وزارة التعليم العالي والبحث العلمي
جامعة الموصل
كلية علوم الحاسوب والرياضيات
قسم البرمجيات

تصميم وتنفيذ أداة لاكتشاف نسخ شفرات دارت

رسالة مقدمة
الى مجلس كلية علوم الحاسوب والرياضيات في جامعة الموصل
كجزء من متطلبات نيل شهادة ماجستير علوم في
البرمجيات

من قبل

ياسر محمد خزعل

بإشراف

د. أسماء ياسين حمو

أستاذ مساعد

1443 هـ

2022 م

المستخلص

يساعد استنساخ الشفرات "Code Cloning" عن طريق النسخ واللصق في تقليل الوقت والجهد المطلوب لبناء البرنامج الحاسوبي. ولكن بالمقابل فان هذا النسخ واللصق يؤدي الى زيادة تكلفة صيانة البرنامج وكذلك انتشار الأخطاء بشكل أكبر. نسخ الشفرات ممكن ان يكون هيكلياً او وظيفياً. الهيكلي يكون نسخاً دقيقاً بدون تغيير (النوع الأول) او نسخ مع إعادة تسمية المعرفات (النوع الثاني) وأخيرا نسخ مُعدل يتم التعديل على الشفرة بالإضافة او الحذف (النوع الثالث). اما النسخ الوظيفي فيتم فقط اخذ الوظيفة للشفرة وبنائها بهيكلية مختلفة عن الشفرة الاصلية.

الكشف عن هذه الشفرات المستنسخة يتم عن طريق كاشف الشفرات المستنسخة (Code Clone Detector) وهو عبارة عن أداة يتم استخدامها للكشف عن الايعازات والخطوات المتشابهة والمكررة على مختلف المستويات، وهناك مجموعة من التقنيات التي تستخدم للكشف عن الشفرات المستنسخة.

تم في هذه الرسالة تصميم وبناء أداة للكشف عن الشفرات المستنسخة بلغة دارت (Dart) وهي لغة انتشرت في الآونة الأخيرة حيث تستخدم لبرمجة تطبيقات الهواتف الذكية التي تعمل على نظامي التشغيل الأندرويد والـ IOS بنفس الشفرة عن طريق إطار العمل فلَتر (Flutter).

تكشف الأداة النسخ باستخدام تقنية الاعتماد على النص (Text-based) والاعتماد على المقاطع (Token-based). وتم الكشف على أنواع النسخ الهيكلي (النسخ الدقيق، النسخ مع إعادة التسمية) كذلك تم اقتراح طريقة هجينة تمزج ما بين الاعتماد على النص والاعتماد على المقاطع للكشف عن النوع الثالث من النسخ (النسخ المُعدل). حيث تم تنفيذ الأداة واختبارها بمجموعة من البيانات (الملفات) تم الحصول عليها من GitHub. حُقنت هذه الملفات بنسخ من الشفرات المستنسخة. استطاعت الأداة الكشف عنها واسترجاعها. حيث كانت نسبة إعادة الاستدعاء (Recall) للنوع الأول والثاني بنسبة 100% اما النوع الثالث كانت النسبة 92%.

تم بناء الأداة باستخدام لغة جافا والبيئة التطويرية (NetBeans) وبناء واجهات المستخدم بالاعتماد على (Java Swing). وهي تعمل على جميع الأجهزة التي تحوي ماكنة جافا الافتراضية (Java Virtual Machine (JVM)).

Ministry of Higher Education and
Scientific Research
University of Mosul
College of Computer Science and
Mathematics
Department of Software

Design and Implementation of a Tool for Dart Code Clone Detection

**A Thesis Submitted to the Council of the College of
Computer Science and Mathematics
University of Mosul
as a Partial Fulfillment of Requirements
for the Degree of Master of Science
in
Software**

By

Yasir Mohammed Khazaal

Supervised by

Dr. Asma' a Yaseen Hamo

Assistant Professor

2022 A.D. **1443 A.H.**

Abstract

Cloning codes by copying and pasting helps reduce the time and effort required to build the software. But on the other hand, this copying and pasting lead to an increase in the cost of maintaining the program, as well as the spread of errors. Code Clone can be structural or functional. The structure is an exact copy without change (Exact Clone) or copies with renamed identifiers (Renamed Clone) and finally a modified copy, the code is modified by the addition or deletion (Gapped Clone). The functional Clone, the function is just taken from the code and built in a different structure from the original code.

The detection of these cloned codes is done by the Code Clone Detector, which is a tool that is used to detect similar and repeated steps at different levels, and there is a set of techniques that are used to detect cloned codes.

In this thesis, a tool was designed and built to detect codes cloned in Dart language, a language that has spread recently, as it is used to program smartphone applications that run on the Android and IOS operating systems with the same code through the framework Flutter.

The tool detects code Clones using text-based and token-based technology. Structural types (exact Clone, Renamed Clone) were detected, also a hybrid method was proposed that combines text-based and Token-based to detect the third type of Code Clone (Gapped Clone). The tool was implemented and tested with a set of data (files). Obtained from GitHub These files were injected with copies of the cloned code. The tool was able to detect and retrieve them. The recall rate for the first and second types was 100% and for the third type was 92%.

The tool is built using the Java language and the development environment (NetBeans) and the user interface is based on (Java Swing). It works on all devices that have a Java virtual machine (JVM).