



УНИВЕРСИТЕТ ИТМО

Разработка информационной системы анализа схожести исходного кода решений задач по программированию в открытых образовательных репозиториях

Автор: Степочкин Н.А.

Научный руководитель: Ефимчик Е.А., к.т.н., доцент
Санкт-Петербург, 2019

Актуальность

- Необходимость в анализе схожести исходного кода решений задач по программированию
- Отсутствие возможности напрямую использовать системы анализа схожести исходного кода для образовательных репозиториев

Системы анализа схожести исходного кода

	Moss	JPlag
Метод анализа	Метод отпечатков	Токенизация
Способ запуска	Отправка файлов на сервер	Исполняемый файл
Открытый исходный код	Нет	Да
Вид результатов анализа	Веб-страницы на сервере системы, хранящиеся 14 дней	Файлы веб-страниц
Виды визуализации результатов	Проценты схожести файлов Пары файлов с указанием их схожих участков	
Поддерживаемые языки	23 языка	6 языков

Система Flaxo

- Заново загружает файлы решений при каждом запуске анализа
- Используется только одна система анализа
- Ограниченное время хранения отчета

Требования к системе

- Загрузка файлов репозиторий в систему
- Поддержание загруженных файлов в обновленном состоянии
- Анализ схожести загруженных файлов системами анализа
- Визуализация результатов

Используемые системы контроля версий



GitHub



GitLab



Bitbucket

Используемые инструменты

- Сервер
 - Kotlin
 - Spring
 - Gradle
- Клиент
 - JavaScript
 - React
- База данных
 - PostgreSQL
- Развертывание
 - Docker



Kotlin



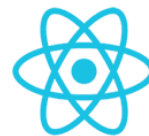
spring®



Gradle



JS



React



PostgreSQL



docker

Страница списка репозиториев

Gitplag

Repositories

Repositories

Add new repository

Name	Git
testns/test_repo	github
thejerome/EPAM_JAVA_Basics_20190225	github
thejerome/IFMO_CPP_programming_20180910	github
thejerome/IFMO_JAVA_Basics_20190225	github

Страница добавления репозитория

New repository

Git

Github Gitlab Bitbucket

Repo name

E.g. myUser/myRepo

Default analyzer

Moss JPlag

Default Moss parameters

See [moss docs](#)

Default JPlag parameters

See [jplag docs](#)

File patterns

.*\.

Split regexps by new lines. Leave empty to download all files.

☒ Enable auto-upload by webhook

Create

Страница репозитория

Gitplag

[Repositories](#) / [thejerome/IFMO_JAVA_Basics_20190225](#)

Analyzes of repository thejerome/IFMO_JAVA_Basics_20190225

Id	Branch	Date	Analyzer	
38	task-1	06/03/2019	jplag	Delete
39	task-2	06/03/2019	jplag	Delete
40	task-3	06/03/2019	jplag	Delete
41	task-1	06/03/2019	moss	Delete
42	task-2	06/03/2019	moss	Delete
43	task-3	06/03/2019	moss	Delete

Run new analysis

Downloaded files

Update files from git

Manage the repository

Страница списка загруженных файлов

Gitplag

[Repositories](#) / [thejerome/IFMO_JAVA_Basics_20190225](#) / [Files](#)

Downloaded files of repository thejerome/IFMO_JAVA_Basics_20190225

Search by regex

Base files [Delete all](#)

Name	Branch	Updated
src/main/java/ru/ifmo/cet/javabasics/BottleSong.java	task-1	2019-02-25 16:52:22
src/main/java/ru/ifmo/cet/javabasics/WAPResult.java	task-2	2019-02-25 17:10:55
src/main/java/ru/ifmo/cet/javabasics/WarAndPeaceExercise.java	task-2	2019-02-25 17:10:55
src/main/java/ru/ifmo/cet/javabasics/WAPResult.java	task-3	2019-02-25 17:15:20
src/main/java/ru/ifmo/cet/javabasics/WarAndPeaceExercise.java	task-3	2019-02-25 17:15:20

Solution files [Delete all](#)

