HINDBR: Heterogeneous Information Network Based Duplicate Bug Report Prediction

Guanping Xiao[1], Xiaoting Du[2], Yulei Sui[3], Tao Yue[1]

[1]Nanjing University of Aeronautics and Astronautics, China

[2]Beihang University, China

[3]University of Technology Sydney, Australia

13/10/2020

Microsoft Teams

Motivation

Duplicate bug reports often exist in bug tracking systems (BTSs), leading to unnecessary maintenance effort such as repeatedly discussing the same bug.

Existing efforts on automatically detecting duplicate bug reports heavily rely on the text similarity calculated with information retrieval (IR) techniques (e.g., TF-IDF).

Motivation

Modern BTSs introduce just-in-time (JIT) retrieval feature in their recent versions, e.g., Bugzilla 4.0.

The built-in JIT feature can suggest possible duplicates when a reporter is filling a bug (i.e., typing in the summary field).

Motivation

Before reporting a bug, please read the bug writing guidelines, please look at the list of most frequently reported bugs, and please search for the bug.

Show Advanced Fields (* = Required Field)

* **Product:** Drivers **Reporter:** xiaogp13@163.com

* **Component:** Bluetooth, Console/Framebuffers, EDAC, Flash/Memory Technology Devices, Hardware Monitoring, HotPlug, I2C

Component Description
Select a component to read its description.

Severity: normal

Hardware: All

* **Kernel Version:**

* **Summary:** bluetooth

JIT feature

Possible Duplicates:

Bug ID	Summary	Status	
20052	Bluetooth mouse only works briefly (Sony Vaio Z, Broadcom Bluetooth)	RESOLVED CODE_FIX	Add Me to the CC List
43199	rfkill block bluetooth when ppp rfcomm connection is active hangs bluetooth, and possibly networking subsystem	RESOLVED CODE_FIX	Add Me to the CC List
43218	Bluetooth driver is drop when stop/restart bluetooth at active connections	RESOLVED CODE_FIX	Add Me to the CC List
51221	[BISECTED]Bluetooth connections not working with 0a5c:201e Broadcom Corp. IBM Integrated Bluetooth IV	NEW	Add Me to the CC List
99371	ath3k Bluetooth adapter not loading: [0cf3:3004] Atheros Communications, Inc. AR3012 Bluetooth 4.0	NEW	Add Me to the CC List
103671	systemd throws Bluetooth: hci0 hardware error 0x37, but bluetooth works fine	NEW	Add Me to the CC List
197121	Bluetooth regression "no bluetooth adapters found" Baytrail T100CHI (SDIO) BCM4324B3	RESOLVED DOCUMENTED	Add Me to the CC List

Motivation

With the advent of the just-in-time (JIT) retrieval feature in modern BTSs, textual-based approaches become ineffective in detecting after-JIT duplicate bug reports[1].

The built-in JIT feature can suggest possible duplicates when a reporter is filling a bug (i.e., typing in the summary field), thereby reducing chances for submitting duplicate reports in the first place.

After JIT filtering, a substantial proportion of duplicate reports still exists in BTSs.

1. Rakha, M. S., Bezemer, C. P., & Hassan, A. E. (2018). Revisiting the performance of automated approaches for the retrieval of duplicate reports in issue tracking systems that perform just-in-time duplicate retrieval. Empirical Software Engineering, 23(5), 2597-2621.

