Plagiarism Detection and Prevention Techniques In Engineering Education

Basel Halak and Mohammed El-Hajjar
Electronics and Computer Science
University of Southampton, Southampton, UK
Email: {bh9,meh}@ecs.soton.ac.uk

Abstract— **Plagiarism seriously damages the education process in a number of ways; it prevents students from developing the skills of creative thinking and critical analysis; it undermines the trust between lectures and students, and if goes undetected, it can impact the reputation of the academic institution and devalue its degrees. In this paper, we present two techniques for plagiarism detection and prevention. The first method is based on the allocation of a unique assignment for each student, and the second approach is based on the use of individual presentation of coursework findings. These techniques are applied to three courses at the Master level in the University of Southampton, where we show that they are effective at reducing plagiarism and improving students' understanding.**

Keywords— Plagiarism Detection, Collusion, Assignment, Engineering.

I. INTRODUCTION

Plagiarism is using someone else's work without indicating that it is not own's work or crediting the original author [1]. Plagiarism has a number of negative effects on education. Firstly, it limits the thought, research and critical thinking involved in developing an original paper or report, which negatively impacts the overall educational experience and growth of a college/university student. Second, it damages the relationship between peers and instructors, due to the loss of trust. Finally widespread systematic plagiarism can damage the reputation of the academic institutions and devalue their awarded degrees [2]. Therefore it is vitally important to detect cases of plagiarism and apply appropriate punishments in order to deter students.

The University of Southampton takes plagiarism very seriously and applies a range of penalties to deter students, which can range from failing an assignment or paper to failing a class or getting suspended or expelled [1, 3].

A number of software tools have recently emerged that assist lecturers in detecting the presence of plagiarism [4-9], these tools typically return a measure of similarity for each student report, which approximates the amount of copied material. Human intervention is then needed to determine whether or not the high similarity is due to plagiarism [4]. An example of such software is the "Turnitin" program [10], widely adopted by the academic institutions in the UK , including the University of Southampton.

Although the use of software-based-plagiarism detection approach is very effective in highlighting potential plagiarism cases, which are caused by direct copying of material or computer software codes, it is less effective in the detection of other forms of plagiarism such as ideas-theft or collusion.

This is because most of the plagiarism detection programs evaluate the similarity metric of a paper by comparing it with submissions available in the repository. Hence, these tools are fundamentally incapable of detecting the plagiarism of undocumented ideas. In addition, some of these tools may sometimes fail to detect plagiarised material due to the limitations of their coverage of all relevant literature [11].

There are many scenarios where plagiarism may not necessarily be detected using existing software tools. One example of such a case is when students have their assignments done by someone else. Another more common example is when students share answers and solutions for a common problem in a class-based assignment. This is particularly relevant to a typical engineering coursework, where students are tasked with developing a piece of hardware or software.

Recent advances in source code plagiarism detections [7, 8] have certainly helped in such cases, but if students decide to use different programming languages in their assignment to describe the same algorithm or solution, then detection of plagiarism using these tools becomes much more difficult.

The aim of this paper is to provide an overview of the methods/tools we use in order to enhance the current practices of plagiarism detection and prevention especially in the cases described above.

We present two techniques to help detect and prevent plagiarism in engineering class assignments. The proposed methods are based on the use of unique assignment specifications and individual presentation sessions. These techniques are currently applied to three modules at the Master level in the University of Southampton.

Our evaluation indicates that the proposed plagiarism detection techniques lead to a marked reduction in plagiarism acts and enhance the learning experience of students.

The remainder of the paper is organised as follows. Section 2 outlines the basic concept of the unique assignment method and evaluates its effectiveness. The use of individual presentation sessions is explained in section 3. Finally, conclusions are drawn in section 4.

II. PLAGIARISM PREVENTION USING UNIQUE ASSIGNMENTS

In this section; we present a technique, which we are using to prevent and detect plagiarism in our MSc class assignments.
In order to show the effectiveness of this method, we consider two study cases, namely a digital design project and a software development coursework.

When students in an engineering class are given identical specifications and asked to develop a piece of hardware or software to meet those requirements, they naturally tend to collaborate, and in certain cases this collaboration becomes collusion, where one student puts all the effort to develop a design, which meets the requirements and then his/her classmates would copy the solution, without understanding it. It should be noted here that some cases of collusion could be unintentional; for example, when two students work on the same challenge but fail to declare their individual contributions accurately.