Student	Name	Branch	Updated
student_1	src/main/java/ru/ifmo/cet/javabasics/WAPResult.java	task-3	2019-05-26 14:55:29
student_1	src/main/java/ru/ifmo/cet/javabasics/WarAndPeaceExercise.java	task-3	2019-05-26 14:55:29
student_2	src/main/java/ru/ifmo/cet/javabasics/WAPResult.java	task-3	2019-05-26 12:57:05
student_2	src/main/java/ru/ifmo/cet/javabasics/WarAndPeaceExercise.java	task-3	2019-05-26 12:57:05

Страница запуска анализа плагиата

New analysis

Branch name

Analyzer

Moss

JPlag

Analyzer parameters



Update files before the analysis

Submit

Страница результата анализа

Gitplag

[Repositories](#) / [thejerome/IFMO_JAVA_Basics_20190225](#) / [Analysis #38](#)

Analysis result #38 of repository thejerome/IFMO_JAVA_Basics_20190225

JPLAG Branch task-1 06/03/2019 Source Graph

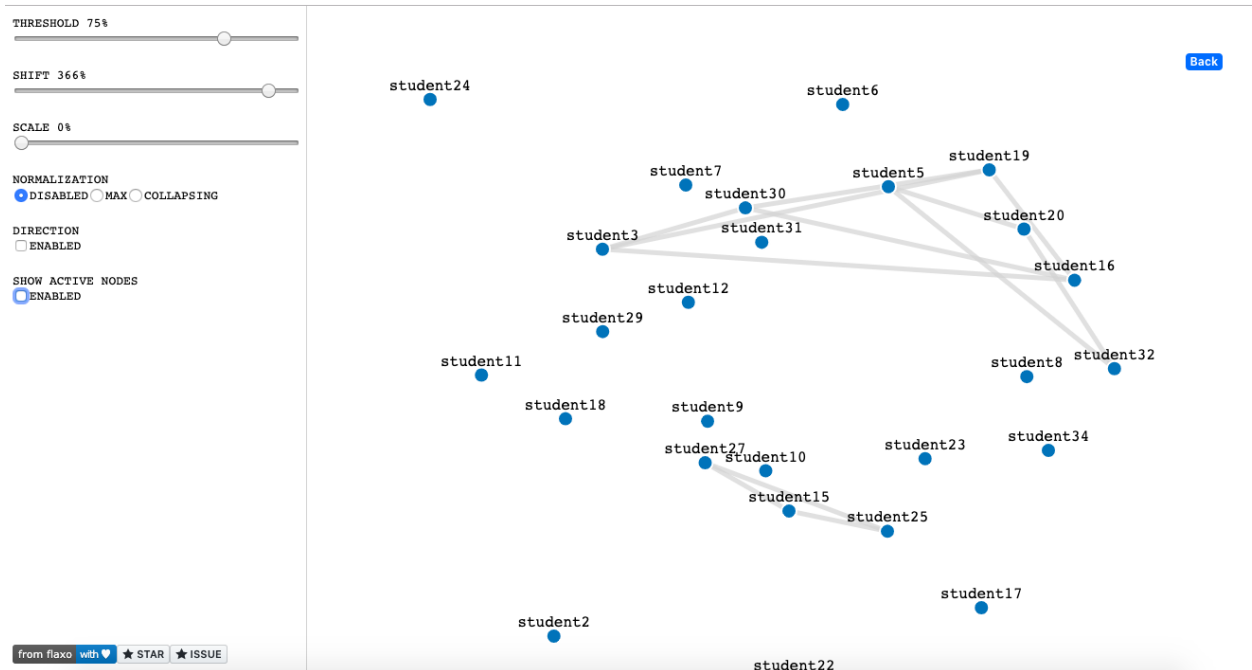
Id	First student	Second student	Similarity
1944	student_3	student_16	100
1945	student_3	student_5	100
1950	student_16	student_5	100
1955	student_10	student_32	100
1956	student_10	student_26	100
1957	student_11	student_7	100
1958	student_11	student_33	100
1959	student_11	student_28	100

