Student Submission Integrity Diagnosis [SSID]: User Guide

Table of Contents

1. Introduction

Student Submission Integrity Diagnosis (SSID) is a web application suite for managing courses, assignments, staff, teaching assistants, students, and student code submissions and most importantly**,** to detect and visualize plagiarism among these student code submissions.

1.1. About this User Guide

This user guide provides a quick start guide to use SSID. This user guide provides documentation of all the various features offered by SSID and frequently asked questions. To navigate between the different sections, you can use the table of contents above.

Additionally, throughout this user guide, there will be various icons used as described below:

|	
❗	This indicates an important note. These are essential notes for you to take note when using SSID.
ⓘ	This indicates an additional note. These are notes that are good for you to know but not essential to note when using SSID.

2. Features

Some of the main features of SSID include:

- Pairwise plagiarism detection in submissions
- Clustering analysis
- Plagiarism history of students

Below, you will get more information on how to use SSID to aid you in detecting and preventing plagiarism.

2.1. Accessing SSID

a) First, ensure that you are connected to the SOC VPN through FortiClient. If you need help in setting up your VPN, please click this link. Once you have logged into SOC VPN successfully, you should see this in your FortiClient console:
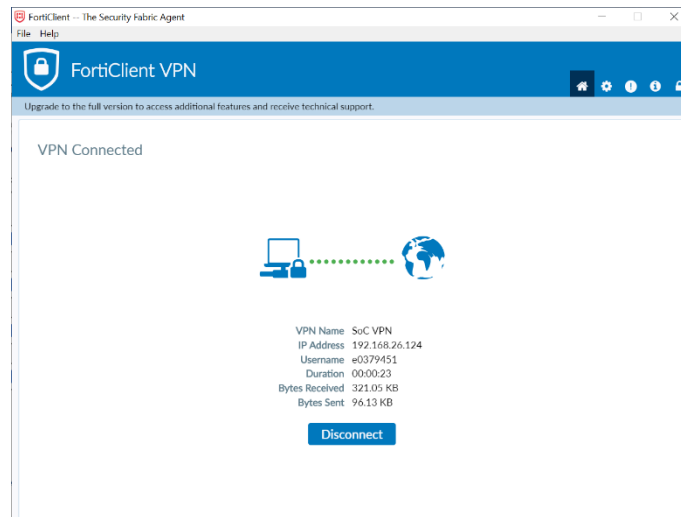
b) After which browse to ssid-i.comp.nus.edu.sg and you will see the below webpage. Please log in to SSID with your given username and password.
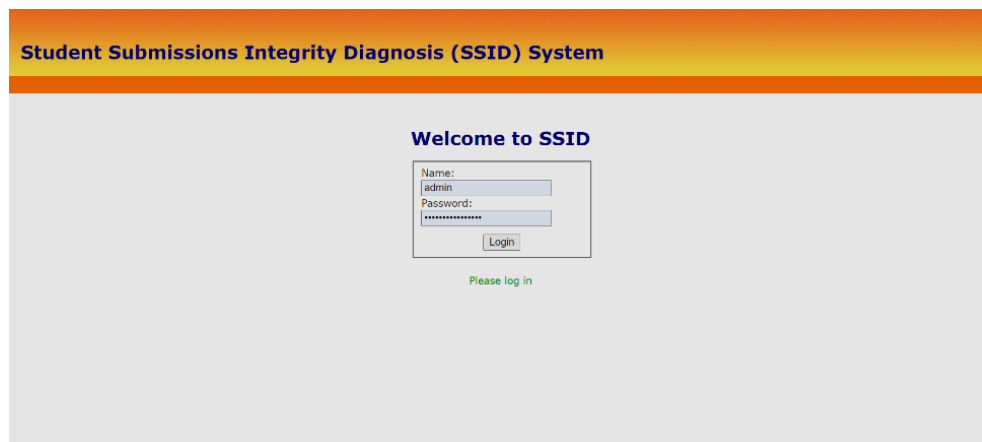
2.2. Viewing assignments

a) To view assignments, click Courses from the top menu.

Student Submissions Integrity Diagnosis (SSID) System

| Home | Courses | Users | Settings | User Guide | Log Out |

b) Under the View tab, click Assignments to view the assignments.

