Senior Project – Computer Science and Mathematics – 2015

Analysis of the PeerRank Method for Peer Grading

Joshua Kline

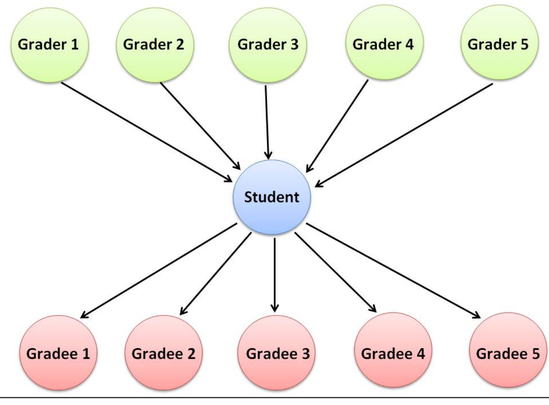
Advisors – Prof. Matthew Anderson & Prof. William Zwicker

Introduction

Peer grading can have both benefits and potential issues.

Pros:
- Faster
- Learning from peers

Cons:
- Inaccurate grades
- Lack of incentive



An example peer grading setup.

PeerRank

PeerRank[1], proposed by Toby Walsh, is a system for producing accurate peer grades using linear algebra similar to that in Google PageRank. Grades are determined as a weighted combination of the peer grades, where the weights used are the graders' own grades. and the grader's accuracy is a component in their own grade.

$$X_i^{n+1} = (1-\alpha-\beta)\cdot X_i^n + \frac{\alpha}{\sum_j X_j^n}\cdot\sum_j X_j^n \cdot A_{i,j} + \frac{\beta}{m}\cdot\sum_j 1 - \left|A_{j,i} - X_j^n\right|$$

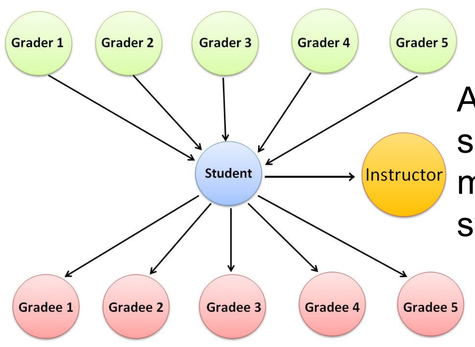
The equation used by PeerRank to determine the grade for student i in iteration $n+1$. $A_{i,j}$ is the peer grade given to student i by student j, X_i^n is the grade given to student i in iteration n, m is the total number of students, and α and β are changeable parameters in the equation. The term with β incorporates the grader's accuracy into their own grade. The equation is repeated iteratively until a fixed point is reached, i.e. until $\overrightarrow{X^{n+1}} = \overrightarrow{X^n}$.

Project Goal

Modify and adapt the PeerRank algorithm to better provide accurate peer grading in a classroom setting.

Ground Truth

- In PeerRank, if a group of incorrect students outnumber a group of correct students, incorrect grades are produced.
- We want to give instructors the ability to establish a basis of "correctness".
- We propose a solution that modifies PeerRank:
  - Instructor submits their own assignment with a known grade
  - Students' accuracies are now determined by how well they grade the instructor, instead of their own grade
  - The grades produced are an average of the peer grades weighted by the graders' accuracies



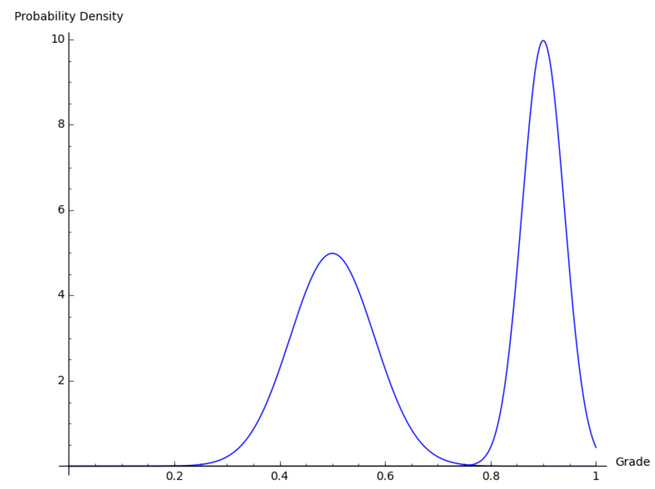
An example setup for the modified system.

$$\vec{X} = \frac{1}{\left\|\overrightarrow{ACC}\right\|}\left(A\cdot\overrightarrow{ACC}\right)$$

The equation used to determine grades in the modified system.