Motivation

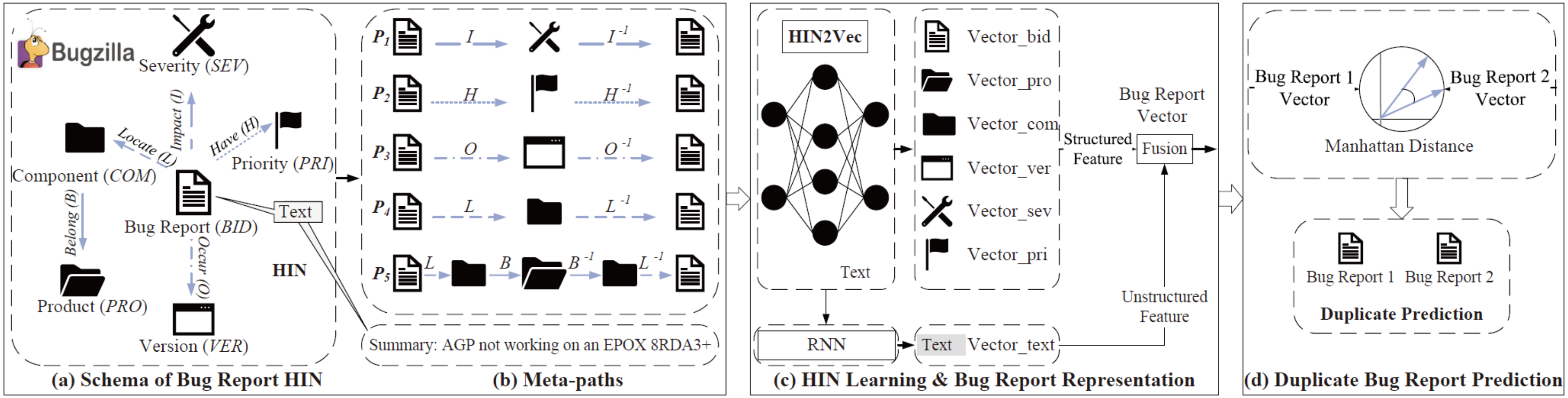

Fig. 1. Overview of HINDBR.

A Motivating Example

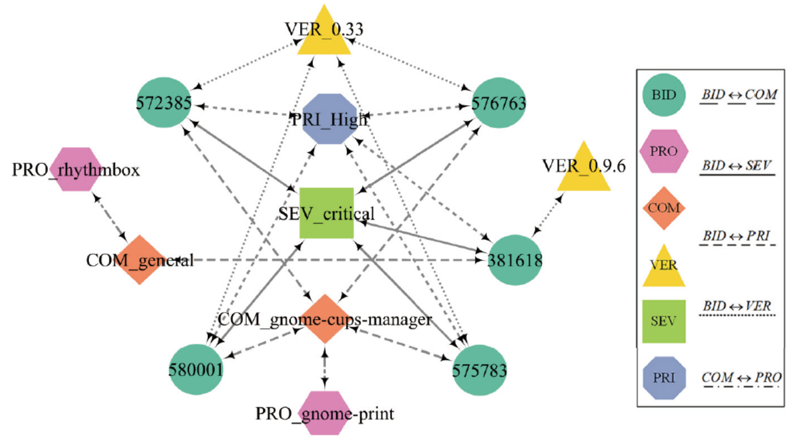

Fig. 2. An excerpt of a bug report HIN for the GNOME project.

TABLE I
SIMILARITIES OF DUPLICATE AND NON-DUPLICATE PAIRS WITH HIN VECTORS

Duplicate	**Similarity**	**Non-Duplicate**	**Similarity**
(576763, 575783)	1	(381618, 576763)	2.14E-224
(580001, 575783)	0.98		
(576763, 572385)	0.96	(381618, 580001)	4.51E-225
(572385, 575783)	0.76	(381618, 572385)	1.28E-225
(580001, 572385)	0.51		
(580001, 576763)	0.41	(381618, 575783)	0

Background

Duplicate Bug Report Prediction


```xml
<bugzilla maintainer="helpdesk@kernel.org" urlbase="https://bugzilla.kernel.org/"
version="5.1.1">
<bug>
<bug_id>200389</bug_id>
<creation_ts>2018-07-02 01:59:58 +0000</creation_ts>
<short_desc>iwlmvm: 7265: stops working after kernel warning / trace</short_desc>
<product>Drivers</product>
<component>network-wireless</component>
<version>2.5</version>
<bug_status>CLOSED</bug_status>
<resolution>DUPLICATE</resolution>
<dup_id>199967</dup_id>
<priority>P1</priority>
<bug_severity>normal</bug_severity>
...
</bug>
</bugzilla>
```


Fig. 3. Linux bug report ID-200389 (XML format).

Background

Duplicate Bug Report Prediction

TABLE II
AN EXAMPLE OF BUG GROUP IN GNOME PROJECT

Type	**Bug ID**	**Summary**
Master	572385	crash in Printing: Just clicked the gnome-c...
Duplicates	575783	crash in Printing:
	576763	crash in Printing: launching gnome-cups-man...
	580001	crash in Printing: Checking to see why I co...