Courses

CODE	NAME	ACADEMIC YEAR (SEMESTER)	YOUR ROLE	ACTIONS		VIEW		
CS1000	Test Platform	2019/2020 Semester 1	Administrator	Edit	Delete	Assignments	Users	Visuals

2.3. Creating assignments

a) To create an assignment, under the assignments listing, click New Assignment to continue.

CS1000 Assignments

No Assignment has been created.

New Assignment   View Visuals
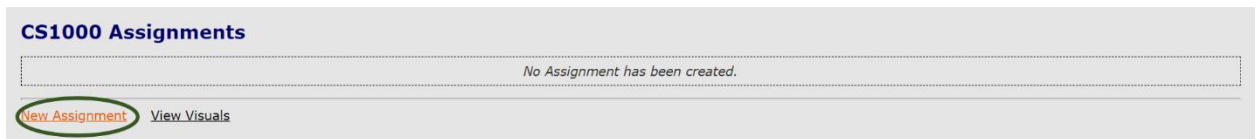
b) Fill in the respective fills and select the zipped submission file (.zip) which contains the student submissions to upload. Then, click Create Assignment to continue.

Courses > CS1000 Assignments > New Assignment

New Assignment

Title	
Language	Java ▾
Minimum Match Length [?]	2
Size of n-gram [?]	4
Student submissions (zip) [?]	Choose File No file chosen

Create Assignment

> ❗ Please ensure that the zipped submission contain 1 folder for each student's code files. Ensure that these folders are named according to the student's ID. If you have any codes/files that you wish to exclude from the system (such as sample code that you gave to the students), put those codes/files under a folder called "skeleton".

	Minimum Match Length refers to the number of words/lines that match between two or more entries. Size of n-gram refers to the size of the consecutive token tiles that are used to match.

	If there are multiple files within a student's folder, do take note that all these files will be merged into a single file by SSID to process it. You would be then able to see this merged file after the processing is done under the Submission Similarities tab which will be explained in the next section.

2.4. Viewing pairwise submissions comparison results

> **!** You must have at least one assignment to view pairwise submissions comparison. To create an assignment, refer to section 2.3

a) To view pairwise submissions comparison for an assignment, firstly, go to the assignments listings. Under your desired assignment, click Submission Similarities under the view tab.

CS1000 Assignments

TITLE	LANGUAGE	SUBMISSIONS	SUSPECTED PLAGIARISM CASES	CONFIRMED PLAGIARISM CASES	VIEW			
Test Assignment	C	3	0	0	Submission similarities	Similarity groups	Upload log	Delete

New Assignment   View Visuals

b) From the rightmost column, click View Submissions of the desired pair of student submissions to view the pairwise submissions comparison result.

CS1000 Assignment: Test Assignment Submission Similarities

- Similarity values are not symmetric so we display both the maximum and average values in the table
- To analyze these similarities, view / create submission cluster groups or use the course visualizations

STUDENT 1	STUDENT 2	MAXIMUM SIMILARITY	AVERAGE SIMILARITY	REMARKS	
as09-001	as09-002	90.909%	46.2065%		View Submissions
as09-001	as09-003	90.909%	45.8745%		View Submissions
as09-002	as09-003	29.925%	23.317%		View Submissions

c) Click on the lines to view the similarities.

Submission Similarities for Students as09-001 and as09-002 (90.909%)

Highlight similarities found in:
Mark these similarities as suspicious or Confirm these similarities as plagiarism

SUBMISSION BY AS09-001	SUBMISSION BY AS09-002	NUM OF MATCHING STATEMENTS	☐	HIDE
Lines 12 — 16	Lines 11 — 15	5	☐	
Lines 18 — 19	Lines 143 — 144	2	☐	

4.	*
5.	* TODO(by student): Brief description of the purpose
6.	* of this program.
7.	*
8.	* @file: digits.c
9.	*/
10.	#include "cs1010.h"
11.	
12.	#define NCOLS 28
13.	#define NROWS 28

|	
11.	#define NCOLS 28
12.	#define NROWS 28
13.	#define WHITE '.'
14.	#define BLACK '#'
15.	#define K 5
16.	#define MAX 784
17.	/**
18.	* This is the structure each sample will be represented as.
19.	* It contains a number which represents the actual number and
20.	* the 2d array corresponding to the drawing.
21.	*/
22.	typedef struct training {
23.	long number;
24.	char **array;
25.	} training;

2.5. Reporting a suspicious case

a) Under the pairwise submissions comparison listing (as seen above), click <u>Mark these similarities as suspicious</u> on the top.

