Fault Localization: Incorporating normalized code length as parameter for coefficient based faulty ranking program

Presented by :이다 and 길랑

Advance Software Engineering

Fall 2022

SEOULTECH
SEOUL NATIONAL UNIVERSITY OF SCIENCE & TECHNOLOGY

Content

- Topic and key issue to be addressed
- Proposed approach and methods
- Experiment results and analysis
- Key findings

SEOULTECH
SEOUL NATIONAL UNIVERSITY OF SCIENCE & TECHNOLOGY

Topic & Key Issue

- Topic selected: Fault Localization
- Key issues to be addressed:
  1. Improving fault localization accuracy by investigating the importance (weights) of different testing coverage information.
  2. Improving suspiciousness scoring accuracy by leveraging syntactical feature.

SEOULTECH
SEOUL NATIONAL UNIVERSITY OF SCIENCE & TECHNOLOGY

Proposed approach

- Based on testing coverage information, suspiciousness factors can be divided into two categories.
  - More suspicious: N_{CF}, N_{US}
  - Less suspicious: N_{CS}, N_{UF}
- Faulty scoring formulation
  - More suspicious $(A) = N_{CF} + N_{US}$
  - Less suspicious $(B) = N_{CS} + N_{UF}$
  - $Score = \alpha A - \beta B$
- Adding syntactical feature L (statement length)
  - Refined score = $\alpha A - \beta B + \gamma L$

SEOULTECH
SEOUL NATIONAL UNIVERSITY OF SCIENCE & TECHNOLOGY

Methods

1. Data preprocessing to extract required features for each Java Program in the evaluation set.

Example: time1

```
5
org.joda.time.Partial#221 F
org.joda.time.field.UnsupportedDurationField#227 F
org.joda.time.field.UnsupportedDurationField#228 F
org.joda.time.field.UnsupportedDurationField#229 F
org.joda.time.Partial#217 T
org.joda.time.Chronology#63 3432 1 513 0 5
org.joda.time.DateTimeField#33 3526 1 419 0 5
org.joda.time.DateTimeFieldType#41 30 0 3915 1 7
org.joda.time.DateTimeFieldType#73 3528 1 417 0 8
```

SEOULTECH
SEOUL NATIONAL UNIVERSITY OF SCIENCE & TECHNOLOGY

Methods

2. Compute the normalized value of each feature
3. Calculate the suspiciousness score for each line of code
4. Compute the suspiciousness ranking and Exam Score

SEOULTECH
SEOUL NATIONAL UNIVERSITY OF SCIENCE & TECHNOLOGY

Experiment Results & Analysis

Performance without weight: $A - B + L$

		Proposed	Dstar	Jaccard
1	'chart19.txt'	0.193462308	0.000667111	0.241827885
2	'closure1.txt'	0.006813075	0.008407625	0.267739364
3	'closure10.txt'	0.015508622	0.013570044	0.258953168
4	'closure100.txt'	0.0183161	0.009748892	0.484194978
5	'closure101.txt'	0.67875383	0.352400409	0.018215867
6	'closure102.txt'	0.523537415	0.514829932	0.240816327
7	'closure103.txt'	0.984585742	0.444123314	0.19894027
8	'closure104.txt'	0.00310559	0.003992902	0.283496007
9	'closure105.txt'	0.009294466	0.020490072	0.396915927
10	'closure106.txt'	0.118925831	0.016943734	0.408407928
11	'closure107.txt'	0.013057546	0.013173099	0.48648024
12	'closure109.txt'	0.003916193	0.000783239	0.241237517
13	'closure11.txt'	0.114109763	0.053251223	0.310994385
14	'closure110.txt'	0.906006006	0.781181181	0.78048048
15	'closure112.txt'	0.001639486	0.006989386	0.416774528
16	'closure113.txt'	0.001757469	0.001318102	0.289762742
17	'closure115.txt'	0.015403235	0.010122126	0.386731214
18	'closure116.txt'	0.000287991	0.000671978	0.327541519
19	'closure117.txt'	0.741925091	0.755376561	0.754732179
⋮				
136	'time4.txt'	0.002604167	0.010416667	0.524553571
137	'time5.txt'	0.679693795	0.601128122	0.601128122
138	'time6.txt'	0.025623736	0.039447067	0.302090357
139	'time7.txt'	0.006245496	0.003843382	0.370405957
140	'time8.txt'	0.345534407	0.324304539	0.324304539
141	'time9.txt'	0.029304029	0.021245421	0.260805861

