How to Gauge Students' Ability to Collaborate?

Christian Servin[1], Olga Kosheleva[2],
Shahnaz Shahbazova[3], and Vladik Kreinovich[2]

[1]Information Technology Systems Department,
El Paso Community College (EPCC)
919 Hunter Dr., El Paso, TX 79915-1908, USA,
cservin1@epcc.edu
[2] University of Texas at El Paso, El Paso, Texas 79968,
olgak@utep.edu, vladik@utep.edu
[3]Azerbaijan Technical University, Baku, Azerbaijan,
shahbazova@gmail.com

1. Gauging ability to collaborate is important

- In most classes, we test the students' individual knowledge and the individual ability to apply this knowledge.
- However, in the modern world, most problems are solved by collaboration, not individually.
- While the need for collaboration seems to have increased, collaboration itself is not a new phenomenon.
- Many historians believe that the ability to successfully collaborate was the main factor that made our species dominant.
- So, to gauge the students' readiness to solve real-life problems, it is important to gauge:
  - not only their individual abilities,
  - but also their ability to collaborate, to solve the problems in collaboration with others.

2. Gauging ability to collaborate is not easy

- A natural way to gauge the ability to collaborate is to combine students into groups, and to assigns tasks to these groups.
- This way, by grading the result, we can gauge the ability of the group to collaborate.
- The problem is that it is not easy to translate this information into individual grades.
- If a group has been successful, this does not necessarily mean that all members of this group mastered the art of collaboration; so:
  - if we give everyone from a successful group a very good grade,
  - for some students who have not yet mastered this skill very well, the resulting grade will be undeserved.

3. Gauging ability to collaborate is not easy (cont-d)

- Similarly:
  - if a group has not been very successful,
  - this does not necessarily mean that all members of this group deserve a bad grade on collaboration abilities.
- A few of them may be better – and so for them, the bad grade based on the project as a whole would also be undeserved.
- So, a fair estimation of the students' ability to collaborate is still an important challenge.
- In this talk, we provide a possible way to solve this challenging problem.

4. What is given

- We want to gauge the students' ability to collaborate.
- For this, it is important to understand how the group's productivity depends on the students' ability to collaborate.
- For this purpose, let us introduce natural notations.
- For each student i, we will denote:
  - this student's individual skills by s_i,
  - this student's ability to collaborate by c_i, and
  - the amount of effort that the student applied by e_i.
- Based on this data, we want to describe the productivity p.
- In other words, we want to come up with a formula that describes productivity of a group of n people as a function of these inputs:

$$p = p(s_1, \ldots, s_n, c_1, \ldots, c_n, e_1, \ldots, e_n).$$

5. How to come up with a model

- In general, any sufficiently smooth function can be described by its Taylor expansion.
- We want to come up with a simple model.
- So, we will use only the smallest terms in the Taylor expansion which are consistent the commonsense understanding of the situation.
- In general, the first terms in the Taylor expansion are linear terms.
- So, from the purely mathematical viewpoint, it may seem reasonable to use these terms here as well, i.e., to take

$$p = p_0 + \sum_{i=1}^{n} p_{si} \cdot s_i + \sum_{i=1}^{n} p_{ci} \cdot c_i + \sum_{i=1}^{n} p_{ei} \cdot e_i.$$

6. How to come up with a model (cont-d)

- However, from the commonsense viewpoint, this formula makes no sense.
- First, if no one has any skills, individual or collective, there is no productivity.
- So, when $s_i = c_i = 0$, we should have $p = 0$. This implies that for all possible values of e_i, we should have $p_0 + \sum_{i=1}^{n} p_{ei} \cdot e_i = 0$.
- This means that $p_0 = 0$ and $p_{ei} = 0$ for all i.
- Similarly, if none of the students applies any effort, there will be no productivity.
- This implies that $p_{si} = p_{ci} = 0$, so all linear terms should be 0s.

7. How to come up with a model (cont-d)

- From the commonsense viewpoint, the only possibility to get some productivity is:
  - either when at least one student has non-zero individual skills s_i and non-zero effort e_i;
  - the simplest term with this property is the product term $e_i \cdot s_i$;
  - or at least two students $i \neq j$ have non-zero ability to collaborate and apply non-zero efforts;
  - the simplest term with this property is $e_i \cdot c_i \cdot e_i \cdot c_j$.
- We decided to limit ourselves to the smallest non-zero terms – which is usually called the first approximation.
- We thus conclude that the desired expression for p should be a linear combination of terms $e_i \cdot s_i$ and $e_i \cdot c_i \cdot e_j \cdot c_j$.