Courses > CS1000 Assignments > Assignment: Test Assignment Submission Similarities > Submissions from Student as09-001 and as09-002

Submission Similarities for Students as09-001 and as09-002 (90.909%)

Highlight similarities found in:
Mark these similarities as suspicious or Confirm these similarities as plagiarism

b) In the confirmation popup box, click <u>OK</u> to report this pair of students as suspicious.

Mark these similarities as suspicious?

Cancel   OK
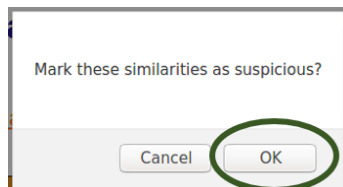
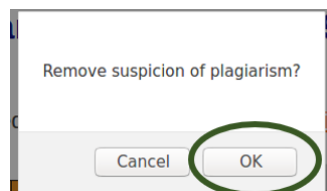## 2.6. Removing a suspicious Case

a) Under pairwise submissions comparison listing, click <u>Remove suspicion</u> on the top.

Submission Similarities for Students as09-001 and as09-002 (90.909%)

Highlight similarities found in:
These similarities were marked as suspicious Remove suspicion or Confirm as plagiarism

b) In the confirmation popup box, click <u>OK</u> to remove this pair of students as suspicious.

Remove suspicion of plagiarism?

Cancel   OK

2.7. Confirming a plagiarism case

|	You must have been assigned with the teaching staff role for the module to confirm plagiarism cases.

a) Under pairwise submissions comparison listing, click <u>Confirm as plagiarism</u> on the top.

Courses > CS1000 Assignments > Assignment: Test Assignment Submission Similarities > Submissions from Student as09-001 and as09-002

Submission Similarities for Students as09-001 and as09-002 (90.909%)

Highlight similarities found in:
These similarities were marked as suspicious. <u>Remove suspicion</u> or <u>Confirm as plagiarism</u>

b) In the confirmation popup box, click <u>OK</u> to report this pair of students as suspicious.

Confirm these similarities as plagiarism?

Cancel   OK
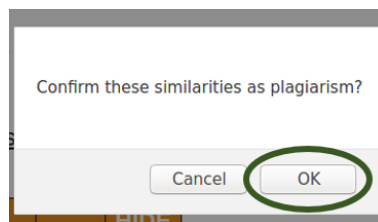
c) Mark the box belonging to the student ID whom is found guilty.

Submission Similarities for Students as09-001 and as09-002 (90.909%)

Highlight similarities found in:
These similarities were confirmed as plagiarism. <u>Remove</u>

Mark student(s) as guilty for plagiarism:
☑ Student as09-001   ☐ Student as09-002

■ SHOW

☐
☐

Submission by as09-001 (Log)   **Submission by as09-002 (Log)**

| 1. | /** | | 1. | /** |
| 2. | * CS1010 Semester 1 AY18/19 | | 2. | * CS1010 Semester 1 AY18/19 |

2.8. Viewing student log

a) Under pairwise submissions comparison listing, click Log beside the student ID of the student.

b) You would be able to see the past history of the student recorded in SSID.

2.9. Creating and viewing similarity cluster groups

!	You must have at least one assignment to create grouping. To create an assignment, refer to section 2.3

a) Under the assignments listing, click Similarity groups.

CS1000 Assignments

TITLE	LANGUAGE	SUBMISSIONS	SUSPECTED PLAGIARISM CASES	CONFIRMED PLAGIARISM CASES	VIEW			
Test Assignment	C	3	0	1	Submission similarities	Similarity groups	Upload log	Delete

b) Click New Cluster Group to create a new similarity cluster group.

Assignment: Test Assignment Submission Cluster Groups

No cluster group has been created.
Cluster groups allow you to identify clusters of students with submission similarity values above a defined threshold.