Such type of plagiarism is more difficult to highlight using common plagiarism detection tools. In these circumstances, allocating a unique set of specifications for each student allows students to collaborate without collusion, as they need to develop their own unique solutions, this may greatly reduce cases of collusion-related plagiarism.

Figure 1 shows an exemplar digital design assignment with unique specifications for each student. A more extended version of the same has been used in the system-on chip design course; the latter is part of System-on-Chip design MSc program at the University of Southampton.

Develop an encryption block based on the shift cipher, and implement your design using 0.35um Technology. Each student is given his own design requirements as listed in the table below:

Student Name	Cipher Shift	Speed	Optimization Target
Student A	5	100 MHz	Minimum Area Overhead
Student B	18	125 MHz	Minimum Power Consumption
Student C	7	200 MHz	Minimum Area Overhead

Fig.1 An exemplar digital design assignment with unique specification for each student

As can be seen from the specification in figure 1, each student is required to design a unique cipher with a specific speed.
In addition, students have different optimization targets, so those who need to achieve minimum area overhead will have to use different design and optimizations strategies from others tasked with minimum power consumption designs.

To measure the effectiveness of this approach in reducing collusion-related plagiarism, we have used as a metric the maximum similarity score between all reports submitted to the same assignments, this metric provides a good indicator of the amount of collusion-related plagiarism. This similarity score have been estimated for two digital design assignments using the Turnitin software [10]: the first uses the unique specification approach as in the example in Fig. 1 and the second uses common specifications for all students. In the first case the maximum similarity score was 5% whereas in the second case, the maximum similarity was found to be 21%.
A similar trend has also been observed in a software development coursework for the same cohort, where in a reduction of the inter-reports similarity index of 13% has been observed.

Therefore, the previous description and our reported results indicate that the use of unique specification approach may lead to a marked decrease in collusion-related plagiarism.

III. PLAGIARISM PREVENTION USING INDIVIDUAL PRESENATION

Although the use of unique specification helps reduce cases of collusion-relate plagiarism, it might not detect plagiarisms of undocumented ideas; for example, when students have their assignment done by someone else. In such cases, we propose the use of 'individual presentation technique', where each student is required to explain his/her own results to the class or to the lecturer.

The use of individual presentation can also be effective in detecting plagiarism of undocumented ideas in 'group design projects', which are very common in the engineering education. In these types of assignments, the class is divided into groups, each of which is tasked with developing a complex system in software or hardware.

It is natural in 'group design assignments' for some students to contribute more ideas to the projects than others. Therefore, having individual presentation, where each student is required to explain his/her own contribution can help reduce cases of plagiarism acts, as each student can only claim credits for his/her own contribution to the project.

In order to estimate the effectiveness of the use of individual presentation approach in reducing cases of plagiarism of undocumented ideas we needed to use some metric. The challenge we faced is how to quantify the plagiarism of undocumented ideas, where similarity index metrics are not useful in this case as they only provide an estimate of the amount of material copied from existing sources. Hence, we decided to use the distribution of marks, which tends to be normal distribution for large sized cohorts [12]. This is a natural result of the wide range of skills and abilities of students.

The distribution of marks has been estimated for an assignment in one of the module in MSc Wireless Communications program at the University of Southampton, which contributes 10% to the overall mark of the module. The distributions of marks have been estimated for similarly sized cohorts before and after the use of the individual presentation approach. The results are shown in figures 2 and 3.
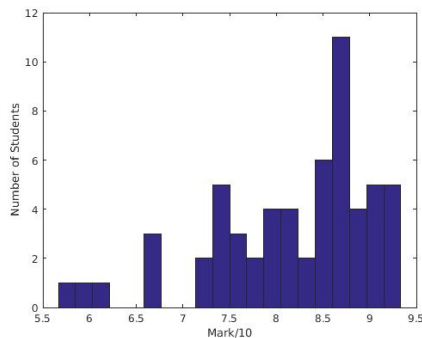

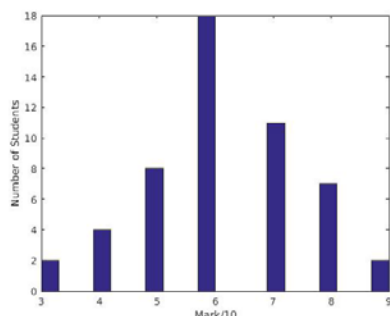Fig.2. Distribution of students' marks before the use of individual presentation method.