8. How to come up with a model (cont-d)

- So, for some coefficients $a_i > 0$ and $b_{ij} > 0$, we should have

$$p = \sum_{i=1}^{n} a_i \cdot e_i \cdot s_i + \sum_{i<j} b_{ij} \cdot e_i \cdot c_i \cdot e_j \cdot c_j.$$

- A priori, we have no reasons to believe that some student's skills affect the resulting productivity in different ways.
- Thus, all the coefficients a_i should be equal: $a_1 = \ldots = a_n$.
- Let us denote the common value of a_i by a.
- Similarly, all the coefficients b_{ij} corresponding to different pairs (i, j) should be equal. Let us denote their common value by b.
- Then, the above formula takes the following simplified form

$$p = a \cdot \sum_{i=1}^{n} e_i \cdot s_i + b \cdot \sum_{i<j} e_i \cdot c_i \cdot e_j \cdot c_j.$$

9. Let us simplify this formula

- According to the this formula, the only way the value s_i enters the formula is via the product $e_i \cdot s_i$.
- There is no way to separate these two quantities – and this makes sense.
- If a student does not even try, how can we determine whether this student has the skills?
- So, the only thing that we can observe are not "hidden" skills s_i, but the actually applied skills $\widetilde{s}_i \stackrel{\text{def}}{=} e_i \cdot s_i$.
- Similarly, we cannot observe the hidden ability to collaborate, we can only observe the product $\widetilde{c}_i \stackrel{\text{def}}{=} e_i \cdot c_i$.
- In terms of these "actual" variables, the above formula takes the following simplified form

$$p = a \cdot \sum_{i=1}^{n} \widetilde{s}_i + b \cdot \sum_{i<j} \widetilde{c}_i \cdot \widetilde{c}_j.$$

10. Let us simplify this formula (cont-d)

- We can also re-scale the student-characterizing parameters $\widetilde{s}_i$ and $\widetilde{c}_i$ into $S_i \stackrel{\text{def}}{=} a \cdot \widetilde{s}_i$ and $C_i \stackrel{\text{def}}{=} \sqrt{b} \cdot \widetilde{c}_i$.

$$p = \sum_{i=1}^{n} S_i + \sum_{i<j} C_i \cdot C_j.$$

- Here:
  - the value S_i describes the individual skills of the i-th student, and
  - the value C_i describe the ability of the i-th student to collaborate.

11. What we want

- We can find the observed productivity values p corresponding to different groups – including “groups” consisting of only one student.
- Based on this information, we want to reconstruct the values C_i (and, of course, the values S_i as well).

12. Simplest case: two students

- Let us start with the simplest case of two students.
- In this case, we do not have much of a choice.
- We can give both students individual assignments.
- Thus, by observing the resulting productivity $p_i = S_i$, we can find their individual skills S_i.
- We can also give them a joint assignment, and observe the joint productivity

$$p_{12} = S_1 + S_2 + C_1 \cdot C_2.$$

- Once we know S_1 and S_2, we can therefore determine the product

$$C_1 \cdot C_2.$$

- However, based only on the product, we cannot determine individual numbers C_1 and C_2.

13. Simplest case: two students (cont-d)

- This impossibility makes perfect mathematical sense:
  - we only have 3 possible measurement results p_1, p_2, and p_{12},
  - so we only have 3 equations for 4 unknowns S_1, S_2, C_1, and C_2,
  - not enough to uniquely determine all desired quantities: S_i, C_i.

14. Next simplest case – three students: analysis

- Let us consider the case when we have three students.
- We can give all students individual assignments.
- Thus, by observing the resulting productivity $p_i = S_i$, we can find find their individual skills S_i.
- We can group them into pairs $\{1,2\}$, $\{2,3\}$, and $\{1,3\}$, and observe the joint productivities

$$p_{12} = S_1 + S_2 + C_1 \cdot C_2,\ p_{23} = S_2 + S_3 + C_2 \cdot C_3, \text{ and}$$
$$p_{13} = S_1 + S_3 + C_1 \cdot C_3.$$