Left solutions are made before right ones

Страница сравнения решений

Back Student testns2, created at 2019-03-11 22:58:32	57%	Student nikita715, created at 2019-03-11 19:10:14
<pre> chmokistriptomokki++; } boolean[][] world = gen(); show(world); System.out.println(); world = nextGen(world); show(world); Scanner s = new Scanner(System.in); while(s.nextLine().length() == 0){ System.out.println(); world = nextGen(world); show(world);//comm } } }) </pre>	<pre> 15 58 16 59 17 60 18 61 19 62 20 63 21 64 22 65 23 66 24 67 25 68 26 69 27 70 28 71 29 72 30 73 31 74 </pre>	<pre> public static void main(String[] args) { show(world); System.out.println(); world = nextGen(world); show(world); Scanner s = new Scanner(System.in); while(s.nextLine().length() == 0){ System.out.println(); world = nextGen(world); show(world); //hehe } } }) </pre>
<pre> package files; public class FileTest { public static void main(String[] args) { final int NUM_FACTS = 100; for(int i = 0; i < NUM_FACTS; i++) System.out.println(i + "! is " + factorial(i)); } } </pre>	<pre> 32 75 33 76 34 77 35 78 36 79 37 80 38 81 39 82 40 83 </pre>	<pre> import java.io.InputStreamReader; import java.net.URL; import java.net.URLConnection; import java.util.Scanner; public class URLExpSimple { </pre>