RESULT

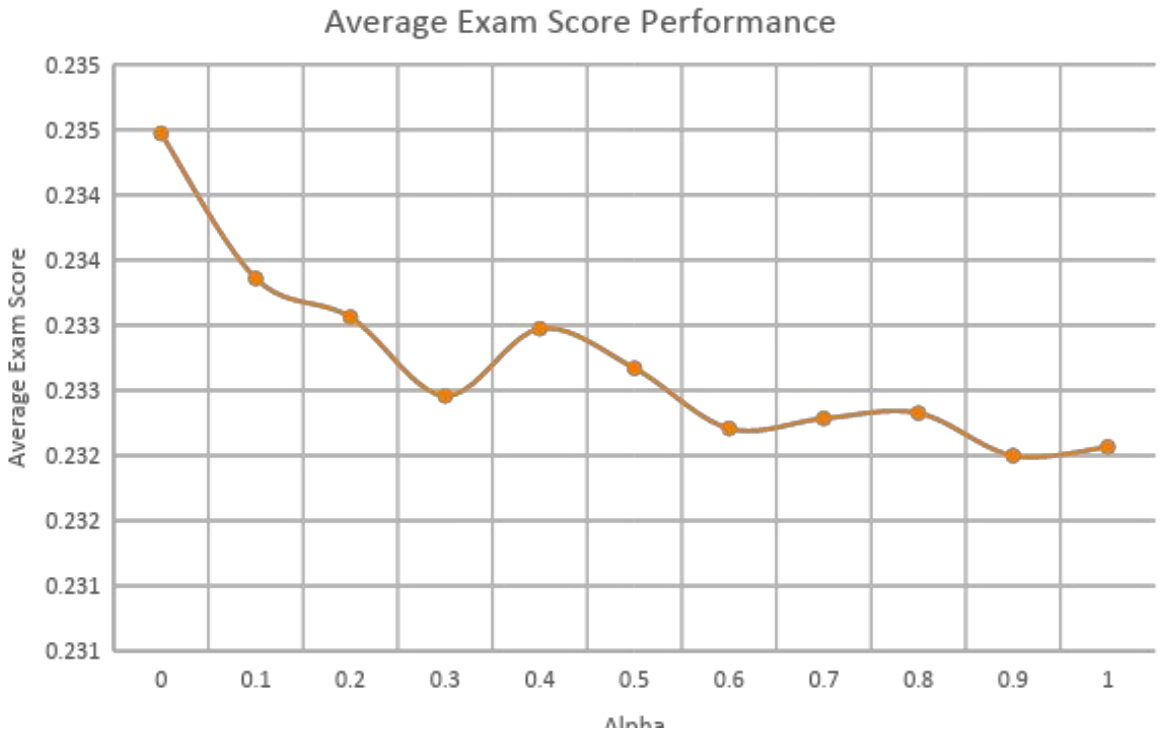
Comparison	Method	Number of Wins	Win Percentage
The best overall	Proposed Method	61	44%
	Dstar	66	47%
	Jaccard	10	7%
	Tie	3	2%
Proposed vs Dstar	Proposed Method	63	45%
	Dstar	75	53%
	Tie	3	2%
Proposed vs Jaccard	Proposed Method	111	79%
	Jaccard	29	21%
	Tie	1	1%