Background

HIN2Vec: a Network Representation Learning for HIN

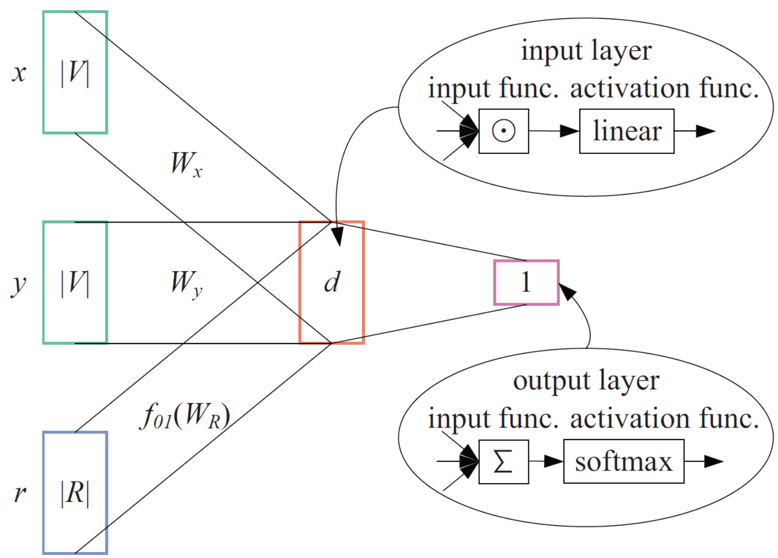

Fig. 4. The HIN2Vec neural network model.

Our HINDBR Approach

Constructing HIN for Bug Reports

- Six Nodes: BID, COM, PRO, VER, PRI, SEV
- Five Relations: R1: Bug-Component
  R2: Component-Product
  R3: Bug-Version
  R4: Bug-Priority
  R5: Bug-Severity

Our HINDBR Approach

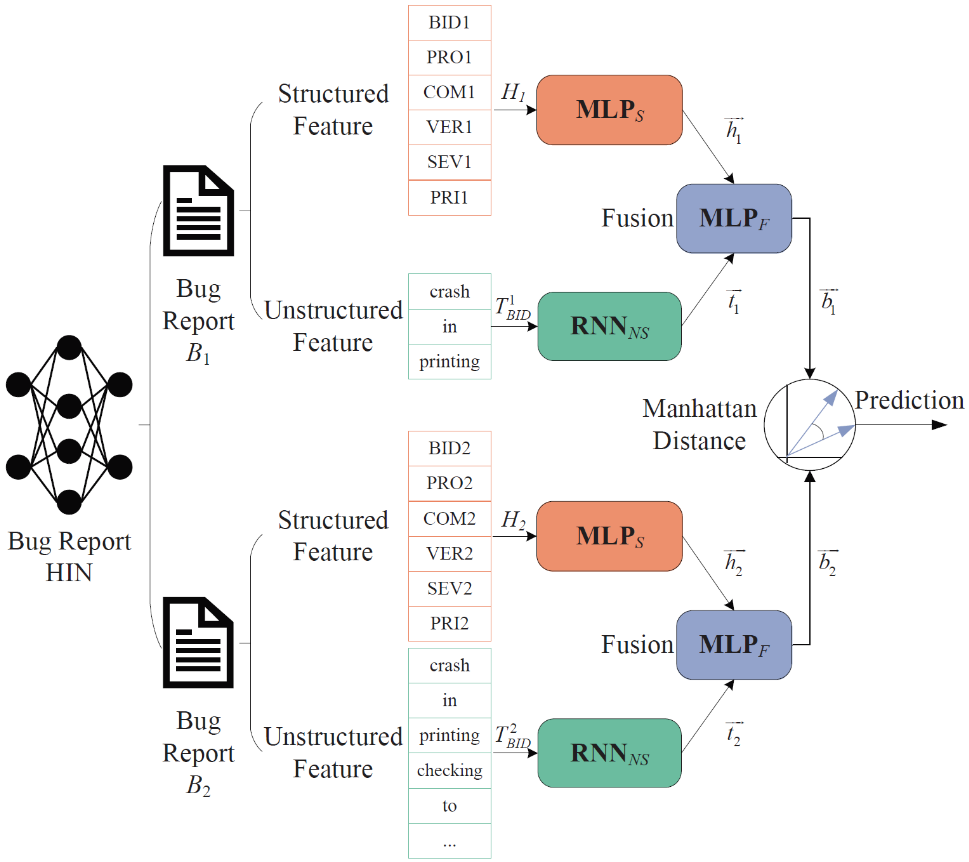

Fig. 5. Detailed structure of HINDBR.

