.git:f36f5c8981	Feb 12 2011 5:16 AM CET	prashant	repo1@6 change c2t3: result (bit 0) to F c2t6: result (bit 0) to F c2t5: result (bit 0) to F
0.759582	...stbed0.git:26c2a86713	Feb 12 2011 5:25 AM CET	nageshwar	repo1@9 change config_2: result (bit 0) to P
0.759582	...stbed0.git:18e57d6808	Feb 12 2011 5:13 AM CET	sharon	repo1@5 change config_2: result (bit 0) to F
0.620521	...stbed1.git:816246c63f	Feb 12 2011 5:17 AM CET	praveen	repo2@6 change c2t3: c2t6: c2t5:
0.587064	...stbed2.git:54abe25cc2	Feb 12 2011 5:21 AM CET	prashant	repo3@7 change c2t2: result (bit 0) to F c2t1: result (bit 0) to F c2t4: result (bit 0) to F
0.043611	...stbed2.git:5800e5fee3	Feb 12 2011 5:18 AM CET	carlos	repo3@6 change c2t3: c2t6: c2t5:
0.000000	...stbed1.git:f0679c2296	Feb 12 2011 5:14 AM CET	hemal	repo2@5 change empty update
0.000000	...stbed1.git:c48c888f86	Feb 12 2011 5:05 AM CET	nageshwar	repo2@2 change empty update
0.000000	...stbed1.git:463089ee6b	Feb 12 2011 5:26 AM CET	prashant	repo2@9 change empty update
0.000000	...stbed1.git:a8f0c9ff4d	Feb 12 2011 5:11 AM CET	prashant	repo2@4 change empty update
0.000000	...stbed2.git:2255f0208e	Feb 12 2011 5:06 AM CET	prashant	repo3@2 change empty update
0.000000	...stbed2.git:5da6eb734b	Feb 12 2011 5:33 AM CET	sharon	repo3@11 change empty update
0.000000	...stbed2.git:16e699db5d	Feb 12 2011 5:30 AM CET	carlos	repo3@10 change empty update
0.000000	...stbed2.git:2286b531e8	Feb 12 2011 5:12 AM CET	sharon	repo3@4 change empty update
0.000000	...stbed2.git:a936a95dbc	Feb 12 2011 5:09 AM CET	carlos	repo3@3 change empty update
0.000000	...stbed1.git:1e01c3e2d0	Feb 12 2011 5:32 AM CET	sharon	repo2@11 change empty update
0.000000	...stbed1.git:f3816f9638	Feb 12 2011 5:29 AM CET	carlos	repo2@10 change empty update
0.000000	...stbed1.git:e58f2c0b20	Feb 12 2011 5:23 AM CET	sharon	repo2@8 change empty update
0.000000	...stbed1.git:9edc996e9e	Feb 12 2011 5:20 AM CET	carlos	repo2@7 change empty update
0.000000	...stbed1.git:7559f5647d	Feb 12 2011 5:08 AM CET	carlos	repo2@3 change empty update
0.000000	...stbed2.git:70cd69b001	Feb 12 2011 5:27 AM CET	praveen	repo3@9 change empty update
0.000000	...stbed2.git:6ff1c0be08	Feb 12 2011 5:24 AM CET	nageshwar	repo3@8 change empty update
0.000000	...stbed2.git:c9537c0d63	Feb 12 2011 5:15 AM CET	hemal	repo3@5 change empty update
0.000000	...stbed0.git:e46bf2274d	Feb 12 2011 5:07 AM CET	praveen	repo1@3 change empty update
0.000000	...stbed0.git:97aeededc1	Feb 12 2011 5:22 AM CET	sharon	repo1@8 change empty update
0.000000	...stbed0.git:28f0f5ed39	Feb 12 2011 5:04 AM CET	nageshwar	repo1@2 change empty update
0.000000	...stbed0.git:9d0fb28415	Feb 12 2011 5:31 AM CET	prashant	repo1@11 change empty update
0.000000	...stbed0.git:55bf132a5c	Feb 12 2011 5:28 AM CET	carlos	repo1@10 change empty update
0.000000	...stbed0.git:3a5b0dc8fc	Feb 12 2011 5:10 AM CET	carlos	repo1@4 change empty update

Generated by PinDown v4.2.5acf331+ Jul 9 2018 2:46 PM CEST

Source: Poster/paper “Predicting Bad Commits” from DVCon US, Feb 2019