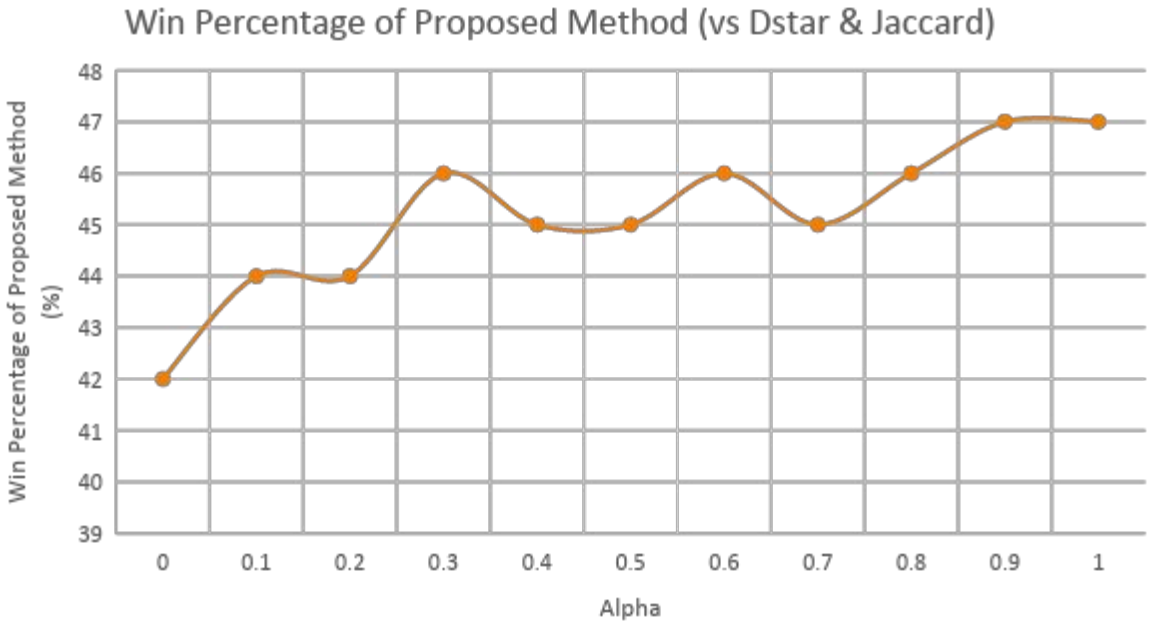
SEOULTECH
SEOUL NATIONAL UNIVERSITY OF SCIENCE & TECHNOLOGY

Experiment Results & Analysis

The significance of weight of each parameter

$$\alpha A - \beta B + \gamma L$$



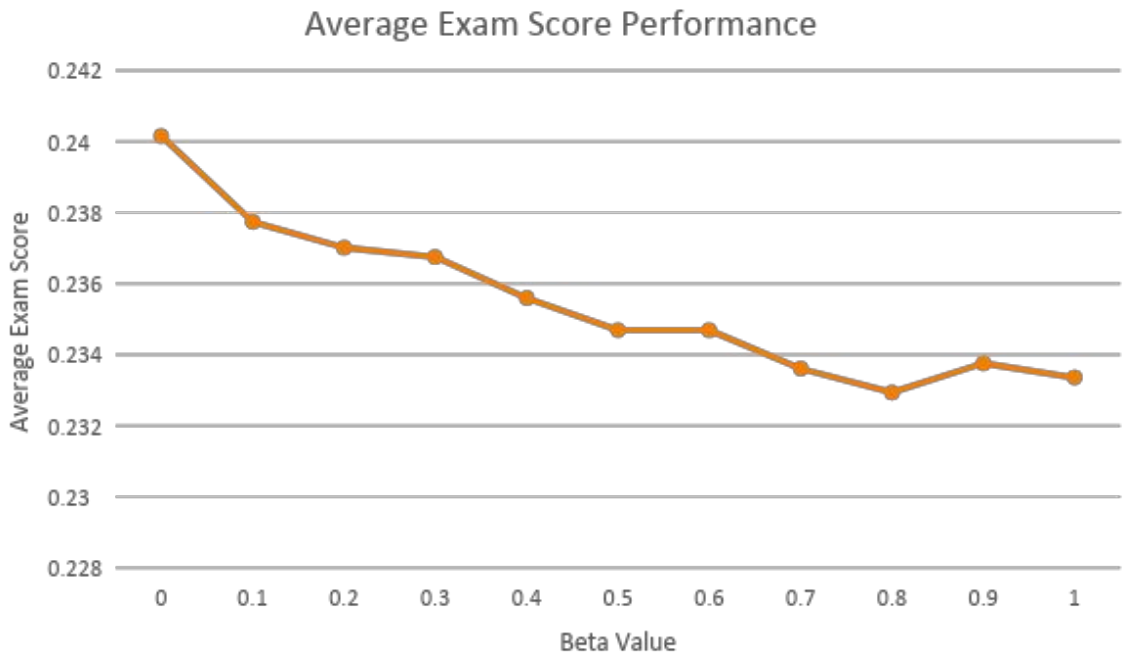


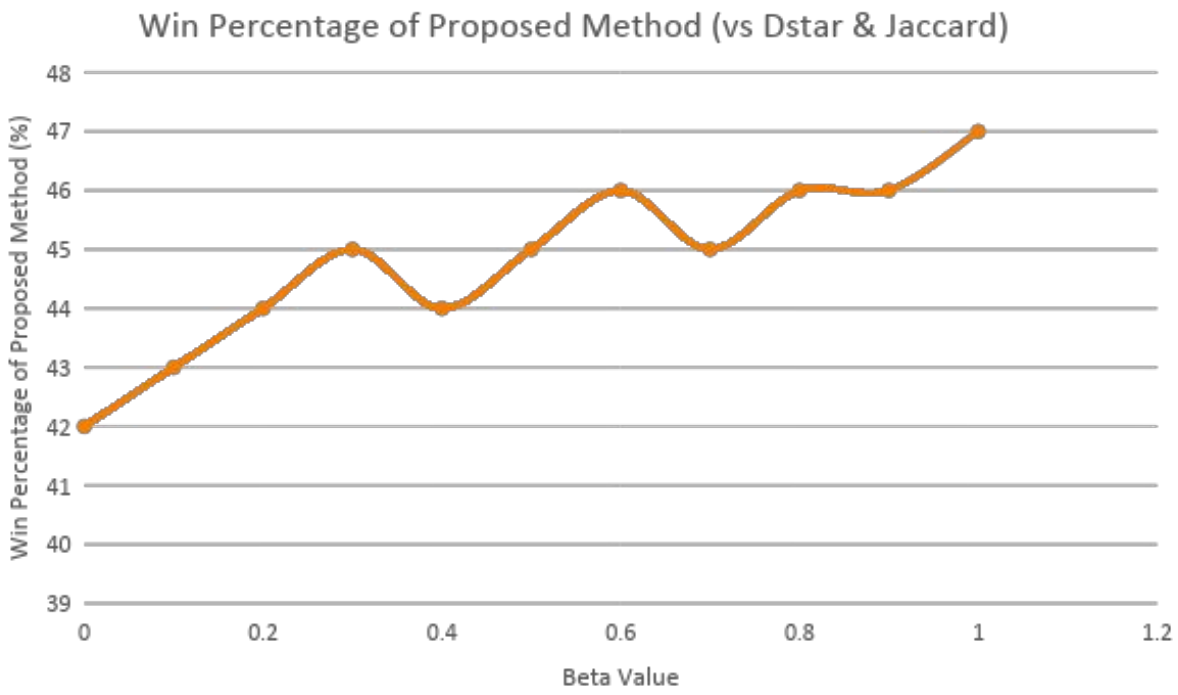
SEOULTECH
SEOUL NATIONAL UNIVERSITY OF SCIENCE & TECHNOLOGY

Experiment Results & Analysis

The significance of weight of each parameter

$$\alpha A - \beta B + \gamma L$$



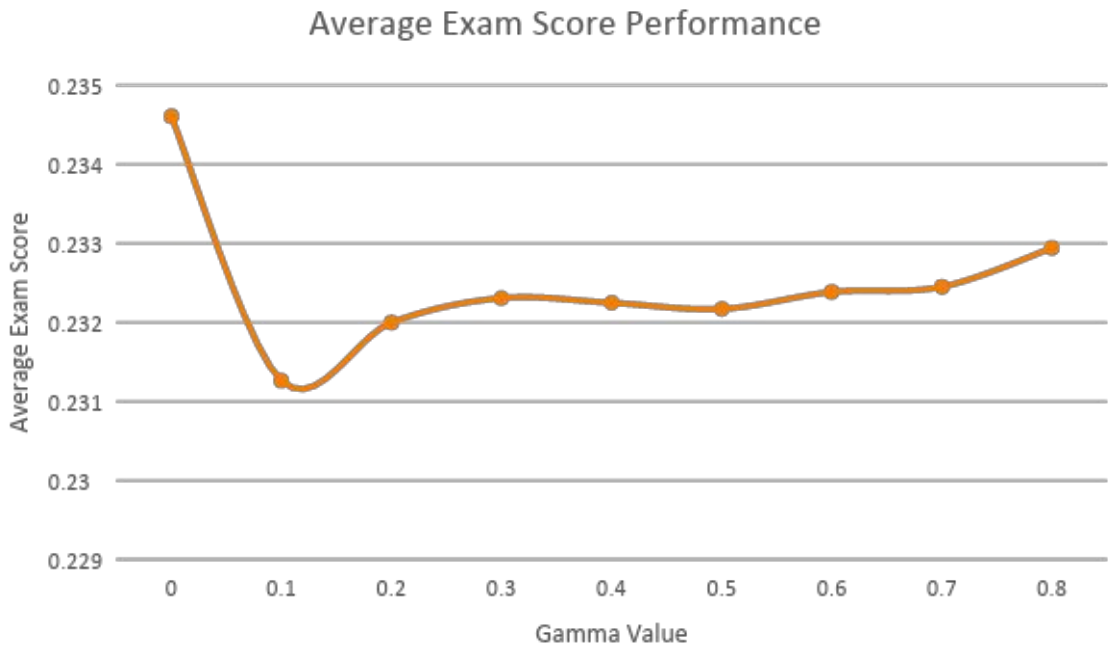


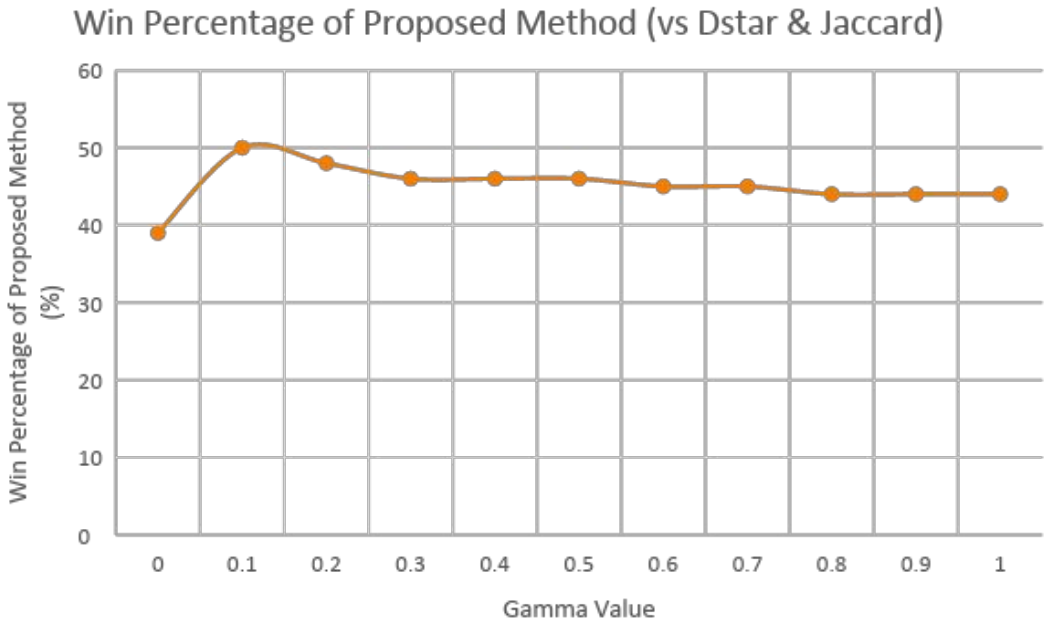
SEOULTECH
SEOUL NATIONAL UNIVERSITY OF SCIENCE & TECHNOLOGY

Experiment Results & Analysis