Система data2graph



Система data2graph после внесения изменений

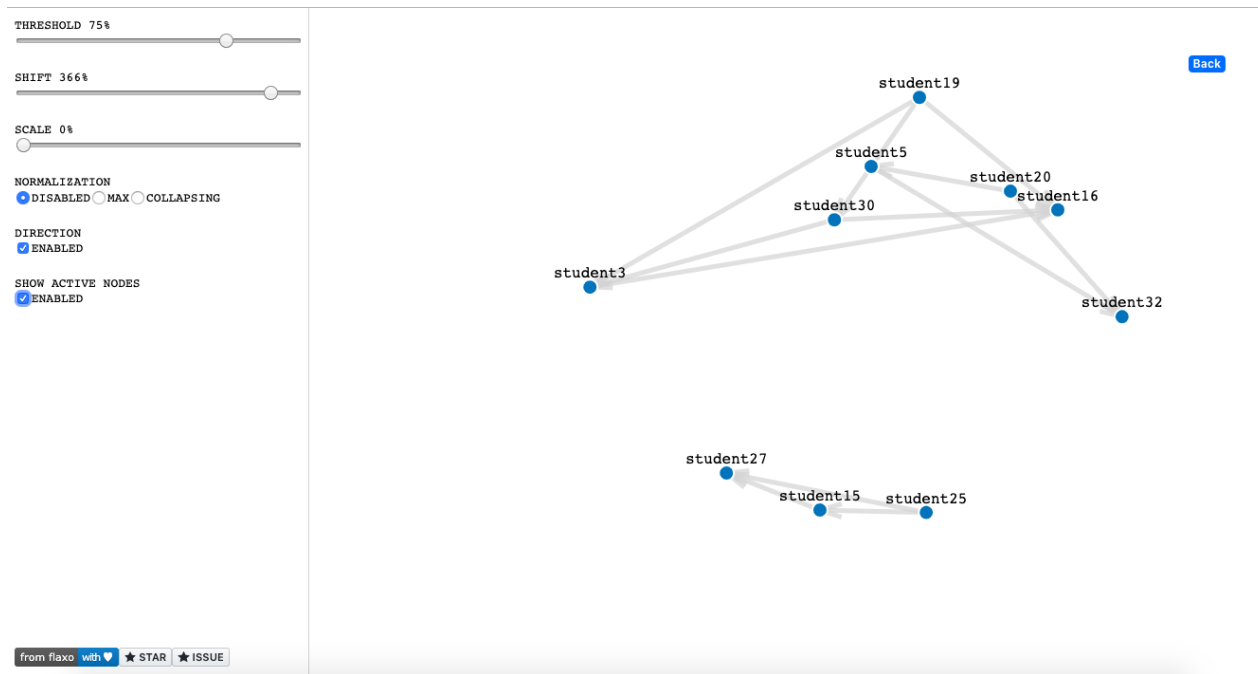
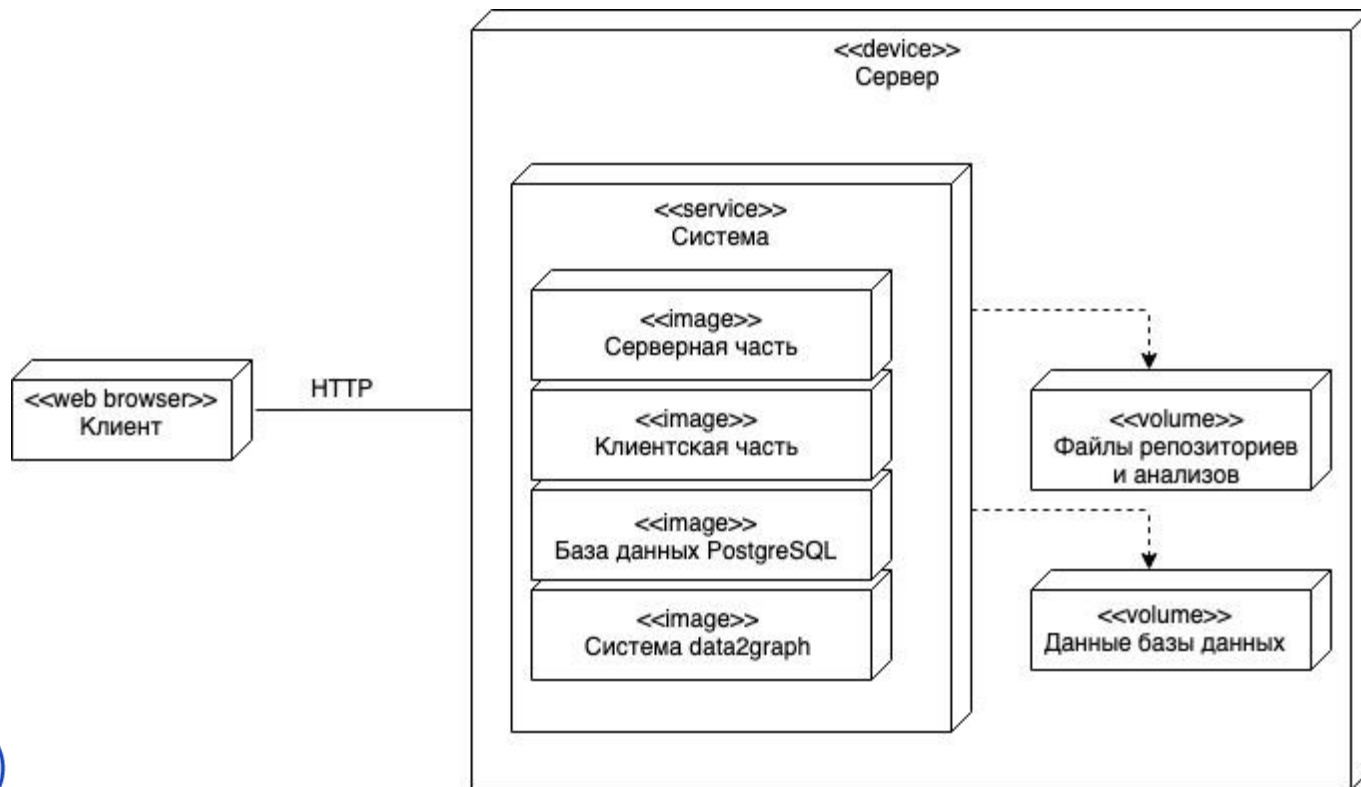