- Based on the results of these assignments, we can find the products

$$P_{12} \stackrel{\text{def}}{=} C_1 \cdot C_2 = p_{12} - p_1 - p_2,$$

$$P_{23} \stackrel{\text{def}}{=} C_2 \cdot C_3 = p_{23} - p_2 - p_3, \text{ and}$$

$$P_{13} \stackrel{\text{def}}{=} C_1 \cdot C_3 = p_{13} - p_1 - p_3.$$

15. Next simplest case – three students (cont-d)

- The product $P_{12} \cdot P_{23} \cdot P_{13}$ of all three products is equal to $(C_1 \cdot C_2 \cdot C_3)^2$.
- Thus $C_1 \cdot C_2 \cdot C_3 = \sqrt{P_{12} \cdot P_{23} \cdot P_{13}}$.
- By dividing this product by the known expression for $C_2 \cdot C_3 = P_{23}$, we conclude that

$$C_1 = \frac{C_1 \cdot C_2 \cdot C_3}{C_2 \cdot C_3} = \frac{\sqrt{P_{12} \cdot P_{23} \cdot P_{13}}}{P_{23}} = \sqrt{\frac{P_{12} \cdot P_{13}}{P_{23}}}.$$

- Similarly, we can determine all three values C_i.
- Thus, we arrive at the following method.

16. Case of three students: how to determine the values C_i describing the students' ability to collaborate

- We give each student an individual assignment, and observe the resulting productivity $p_i = S_i$.
- This way, we determine the values S_i.
- We then give each pair of students a group assignment and thus determine the corresponding group productivities p_{12}, p_{23}, and p_{13}.
- Based on these values, we compute $P_{ij} = p_{ij} - p_i - p_j$, and then compute

$$C_1 = \sqrt{\frac{P_{12} \cdot P_{13}}{P_{23}}}; \quad C_2 = \sqrt{\frac{P_{12} \cdot P_{23}}{P_{13}}}; \quad C_3 = \sqrt{\frac{P_{13} \cdot P_{23}}{P_{12}}}.$$

17. Case of three students: possible alternative methods

- For each student i, we need to determine 2 values S_i and C_i.
- So, for 3 students, we need to determine $3 \cdot 2 = 6$ parameters.
- For this, we need to perform 6 experiments – which is exactly what the above method does.
- In addition to these 6 experiments, we could also make a group of all 3 students.
- So overall, we have 7 possible experiments, corresponding to groups

$$\{1\}, \{2\}, \{3\}, \{1,2\}, \{2,3\}, \{1,3\}, \text{ and } \{1,2,3\}.$$

- Let us show that any 6 of these experiments enable us to uniquely determine all the desired values S_i and C_i.
- Indeed, in the above method, we omitted the $\{1,2,3\}$ experiment.

18. Case of 3 students: possible alternative methods (cont-d)

- What if we omit one the individual-measuring experiments?
- Without losing generality, let us assume that we miss experiment $\{1\}$. In this case, we get $S_2 = p_2$, $S_3 = p_3$, and $C_2 \cdot C_3 = p_{23} - p_2 - p_3$.
- We also know the values $p_{12} = S_1 + S_2 + C_1 \cdot C_2$,

$$p_{13} = S_1 + S_3 + C_1 \cdot C_3, \text{ and}$$

$$p_{123} = S_1 + S_2 + S_3 + C_1 \cdot C_2 + C_2 \cdot C_3 + C_1 \cdot C_3.$$

- Here, $p_{12} + p_{23} + p_{13} = 2 \cdot (S_1 + S_2 + S_3) + C_1 \cdot C_2 + C_2 \cdot C_3 + C_1 \cdot C_3$.
- Thus, $p_{12} + p_{23} + p_{13} - p_{123} = S_1 + S_2 + S_3$.
- Since we know S_2 and S_3, we can therefore determine S_1 as the difference $S_1 = (p_{12} + p_{23} + p_{13} - p_{123}) - p_2 - p_3$.
- Once we know S_1, we can determine all the values C_i as above.

