

Propozycja tematu projektu

Patrycja Kałużna (252864)

Spis treści:

- tytuł, autorzy i źródło artykułu naukowego,
- streszczenie artykułu,
- przeprowadzone badania naukowe,
- wyniki badań,
- ocena reprodukowalności badań,
- ocena potencjału rozwoju badań.

Tytuł, autorzy i źródło artykułu naukowego

84

IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, VOL. 49, NO. 1, JANUARY 2023

DeepLineDP: Towards a Deep Learning Approach for Line-Level Defect Prediction

Chanathip Pornprasit^{ID}, *Student Member, IEEE* and Chakkrit (Kla) Tantithamthavorn^{ID}, *Member, IEEE*

Streszczenie artykułu naukowego

Abstract—Defect prediction is proposed to assist practitioners effectively prioritize limited Software Quality Assurance (SQA) resources on the most risky files that are likely to have post-release software defects. However, there exist two main limitations in prior studies: (1) the granularity levels of defect predictions are still coarse-grained and (2) the surrounding tokens and surrounding lines have not yet been fully utilized. In this paper, we perform a survey study to better understand how practitioners perform code inspection in modern code review process, and their perception on a line-level defect prediction. According to the responses from 36 practitioners, we found that 50% of them spent at least 10 minutes to more than one hour to review a single file, while 64% of them still perceived that code inspection activity is challenging to extremely challenging. In addition, 64% of the respondents perceived that a line-level defect prediction tool would potentially be helpful in identifying defective lines. Motivated by the practitioners' perspective, we present DeepLineDP, a deep learning approach to automatically learn the semantic properties of the surrounding tokens and lines in order to identify defective files and defective lines. Through a case study of 32 releases of 9 software projects, we find that the risk score of code tokens varies greatly depending on their location. Our DeepLineDP is 17%-37% more accurate than other file-level defect prediction approaches; is 47%-250% more cost-effective than other line-level defect prediction approaches; and achieves a reasonable performance when transferred to other software projects. These findings confirm that the surrounding tokens and surrounding lines should be considered to identify the fine-grained locations of defective files (i.e., defective lines).

Streszczenie artykułu naukowego

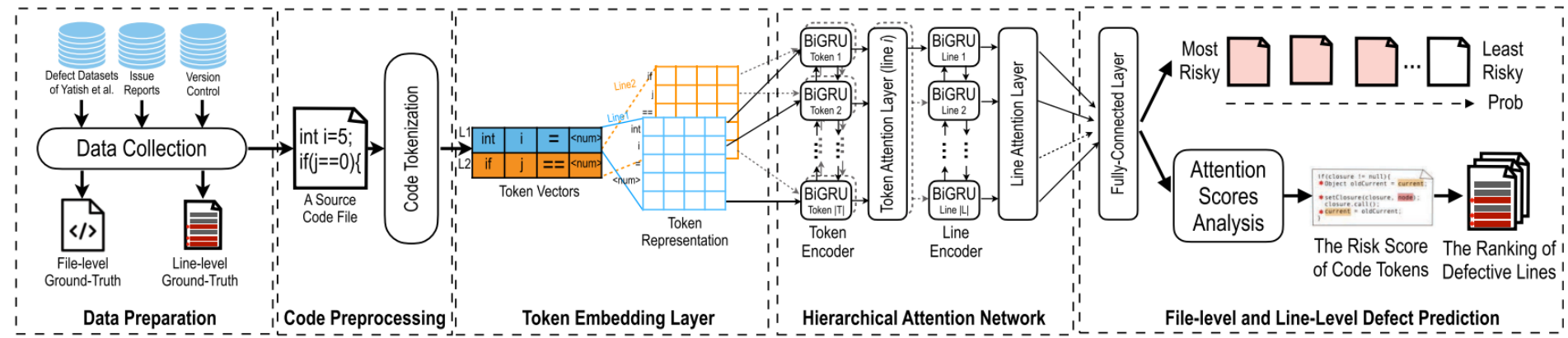


Fig. 3. An overview diagram of our DeepLineDP approach.

Przeprowadzone badania naukowe

- (RQ1) *Can our DeepLineDP be used to differentiate the riskiness of code tokens in defective and clean lines?*
- (RQ2) *What is the accuracy of our DeepLineDP for predicting defective files?*
- (RQ3) *What is the cost-effectiveness of our DeepLineDP for locating defective lines?*
- (RQ4) *What is the accuracy of our DeepLineDP for line-level cross-project defect predictions?*

Wyniki przeprowadzonych badań naukowych

(RQ1) Can our DeepLineDP be used to differentiate the riskiness of code tokens in defective and clean lines?

Tak - wyniki dla tego samego symbolu kodu, pojawiającego się w różnych liniach, różnią się w zależności od lokalizacji tych linii w pliku. Ponadto wyniki dla linii zawierających defekty są znacząco wyższe niż dla linii niezawierających defektów.

(RQ2) What is the accuracy of our DeepLineDP for predicting defective files?

Wartość mediany AUC (Area Under the ROC Curve) wynosi 0.81. Z kolei wartość mediany BA (Balanced Accuracy) wynosi 0.63. Oznacza to, że przedstawiony model jest o 17-37% oraz 3-26% dokładniejszy niż najnowocześniejszy model w kontekście wspomnianych wskaźników.

Wyniki przeprowadzonych badań naukowych

(RQ3) What is the cost-effectiveness of our DeepLineDP for locating defective lines?

DeepLineDP jest o 47-250% wydajniejszy niż najnowsze modele służące do predykcji defektów oprogramowania na poziomie pojedynczych linii.

(RQ4) What is the accuracy of our DeepLineDP for line-level cross-project defect predictions?

DeepLineDP może być wykorzystywany w innych projektach programistycznych i osiąga na nich AUC z zakresu od 0.69 do 0.79 oraz Recall@Top20%LOC z zakresu od 0.31 do 0.46. Może być nieco mniej dokładny i wydajny niż modele nauczone na wcześniejszych release'ach badanego projektu, jednak wciąż osiąga rozsądne wyniki.

Ocena reprodukowalności przeprowadzonych badań naukowych

- **kod źródłowy** zaimplementowanego modelu jest dostępny na zdalnym repozytorium:
<https://github.com/awsm-research/DeepLineDP>
- **zbiór danych** wykorzystany do nauczania zaimplementowanego modelu jest dostępny na zdalnym repozytorium:
<https://github.com/awsm-research/line-level-defect-prediction>

Ocena potencjału rozwoju przeprowadzonych badań naukowych

- możliwość rozszerzenia zbioru danych
- możliwość modyfikacji wartości parametrów modelu
- możliwość modyfikacji architektury modelu

Dziękuję za uwagę!