Our HINDBR Approach

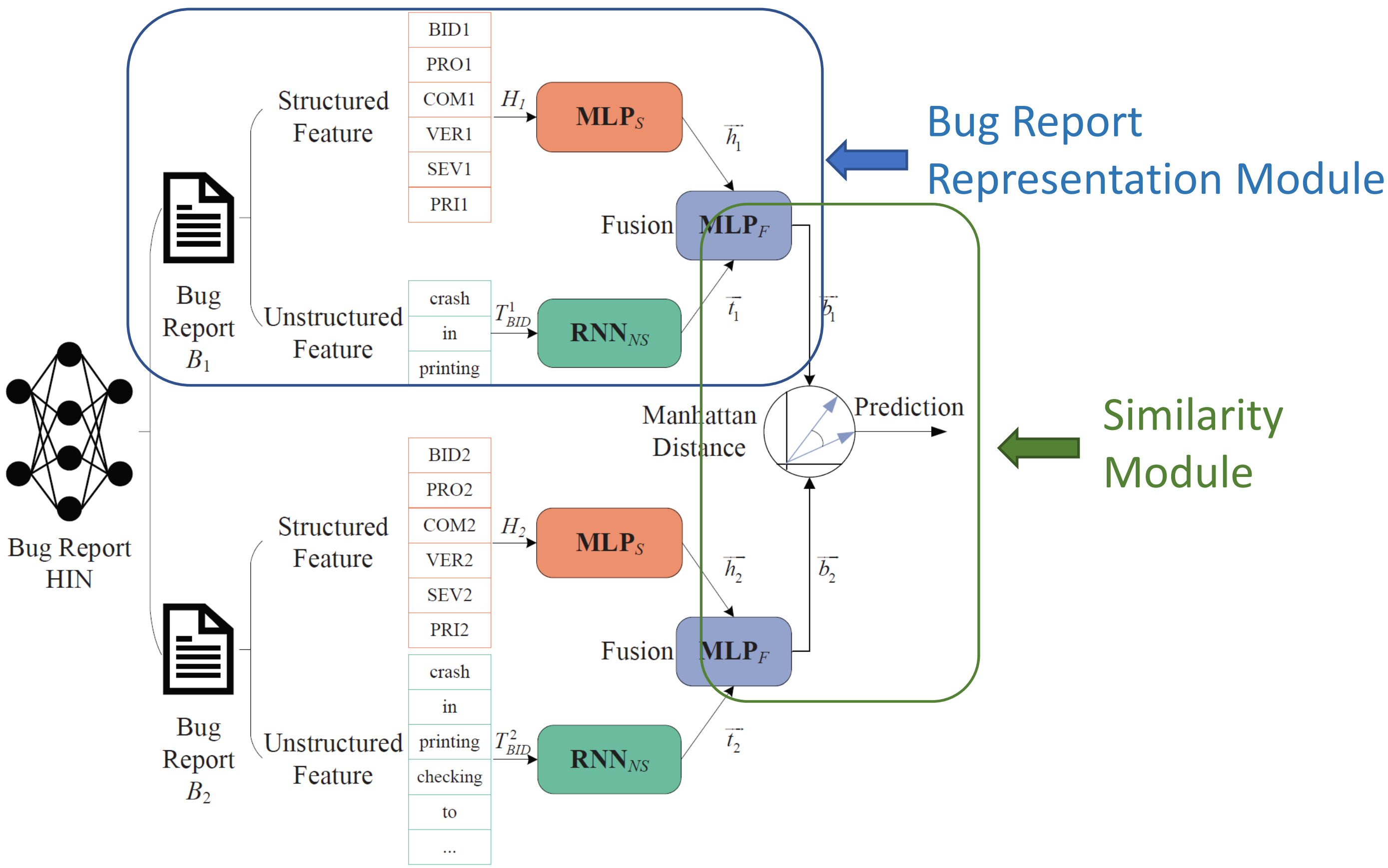

Fig. 5. Detailed structure of HINDBR.