Fig.3. Distribution of students' marks after the use of individual presentation method.

Figure 2 shows the distribution of marks for 2014/2015 cohorts with a size of 48 students, which was obtained before the use of individual presentation methods.

In 2015/2016 academic year, students were told they should be ready to prepare a presentation to explain their results if asked to do so. It has been made clear to them that failure to explain and justify any result of finding included in their submitted assignments will trigger a plagiarism investigation case. After the submission of assignments, we asked a sample of students to give presentations. The use of individual presentation approach seems to have a clear effect on the distributions of marks as can be seen from figure 3, where in the marks have a more normal distribution
In 2014/2015 case the average mark was 8.1 out of 10 with a very small standard deviation of 0.9. After the use of individual presentation method in 2015/2016, the average mark dropped to 6.1 out of 10 with a larger standard deviation of 1.4.
The authors strongly believe that the use of individual presentation technique has motivated students to avoid

plagiarising solutions and ideas. That is why; the overall result in 2015/2016 is much closer to a normal distribution, which better represents the wide range of skills and the abilities of large sized cohorts.

IV. CONCLUSIONS

In this paper, we presented two plagiarism detection and prevention techniques based on the use of individual design specification in the course works and individual presentation of the work. We have given examples where these techniques have been used in different MSc modules in the University of Southampton and have shown how these techniques can be effective in reducing the plagiarism and at the same time improving students' understanding.

REFERENCES

[1] University of Southampton Plagiarism Policy: Student Handbook: http://www.southampton.ac.uk/studentservices/handbook/index.html
[2] S.K. Dey, M.A. Sobhan, "Impact of Unethical Practices of Plagiarism on Learning, Teaching and Research in Higher Education: Some Combating Strategies," in 7th International Conference on Information Technology Based Higher Education and Training, pp.388-393, July 2006
[3] University of Southampton Calendar: Plagiarism and Cheating: Policy and Procedures http://www.calendar.soton.ac.uk/sectionIV/sectIV-index.html
[4] M. Li, X. Chen, X. Li, B. Ma and P.M.B. Vitanyi, "The similarity metric," *IEEE Transactions on Information Theory*, vol. 50, no.12, pp.3250-3264, Dec. 2004.
[5] Y.-C. Jhi, X. Jia, X. Wang, S. Zhu, P. Liu and D. Wu, "Program Characterization Using Runtime Values and Its Application to Software Plagiarism Detection," *IEEE Transactions on Software Engineering*, vol. 41, no. 9, pp. 925-943, Sept. 2015.
[6] Z. Tian, Q. Zheng, T. Liu, M. Fan, E. Zhuang and Z. Yang, "Software Plagiarism Detection with Birthmarks Based on Dynamic Key Instruction Sequences," in *IEEE Transactions on Software Engineering*, vol. 41, no. 12, pp. 1217-1235, Dec. 2015.
[7] F. Rosales, A. Garcia, S. Rodriguez, J.L. Pedraza, R. Mendez and M.M. Nieto, "Detection of Plagiarism in Programming Assignments," in *IEEE Transactions on Education*, vol. 51, no. 2, pp. 174-183, May 2008.
[8] G. Cosma and M. Joy, "An Approach to Source-Code Plagiarism Detection and Investigation Using Latent Semantic Analysis," in *IEEE Transactions on Computers*, vol. 61, no. 3, pp. 379-394, March 2012.
[9] M. Jiffriya, M. Jahan, R. Ragel and S. Deegalla, "AntiPlag: Plagiarism detection on electronic submissions of text based assignments," in *8th IEEE International Conference on Industrial and Information Systems* pp.376-380, Dec. 2013.
[10] URL: http://turnitin.com/
[11] C. Kaner and R.L. Fiedler, "A Cautionary Note on Checking Software Engineering Papers for Plagiarism," in *IEEE Transactions on Education,* vol. 51, , pp. 184-188, May 2008.
[12] Sl.L. Hoskins J.C van Hooff, "Motivation and ability: Which students use online learning and what influence does it have on their achievement?", *in British Journal of Educational Technology*, 36(2), 177−192.