Correlation of a Flexible Late Day Policy with Student Stress and Programming Assignment Plagiarism

Samuel Breese Evan Maicus Matthew Peveler Barbara Cutler

Abstract

We investigate the potential for a flexible late submission policy to help address breaches of academic integrity and student stress as they relate to programming assignments. To this end, we examine the use of late days (penalty-free deadline extensions) during our large format, two-semester introductory programming sequence. In particular, we examine correlations between patterns of student assignment submission and student plagiarism. Lexical-analysis-based techniques are used to identify probable cases of plagiarism across student submissions. Furthermore, we present metrics for estimating student stress at submission time. Our metric utilizes the automated grading score of the student's current and previous submission attempts, the imminence of the assignments deadline, and the availability of late days. We corroborate these patterns with self-reported student surveys conducted at the end of the academic year.

Motivation

- In large introductory programming classes, it is important to have well-defined deadlines for regularly-assigned programming assignments.
- Deadlines ensure that students maintain their pace learning the material, promote good time management skills, and facilitate efficient and consistent TA grading.
- Deadline pressures may tempt students to plagiarize assignment submissions from other students or from external sources.
- Novel techniques for introducing leniency to homework deadlines may be useful in reducing student stress, minimizing student plagiarism, and thus maximizing performance.

Late Day Policy

- We use Submitty to implement a flexible late-day policy.
- Each student is given a fixed number of late days at the start of the term.
- Students may use these days to submit an assignment after the deadline.
- We believe that this penalty-free policy does not discourage late submissions. Rather, students are encouraged to finish incomplete work.
- Since Submitty manages late submissions automatically, we can easily experiment with extensions to our policy.

Student Plagiarism

- Our historical data indicates that approximately five to ten percent of students are prosecuted for assignment plagiarism once during the term.
- Each term we have approximately 10 assignments.
- For each assignment, we first compute a measure of similarity for all pairs of students.
- If a student's highest measure of similarity occurs in the top one percent of students on that assignment, we say (for the purposes of our study) that the student is likely to have cheated.

Plagiarism Detection

- In the past, our instructors used the online tool MOSS, provided as a free service by Stanford, to detect likely plagiarism cases.
- We have replicated its functionality in a form more conducive to automated execution and integration with our automatic grading platform.
- Our tool produces a ranked list of pairs of students, sorted by their detected similarity.
- Instructors then manually inspect the top matches to determine whether or not they are actually a violation of academic integrity.

Assignment-Related Stress Metrics

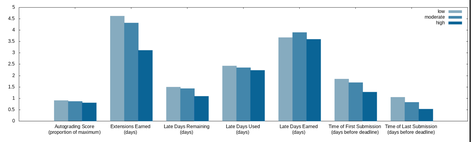
- time of first submission
- time of last submission
- number of submissions
- the autograding score
- available late days to the student for that homework
- earned bonus extension days for that homework

Student Survey

- We conducted a large-scale student survey to provide evidence for the validity of the stress metrics.
- The survey was administered via email, and took approximately 30 minutes to complete.
- We sent the survey to all the students who had completed at least half of the assignments in one of our introductory courses in the previous 2 years.

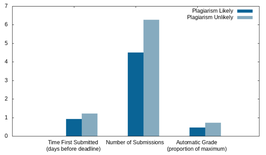
Results

- We compare survey respondents end-of-term stress (low, moderate, high) vs. the proposed metrics for assignment estimates of student stress.



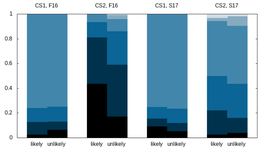
This data confirms our experience as teachers that students who start assignments early, do early test submissions of their assignment, and make their final submission well before the deadline feel less stressed. The data supports use of these metrics as a estimate of student stress.

- Next we compare the stress estimates that can be assessed per assignment with the per assignment incidents of likely plagiarism.



We find a correlation between likely plagiarism and students who start late, use the server less frequently, and earn lower autograding scores.

- Finally, we test the hypothesis that students who have used most or all of their late days on earlier assignments may be more likely to plagiarism on this assignment.



Late days available for assignments categorized as likely or unlikely cases of plagiarism. Black indicates no late days available. Progressively lighter colors indicate progressively more available late days.

Discussion

Stress and Late Days

- On the survey, we directly asked students about assignment-related stress.
- A comparison of student reported stress (no/low, moderate, or high) with our stress metrics is shown at left:
  - 19% of students reported experiencing no/low stress,
  - 62% report moderate stress about homework deadlines, and
  - 19% report high stress about homework deadlines.
- Students also tended to report a positive connection between late day usage and stress:
  - 83% said that the availability of late days decreased their stress about deadlines.
  - 15% reported that late days had no impact on their stress.
  - 2% said that the availability of late days increased their stress.

Plagiarism and Late Days

- It is immediately noticeable that in CS2 F16, student assignment submissions categorized as likely to be plagiarized were much more likely to have been submitted with zero late days remaining.
- However, this pattern is not observable in any of the other semesters.
- A possible explanation for the CS1 F16 and CS1 S17 data is that our CS1 course is significantly less time-intensive when compared to our CS2 course.
- In the CS2 S17 bar, we expected similar late day usage to that in CS2 F16.
- Upon examining that data we found that late days had been awarded much more generously than in the F16 course due to a change in policy.
- In the survey results, students claimed that the availability of late days decreased the incidence of plagiarism:
  - 54% said that the availability of late days decreases the likelihood of plagiarism / file sharing / academic integrity violations.
  - 43% said that the availability of late days has no effect on incidence of plagiarism.
  - 3% said that the availability of late days would tend to increase the rate of plagiarism.

Submitty http://submitty.org

Submitty is an open source programming assignment submission system from the Rensselaer Center for Open Source Software (RCOS), launched by the Department of Computer Science at Rensselaer Polytechnic Institute.

Related Publications

- *Analysis of Container Based vs. Jailed Sandbox Autograding Systems*
  Peveler, Maicus, Holzbauer, and Cutler, SIGCSE 2018 Poster
- *Program Analysis Tools in Automated Grading*
  Dinella, Breese, Maicus, Cutler, Holzbauer, and Milanova
  SIGCSE 2018 Poster
- *Supporting Team Submission and Peer Grading within Submitty*
  Peveler, Breese, Maicus, Aikens, Cyrus, Dinella, Anderson, Barthelmess, Lee, Montealegre, Wang, Holzbauer, Cutler, and Milanova
  SIGCSE 2018 Demo

Acknowledgments

- Rensselaer Center for Open Source (RCOS)
- Red Hat Software
- Google Summer of Code
- https://github.com/Submitty/