Our HINDBR Approach

Bug Report Representation Module

- Structured Feature: MLP_S

$$h = tanh(W^H H),$$

- Unstructured Feature: RNN_{NS}

$$t_i = tanh(W^T [x_i, t_{i-1}]), \forall i = 1, 2, \ldots, N_T,$$

- Feature Fusion: MLP_F

$$b = tanh(W^B [h, t]),$$

Our HINDBR Approach

Similarity Module

- Manhattan Distance

$$S(b_1, b_2) = exp(-||b_1 - b_2||_1),$$

Model Training: Training Instance $< B_1, B_2 >$. Label: 1 for duplicate

- Loss Function: binary cross entropy loss

$$\mathcal{L}(\theta) = -(ylog(\hat{y}) + (1-y)log(1-\hat{y})),$$

- Similarity Threshold: 0.5

$$\hat{y} = \begin{cases} 1, & S(B_1, B_2) \geq 0.5 \\ 0, & S(B_1, B_2) < 0.5 \end{cases},$$

Data Collection & Aggregation

Data Collection

TABLE III
COLLECTED BUG REPORTS

Project Type	Project	Time Frame	JIT Year	# of Reports
Development Tool	Eclipse	10/10/01 - 09/30/18	2011 [29]	528,862
	GCC	08/03/99 - 09/30/18	2011 [30]	81,463
	LLVM	10/07/03 - 09/30/18	Unknown	38,107
Desktop Environment	Freedesktop	01/09/03 - 09/30/18	2011 [31]	106,065
	GNOME	02/05/99 - 09/30/18	Unknown	673,301
	KDE	01/21/99 - 09/30/18	2012 [32]	388,711
Office Suite	LibreOffice	08/03/10 - 09/30/18	Unknown	62,029
	OpenOffice	10/16/00 - 09/30/18	2012 [33]	127,797
Operating System	Linux kernel	11/06/02 - 09/30/18	2012 [34]	32,340
Total				2,038,675

Data Collection & Aggregation

Feature Extraction

- HIN Construction
- Text Extraction

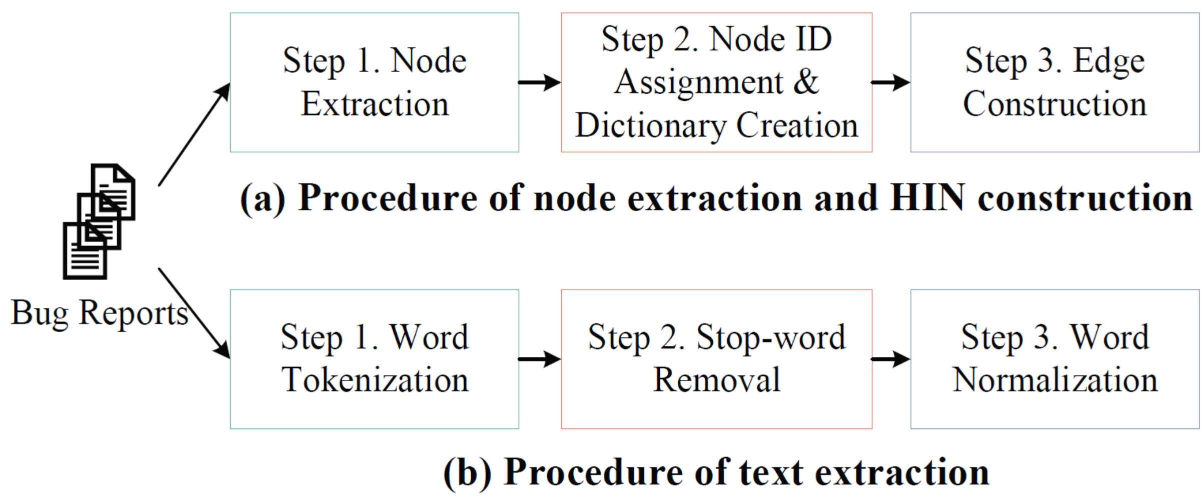

Fig. 7. Procedure of constructing HIN and extracting text.