Диаграмма развертывания системы



Список проведенных анализов

Analyzes of repository thejerome/IFMO_JAVA_Basics_20190225

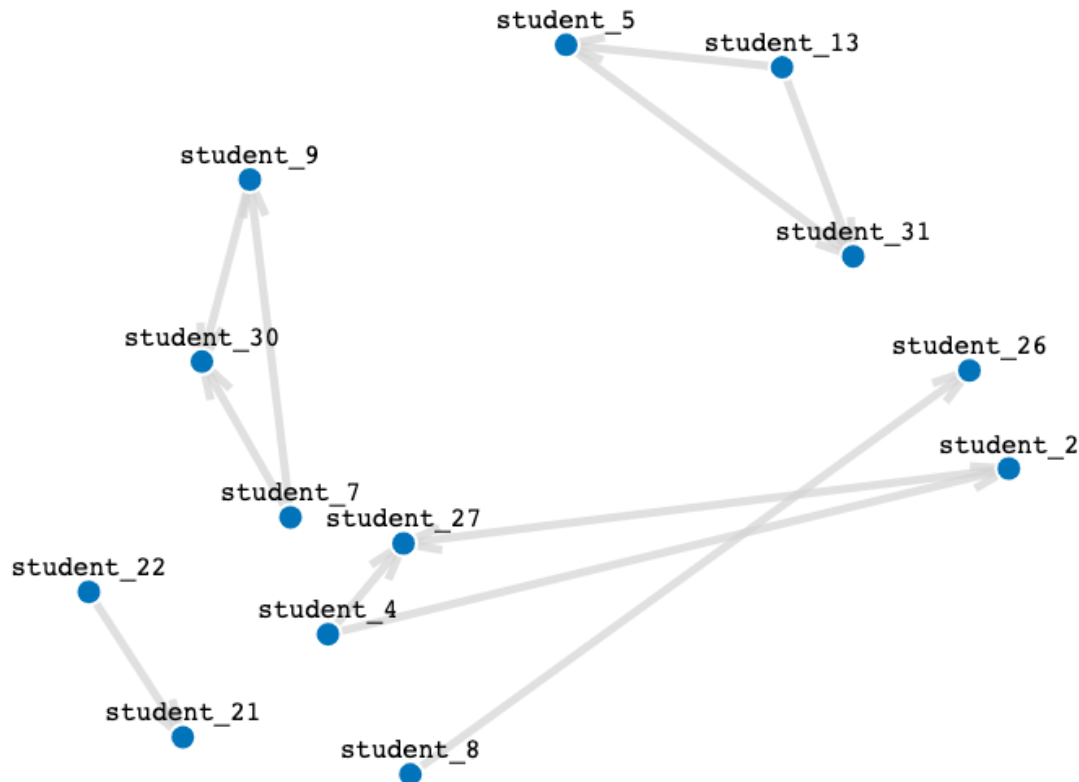
Id	Branch	Date	Analyzer	
38	task-1	06/03/2019	jplag	Delete
39	task-2	06/03/2019	jplag	Delete
40	task-3	06/03/2019	jplag	Delete
41	task-1	06/03/2019	moss	Delete
42	task-2	06/03/2019	moss	Delete
43	task-3	06/03/2019	moss	Delete

Результаты анализов одной из задач

First student	Second student	Similarity
student_9	student_30	78
student_7	student_9	78
student_7	student_30	78
student_8	student_26	75
student_5	student_31	71
student_13	student_5	71
student_13	student_31	71
student_2	student_27	69
student_4	student_2	69
student_4	student_27	69
student_22	student_21	67
student_18	student_25	56
student_3	student_11	54
student_12	student_12	52

First student	Second student	Similarity
student_4	student_27	100
student_2	student_27	100
student_4	student_2	100
student_22	student_21	100
student_18	student_25	100
student_18	student_20	100
student_20	student_25	100
student_8	student_26	100
student_9	student_30	100
student_7	student_9	100
student_10	student_32	100
student_10	student_16	100
student_7	student_30	100
student_16	student_32	100

Граф результатов анализа системой Moss



Заключение

- Спроектирована и разработана система анализа схожести исходного кода в образовательных репозиториях
- Проведено тестирование на реальных данных
- Система готова к использованию в образовательном процессе