New Cluster Group

c) Create a cluster group based on your requirements. Then, click Create Submission Cluster group to continue

Courses > CS1000 Assignments > Assignment: Test Assignment Submission Cluster Groups > New Submission Cluster Group

New Submission Cluster Group

Create the cluster group with a cut-off criterion...
- ⦿ To include all currently Confirmed or Suspected Plagiarism Cases
- ○ To include all currently Confirmed Plagiarism Cases
- ○ With a User-Defined Value of: [   ] %

The cut-off criterion provides a threshold value for the cluster group. Any submission cluster with a maximum similarity value lower than the threshold will not be included in the cluster group. Cut-off criterion values to include all currently confirmed or suspected plagiarism cases are computed automatically once submission similarities have been marked as confirmed or suspected for plagiarism.

[Create Submission cluster group]

!	Note that the options "Confirmed plagiarism cases" and "Suspicious or confirmed plagiarism cases" are only available if there exists at least one confirmed plagiarism case and/or reported suspicious case respectively

d) Click <u>View Clusters</u> to view the similarity cluster group created.

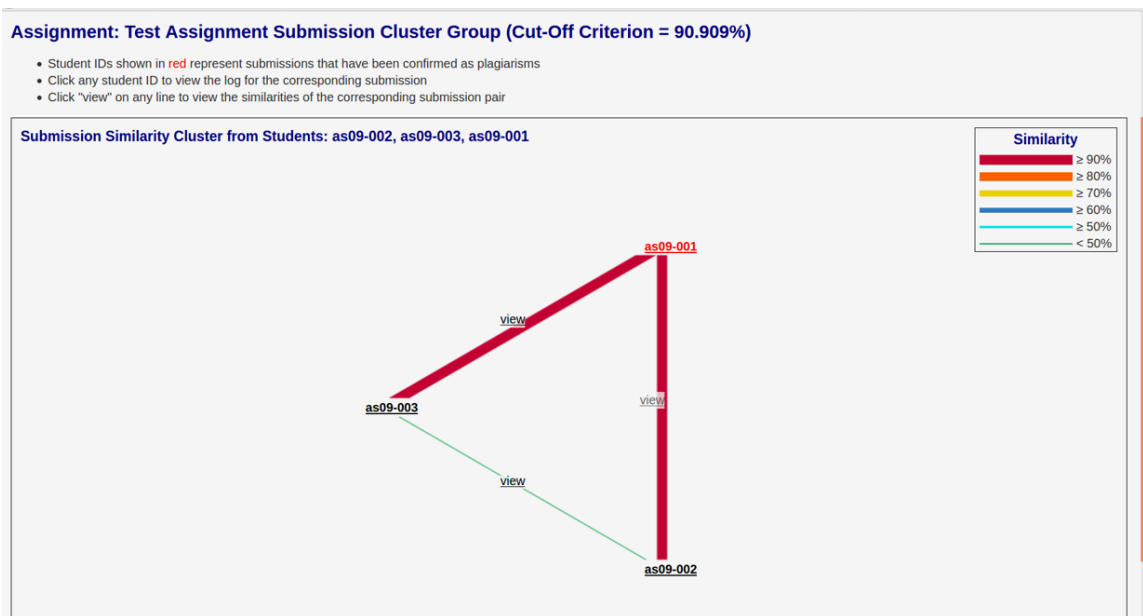Courses > CS1000 Assignments > Assignment: Test Assignment Submission Cluster Groups

Assignment: Test Assignment Submission Cluster Groups

Submission cluster group was successfully created.

CUT-OFF CRITERION	NUMBER OF CLUSTERS	DESCRIPTION	ACTION	
90.909	1	[Created on 2020-09-23 01:25:00 +0800] Based on cases confirmed or suspected as plagiarism	View Clusters	Delete

New Cluster Group

e) View the cluster group to make observations.

Assignment: Test Assignment Submission Cluster Group (Cut-Off Criterion = 90.909%)

- Student IDs shown in red represent submissions that have been confirmed as plagiarisms
- Click any student ID to view the log for the corresponding submission
- Click "view" on any line to view the similarities of the corresponding submission pair

Submission Similarity Cluster from Students: as09-002, as09-003, as09-001

Similarity	
	≥ 90%
	≥ 80%
	≥ 70%
	≥ 60%
	≥ 50%
	< 50%

as09-001

view

view

as09-003

view

as09-002

3. Frequently Asked Questions & Answers

- What is the recommended percentage value to define a pair of plagiarized submissions?

  There exists no static lower-bound percentage of matching. However, based on our experiment, the lowest similarity between a pair of plagiarized submission is 57%, with the minimum-match-length be 2 and size of *N*-gram be 4.