Experimental Method

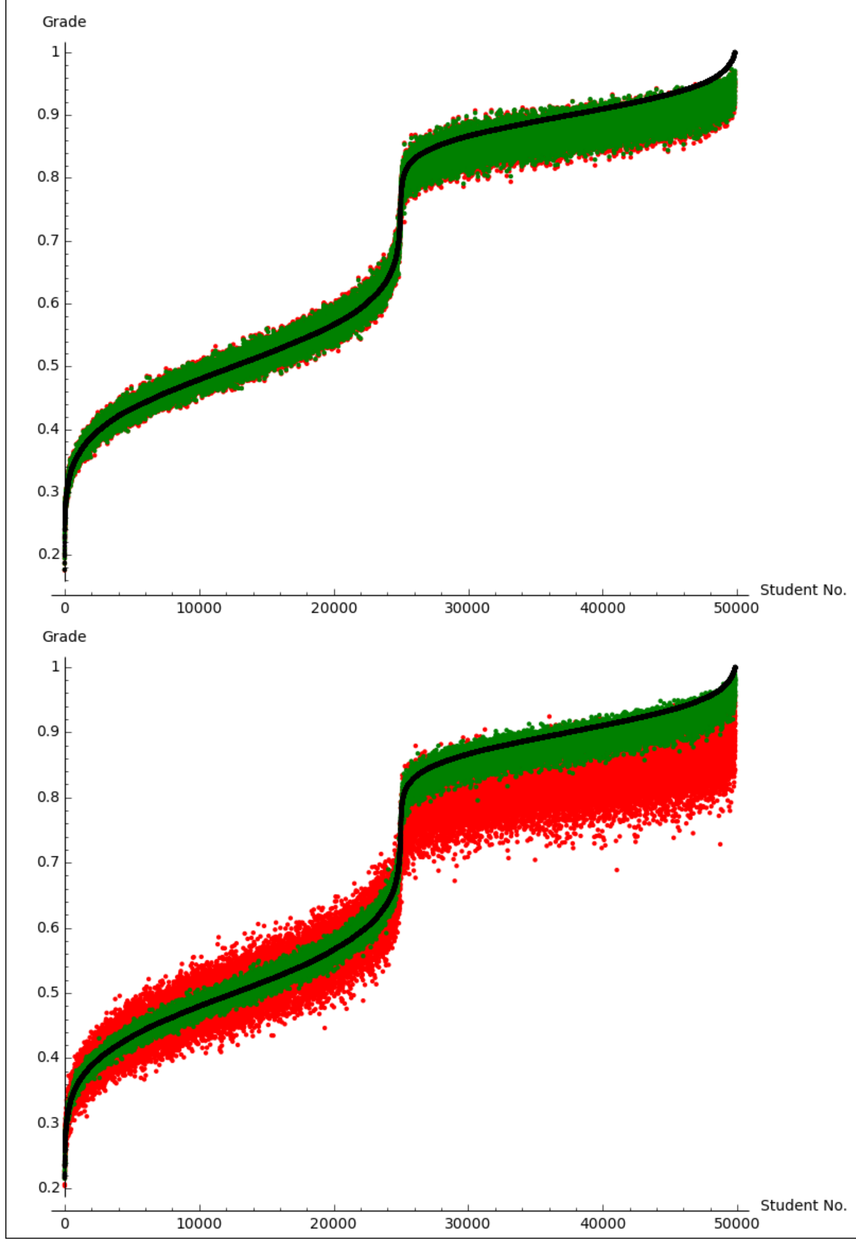
- We want to compare the accuracy of our method in determining correct grades. against that of PeerRank
- Simulated grades were drawn from a bimodal distribution, and accuracies were drawn from normal distributions around the grader's grade[2]. These were then used to generate peer grades.
- Grades were then produced by PeerRank and our method, and compared to the ground truth grades.



The distribution used to simulate ground truth grades.

Results

As we reduce the connection between a grader's grade and their accuracy, our system generates grades that are closer to the correct grades than PeerRank, as our system does not assume this connection. This shows that our method is more accurate in determining correct grades than PeerRank.



Graphs depicting the results of 1000 tests on two different cases. The black line is the ground truth grades, the green area is the grades produced in our system, and the red area is the grades produced by PeerRank. In the first case each grader's accuracy is strongly correlated with their grade, and PeerRank and our system produce similar results. In the second case there is almost no correlation between grade and accuracy, and our system produces results closer to the correct grades than PeerRank.

Future Work

- Implement "partial grading", where each student only grades a subset of the class
- Determine other methods of integrating ground truth

[1] Walsh, Toby. "The PeerRank Method for Peer Assessment." (2014).
[2] Model distributions were suggested by Prof. Roger Hoerl, Department of Mathematics, Union College.