19. Case of 3 students: possible alternative methods (cont-d)

- What if we omit one of the paired experiments?
- Without losing generality, let us assume that we miss experiment $\{2, 3\}$.
- In this case, we have all the values $S_i = p_i$, and we also have

$$p_{12} = S_1 + S_2 + C_1 \cdot C_2, \quad p_{13} = S_1 + S_3 + C_1 \cdot C_3, \text{ and}$$

$$p_{123} = S_1 + S_2 + S_3 + C_1 \cdot C_2 + C_2 \cdot C_3 + C_1 \cdot C_3.$$

- Thus, we can determine $C_1 \cdot C_2 = p_{12} - p_1 - p_2$, $C_1 \cdot C_3 = p_{13} - p_1 - p_3$, and $C_1 \cdot C_2 + C_2 \cdot C_3 + C_1 \cdot C_3 = p_{123} - p_1 - p_2 - p_3$.
- Thus, we can find the remaining value $C_2 \cdot C_3$ as

$$C_2 \cdot C_3 = (C_1 \cdot C_2 + C_2 \cdot C_3 + C_1 \cdot C_3) - C_1 \cdot C_2 - C_1 \cdot C_3 =$$

$$(p_{123} - p_1 - p_2 - p_3) - (p_{12} - p_1 - p_2) - (p_{13} - p_1 - p_3) =$$

$$p_{123} - p_{12} - p_{13} + p_1.$$

- Once we know $C_2 \cdot C_3$, we can determine the values C_i as above.

20. General case

- In the general case, we can divide students into groups of 3 and follow one of the above procedures for each triple.
- Some caution is needed.
- As we have mentioned, to determine $2n$ unknowns S_i and C_i, we need to have at least $2n$ results.
- So, we need to perform at least $2n$ measurements.
- It is important to notice that:
  - the very fact that we have performed $2n$ measurements
  - does not necessarily mean that we can uniquely determine all $2n$ values.
- An important counterexample is when all the groups have the same size k.
- Let us show that in this case, the unique determination is not possible.

21. General case (cont-d)

- Indeed, let us show that in this case, the same observations p_g corresponding to different k-element groups $g \subset \{1, \ldots, n\}$ are consistent:
  - not only with the actual values C_i
  - but also with modified values $C_i' = C_i + \delta$.
- Indeed, for each i and j, we have

$$C_i' \cdot C_j' = (C_i + \delta) \cdot (C_j + \delta) = C_i \cdot C_j + \delta \cdot C_i + \delta \cdot C_j + \delta^2.$$

- Thus, if we add up these products for all $(k-1) \cdot k/2$ pairs $i, j \in g$, we get

$$\sum_{i,j \in g,\ i<j} C_i' \cdot C_j' = \sum_{i,j \in g,\ i<j} C_i \cdot C_j + (k-1) \cdot \delta \cdot \sum_{i \in g} C_i + \frac{(k-1) \cdot k}{2} \cdot \delta^2,$$

$$\text{thus} \sum_{i,j \in g,\ i<j} C_i' \cdot C_j' = \sum_{i,j \in g,\ i<j} C_i \cdot C_j + \sum_{i \in g} \delta_i.$$

- Here we denoted $\delta_i \stackrel{\text{def}}{=} (k-1) \cdot \delta \cdot C_i + \dfrac{k-1}{2} \cdot \delta^2$.

22. General case (cont-d)

- Therefore, $\sum\limits_{i,j\in g,\ i<j} C_i\cdot C_j = \sum\limits_{i,j\in g,\ i<j} C'_i\cdot C'_j - \sum_{i\in g}\delta_i.$
- Thus, for each k-element group g, we have

$$p_g = \sum_{i\in g} S_i + \sum_{i,j\in g,\ i<j} C_i\cdot C_j = \sum_{i\in g} S_i + \sum_{i,j\in g,\ i<j} C'_i\cdot C'_j - \sum_{i\in g}\delta_i.$$

- So, for $S'_i \stackrel{\text{def}}{=} S_i - \delta_i$, we get $p_g = \sum\limits_{i\in g} S'_i + \sum\limits_{i,j\in g,\ i<j} C'_i\cdot C'_j$.
- Thus, indeed, the same observations p_g are consistent:
  - not only with the actual values S_i and C_i,
  - but also with different values S'_i and $C'_i = C_i + \delta$.
- Thus, *to uniquely determine the values S_i and C_i, we need to have groups of different sizes.*

23. Acknowledgments

- This work was supported in part by the National Science Foundation grants:
  - 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science), and
  - HRD-1834620 and HRD-2034030 (CAHSI Includes).
- It was also supported by the AT&T Fellowship in Information Technology.
- It was also supported by the program of the development of the Scientific-Educational Mathematical Center of Volga Federal District No. 075-02-2020-1478.