Data Collection & Aggregation

Bug Pairs Generation: Model Training and Testing

TABLE IV
NUMBER OF BUG PAIRS FOR MODEL TRAINING AND TESTING

Project	**Duplicate Pair**	**Non-Duplicate Pair**	**Pair**
Eclipse	54,742	218,968	273,710
Freedesktop	11,316	45,264	56,580
GCC	7,819	31,276	39,095
GNOME	69,381	277,524	346,905
KDE	41,094	164,376	205,470
LibreOffice	6,771	27,084	33,855
Linux kernel	2,998	11,992	14,990
LLVM	3,093	12,372	15,465
OpenOffice	12,821	51,284	64,105
Total	210,035	840,140	1,050,175

Data Collection & Aggregation

Bug Pairs Generation: Before-JIT & After-JIT Evaluation

TABLE V
NUMBER OF BUG PAIRS FOR BEFORE-JIT AND AFTER-JIT EVALUATION

Project	Before-JIT		After-JIT	
	Duplicate	Non-Duplicate	Duplicate	Non-Duplicate
Eclipse	5,474	21,896	5,474	21,896
Freedesktop	1,131	4,524	1,131	4,524
GCC	781	3,124	781	3,124
KDE	4,109	16,436	4,109	16,436
Linux kernel	299	1,196	299	1,196
OpenOffice	1,282	5,128	1,282	5,128

Implementation Details

Settings of Pre-trained Embeddings

- Word2Vec: d_1 – 100
- HIN2Vec: d_2 – 128

Settings of Neural Networks in HINDBR

- Text
- HIN1 (no Text)
- HIN2 (with Text)

Implementation Details

Settings of Model Training: Keras, Dell Precision Tower, RTX2080Ti

- Training Parameters: epochs – 100; batch size – 128;
- Stratified Cross-Validation Evaluation: 5-fold cross-validation
- Dealing with Imbalanced Data: SMOTE + TL

Evaluation Metrics

- Accuracy
- Precision
- Recall
- F1 Score

Model Evaluation

Comparison Method

- DLDbr: Text feature (Long: CNN + Short: LSTM)
  Structure feature (Numerical Vectors)

Research Questions (RQs)

- RQ1: HINDBR Effectiveness
- RQ2: Impacts of Feature Settings
- RQ3: Impacts of Before-JIT and After-JIT Duplicates

Model Evaluation

RQ1: HINDBR Effectiveness

TABLE VI
PREDICTION RESULTS OF HINDBR COMPARED WITH BASELINE APPROACH DLDBR

Project	Accuracy			Precision			Recall			F1 Score		
	DLDBR	HINDBR	Impro.	DLDBR	HINDBR	Impro.	DLDBR	HINDBR	Impro.	DLDBR	HINDBR	Impro.
Eclipse	0.8910	**0.9489**	**6.51%**	0.8196	**0.9005**	**9.87%**	0.7930	**0.8374**	**5.60%**	0.8037	**0.8671**	**7.89%**
Freedesktop	0.9161	**0.9621**	**5.01%**	0.8503	**0.9184**	**8.01%**	0.8519	**0.8891**	**4.36%**	0.8504	**0.9035**	**6.24%**
GCC	0.9061	**0.9587**	**5.81%**	0.8523	**0.9205**	**8.01%**	0.8306	**0.8721**	**5.01%**	0.8392	**0.8957**	**6.73%**
GNOME	0.9843	**0.9883**	**0.42%**	0.9620	**0.9709**	**0.93%**	**0.9769**	0.9707	-0.63%	0.9693	**0.9708**	**0.15%**
KDE	0.9639	**0.9834**	**2.02%**	0.9363	**0.9651**	**3.08%**	0.9312	**0.9508**	**2.11%**	0.9333	**0.9579**	**2.64%**
LibreOffice	0.8440	**0.9277**	**9.91%**	0.7708	**0.8538**	**10.76%**	0.7022	**0.7703**	**9.69%**	0.7259	**0.8096**	**11.53%**
Linux kernel	0.8943	**0.9578**	**7.10%**	0.8242	**0.8961**	**8.72%**	0.8321	**0.8925**	**7.26%**	0.8274	**0.8942**	**8.08%**
LLVM	0.8388	**0.9296**	**10.82%**	0.7500	**0.8423**	**12.31%**	0.7033	**0.7903**	**12.38%**	0.7115	**0.8154**	**14.61%**
OpenOffice	0.8432	**0.9487**	**12.51%**	0.7508	**0.8969**	**19.45%**	0.7464	**0.8369**	**12.13%**	0.7454	**0.8658**	**16.16%**