The significance of weight of each parameter

$$\alpha A - \beta B + \gamma L$$





SEOULTECH
SEOUL NATIONAL UNIVERSITY OF SCIENCE & TECHNOLOGY

Experiment Results & Analysis

The threshold of parameter setting

Based on those exploration, we get a rule of Parameter Setting

1. Higher α, higher performance.
2. Higher β, higher performance.
3. The optimum performance of γ at 0.1

- By iterating each parameter based on that rule, we define threshold setting at α = 0.9; $\beta = 0.9$; γ = 0.1.
- This is the minimum required setting to met Dstar performance.

SEOULTECH
SEOUL NATIONAL UNIVERSITY OF SCIENCE & TECHNOLOGY

Experiment Results & Analysis

Performance comparison with Jaccard and Dstar

$\alpha = 0.9; \beta = 0.9; \gamma = 0.1$ → (Increase only 0.1) → $\alpha = 1; \beta = 1; \gamma = 0.1$

RESULT

Comparison	Method	Number of Wins	Win Percentage
The best overall	Proposed Method	69	49%
	Dstar	59	42%
	Jaccard	11	8%
	Tie	1	1%
Proposed vs Dstar	Proposed Method	70	50%
	Dstar	70	50%
	Tie	1	1%
Proposed vs Jaccard	Proposed Method	111	79%
	Jaccard	29	21%
	Tie	1	1%

RESULT

Comparison	Method	Number of Wins	Win Percentage
The best overall	Proposed Method	72	51%
	Dstar	56	40%
	Jaccard	11	8%
	Tie	1	1%
Proposed vs Dstar	Proposed Method	73	52%
	Dstar	67	48%
	Tie	1	1%
Proposed vs Jaccard	Proposed Method	112	79%
	Jaccard	28	20%
	Tie	1	1%