Model Evaluation

RQ2: Impacts of Feature Settings

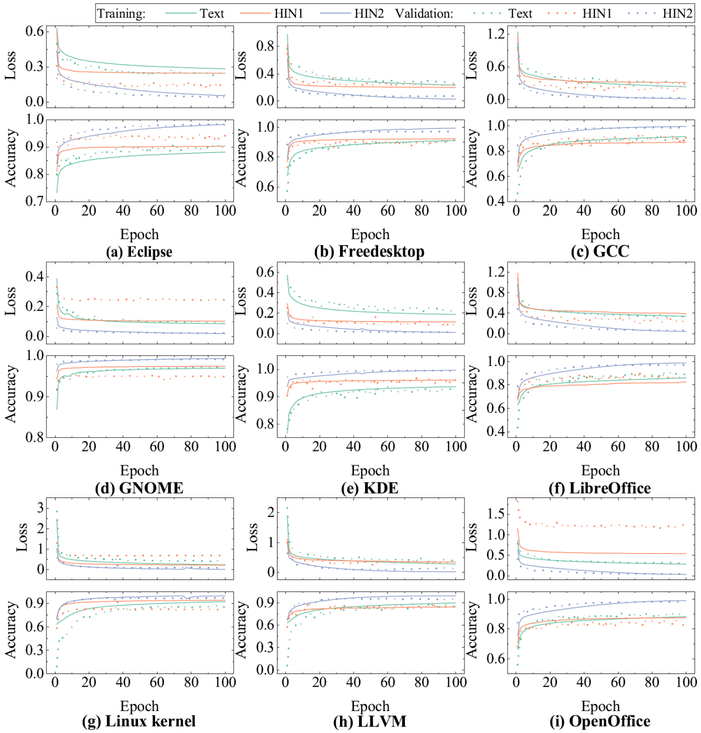

Fig. 8. Comparison of training history under different feature settings.

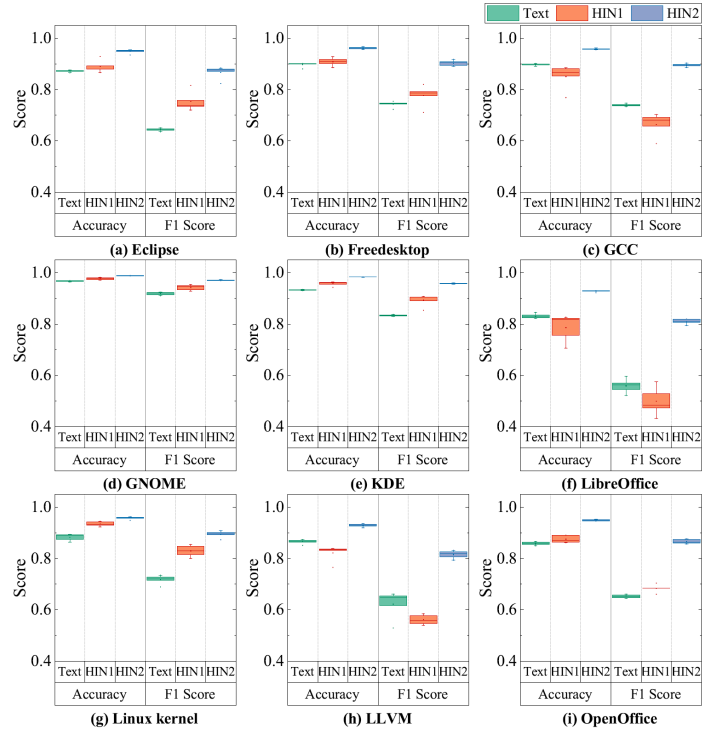

Fig. 9. Comparison of performance under different feature settings.