SEOULTECH
SEOUL NATIONAL UNIVERSITY OF SCIENCE & TECHNOLOGY

Experiment Results & Analysis

Ablation Study

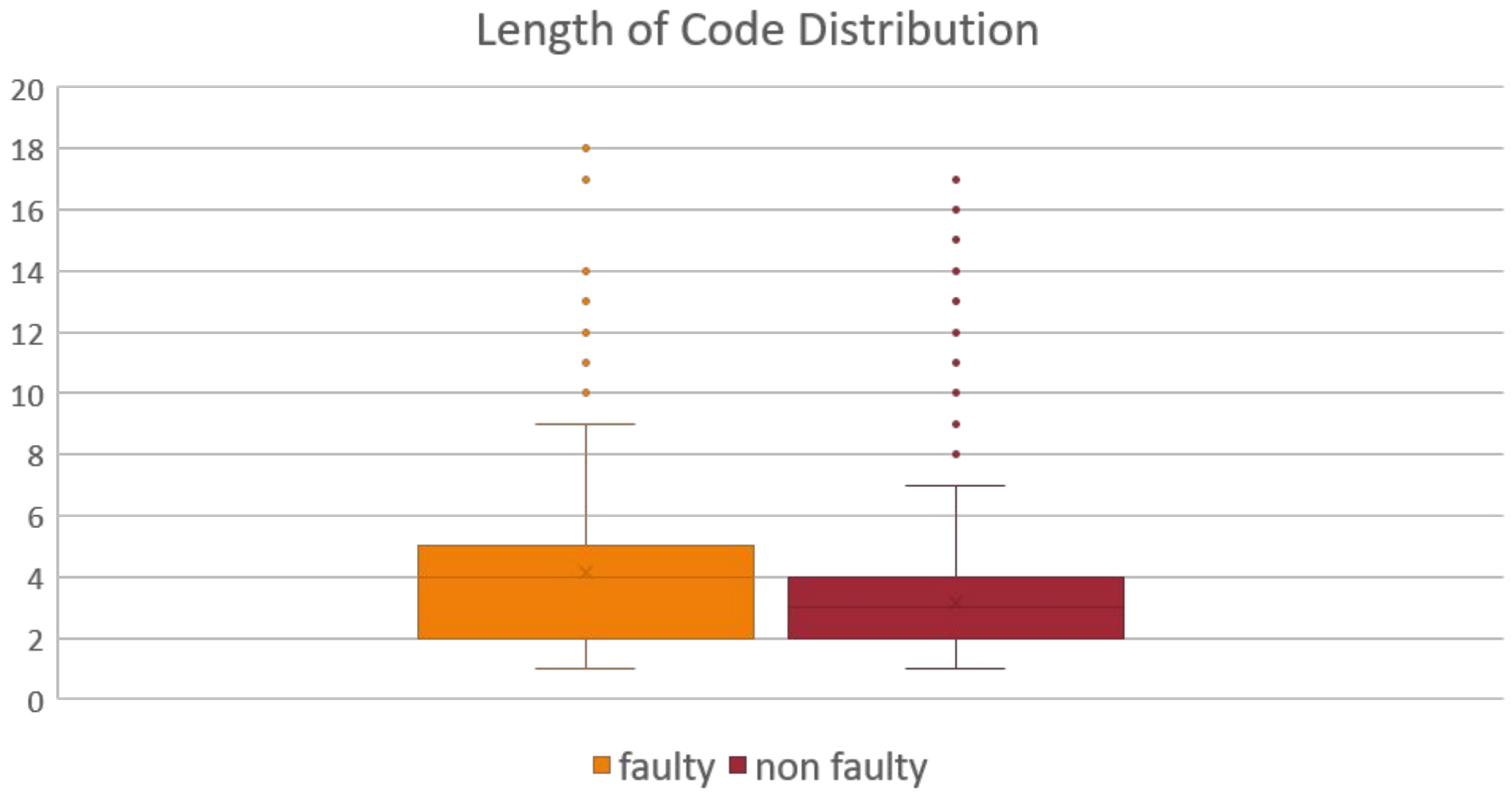
		Proposed	Dstar	Without Length
1	'chart19.txt'	0.152101401	0.000667111	0.133755837
2	'closure1.txt'	0.007392912	0.008407625	0.008335145
3	'closure10.txt'	0.011835527	0.013570044	0.013570044
4	'closure100.txt'	0.011225997	0.009748892	0.210044313
5	'closure101.txt'	0.632107593	0.352400409	0.652400409
6	'closure102.txt'	0.518639456	0.514829932	0.524829932
7	'closure103.txt'	0.975313102	0.444123314	0.973747592
8	'closure104.txt'	0.002218279	0.003992902	0.003992902
⋮				
134	'time27.txt'	0.111111111	0.116421569	0.116421569
135	'time3.txt'	0.762943787	0.542159763	0.600221893
136	'time4.txt'	0.005952381	0.010416667	0.010416667
137	'time5.txt'	0.669621273	0.601128122	0.667203868
138	'time6.txt'	0.009777478	0.039447067	0.010114633
139	'time7.txt'	0.004083594	0.003843382	0.003843382
140	'time8.txt'	0.321376281	0.324304539	0.485358712
141	'time9.txt'	0.022710623	0.021245421	0.021245421

RESULT

Comparison	Method	Number of Wins	Win Percentage
Proposed vs Without Length	Proposed Method	78	55%
	Without Length	55	39%
	Tie	8	6%
Proposed vs Dstar	Proposed Method	73	52%
	Dstar	67	48%
	Tie	1	1%
Without Length vs Dstar	Without Length	52	37%
	Dstar	75	53%
	Tie	14	10%