Model Evaluation

RQ2: Impacts of Feature Settings

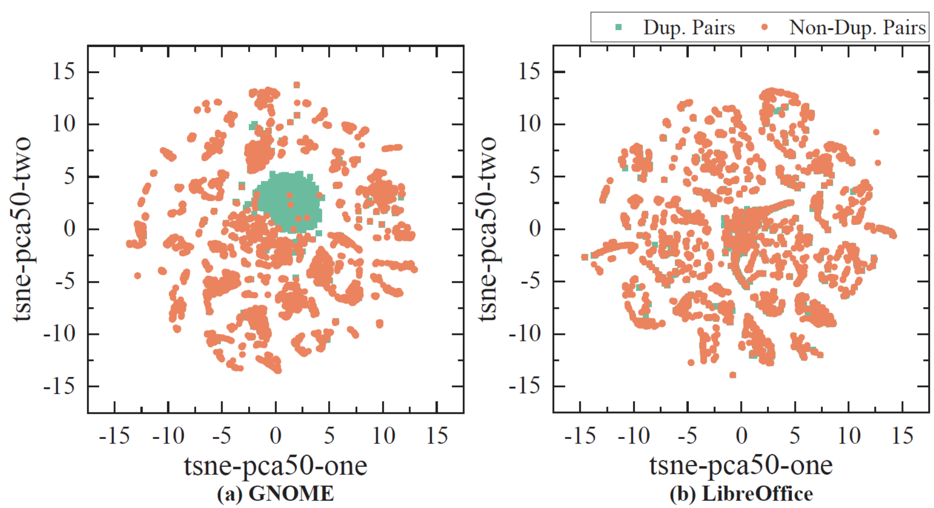

Fig. 10. t-SNE visualization of structured feature vectors of bug pairs.

Model Evaluation

RQ3: Impacts of Before-JIT & After-JIT Duplicates

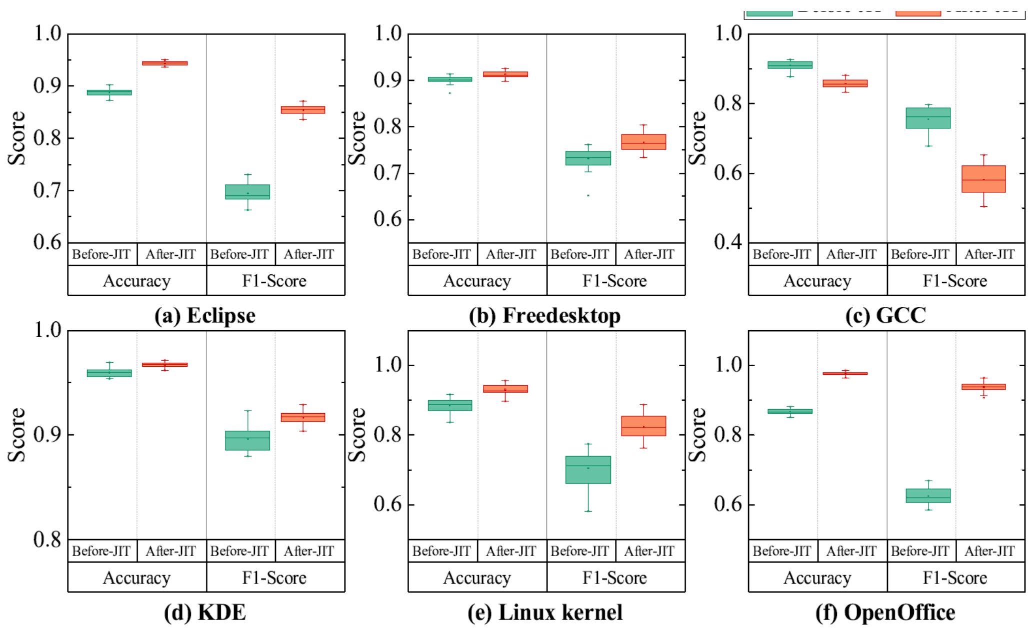

Fig. 11. Comparison of performance on before-JIT and after-JIT datasets.

Thank you for your listening!

Q&A

https://guanpingxiao.github.io/