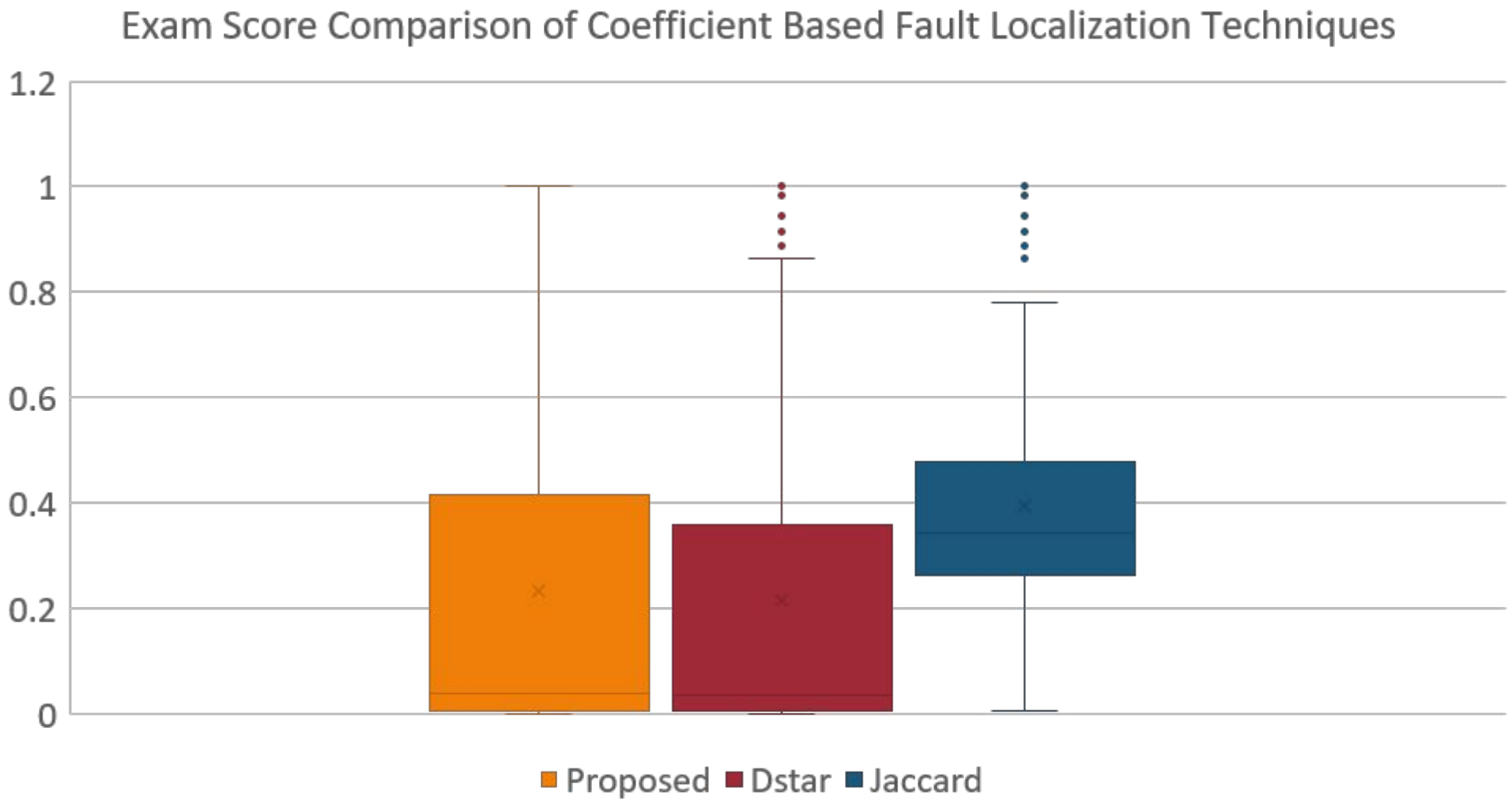
SEOULTECH
SEOUL NATIONAL UNIVERSITY OF SCIENCE & TECHNOLOGY

Experiment Results & Analysis



SEOULTECH
SEOUL NATIONAL UNIVERSITY OF SCIENCE & TECHNOLOGY

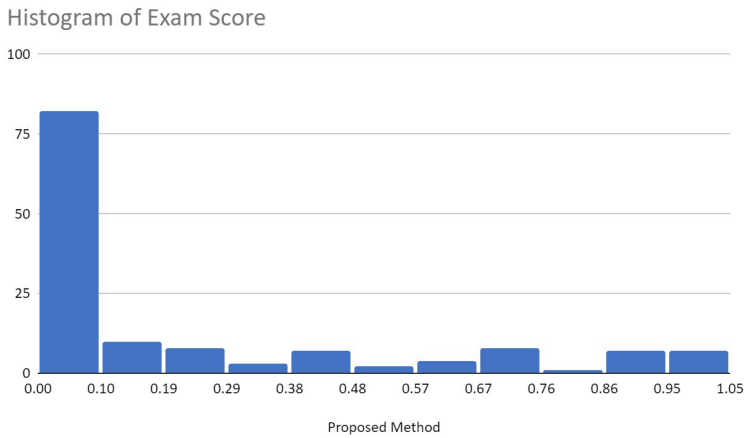
Experiment Results & Analysis

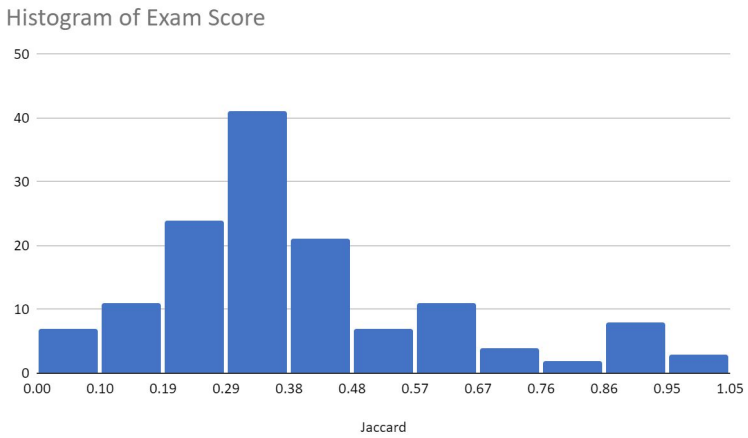


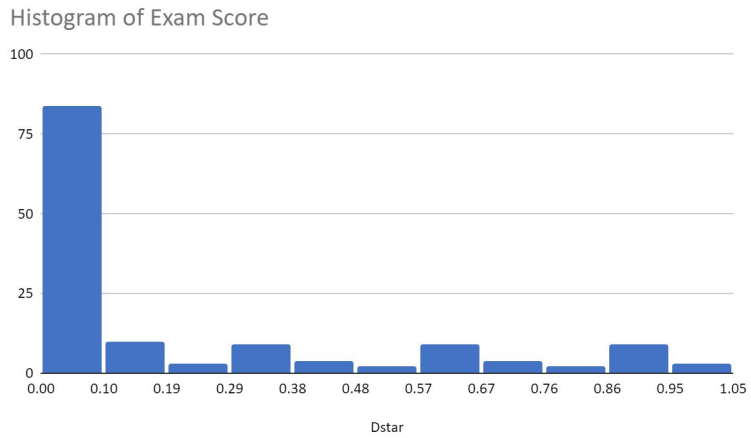
SEOULTECH
SEOUL NATIONAL UNIVERSITY OF SCIENCE & TECHNOLOGY

Experiment Results & Analysis

- From the histogram we can see that the Exam score variance is 1 with mean close to 0







SEOULTECH
SEOUL NATIONAL UNIVERSITY OF SCIENCE & TECHNOLOGY

Experiment Results & Analysis

- t-Test: Proposed vs Jaccard

	Proposed	*Jaccard*
Mean	0.2314	0.3952
Variance	0.1072	0.0509
Observations	141	141
Hypothesized Mean Difference	0	
df	140	
Alpha	0.05	
P(T<=t) one-tail	1.2576E-11	
P(T<=t) two-tail	2.5152E-11	

SEOULTECH
SEOUL NATIONAL UNIVERSITY OF SCIENCE & TECHNOLOGY

Experiment Results & Analysis

- t-Test: Proposed vs DStar

	Proposed	*DStar*
Mean	0.2314	0.2147
Variance	0.1072	0.0934
Observations	141	141
Hypothesized Mean Difference	0	
df	140	
Alpha	0.05	
P(T<=t) one-tail	0.124515	
P(T<=t) two-tail	0.24903	

SEOULTECH
SEOUL NATIONAL UNIVERSITY OF SCIENCE & TECHNOLOGY

Key Findings

- The syntactical feature length helps in improving the fault localization performance.
- The proposed method outperforms Jaccard in all parameters setup.
- The proposed method reaches Dstar performance with threshold parameter setting (α = 0.9; β = 0.9; γ = 0.1)

SEOULTECH
SEOUL NATIONAL UNIVERSITY OF SCIENCE & TECHNOLOGY

Public Repository

https://github.com/archive05/fault-localization

SEOULTECH
SEOUL NATIONAL UNIVERSITY OF SCIENCE & TECHNOLOGY