STUDENTS' NETWORKING AND ACADEMIC PERFORMANCE: AN ANALYSIS USING SOCIAL NETWORKS MEASURES

A. Pelaez-Verdet, J. Garcia-Mestanza, M.J. Carrasco-Santos, E.I. Gonzalez-Guerrero

University of Malaga (SPAIN)

Abstract

According to literature, networking plays an important role when it comes to reinforce the student performance. In this paper we try to deepen in this concept, presenting data gathered within the period 2014-2018, coming from a single Human Resources subject, belonging to Tourism Faculty. Thus, data from 994 students was registered, using their final grade as the explained variable. Independent variables were students' partial grades and whether they co-authored works or assignments through their assessment, using this for building a network model in software Gephi 0.9.2. Right after this, we could execute a statistical study and calculate for any given student their measures according to the place he or she was keeping within the network. Students who co-authored more assignments received a higher score in centrality, for instance. We also calculated their respective degree, betweenness, closeness, eccentricity, clustering coefficient, etc. Furthermore, apart from the academic performance variables, we also included in our study dummy variables, such as the language of teaching, or the group they belonged. With this we could calculate a model that suggested that not the group nor the language was relevant for predicting the final score got by the student, since no effect seemed to have on the final grades. On the other hand, their final score appeared significantly linked with four statistical measures associated with their networking behaviour (measured through the papers and assignments co-authored). Namely, these variables were eccentricity and eigencentrality (positive relationship), or degree and clustering (negative relationship). From our perspective, and according to our results, networking cannot be understood as a tool plainly useful in any way, for it needs to be qualified. The student could need to overcome several drawbacks even in the event they freely choose their networking team (e.g. lack of time, disparity of mates, disagreements with other members of the network, disengagement, etc). Turns out that the limitations of this study, like the short period available for analysis, or the need for including information from other fields and courses, also constrains the chances of applying it in a wider context. The study is part of a wider project which intends to improve the understanding of the student's behaviour when it comes to organise their work on their own, the responsibilities they take, and even how they tackle their job search strategy.

Keywords: student's performance, academic behaviour, social network analysis, networking.

1 INTRODUCTION

One of the first authors in pointing out that there are implications to be considered when it comes to design a course -offline, online or blended alike-, were Fulford and Zhang. They suggested that the students' perception on the interaction is associated to their satisfaction on the course [1]. However, their method was constrained to TV courses in a very primitive stage of distance learning. Later, in 2000, Wigfield broadened the scope of the study and found evidence of such that relationship. Later on, the notion of “interaction” was replaced by “networking”. Different theories have been deployed underpinning the idea of the need for networking when it comes to achieve a better academic performance, among the most relevant of them the one carried out by Picciano in 2002 [2], or that presented in 2003 by Yang and Tang [3]. In 2006 we need to mention the work from Mupinga et al., who provided ideas to accommodate the course to previous experiences and background for every student [4], backed up by subsequent works concerning skills e-portfolios [5] or knowledge management [6]. In 2013, else, we can find two relevant papers introducing the idea of student communities, whereas they remark the attention the students pay to the relationship with other students, and its link to their perception about the course quality [7], [8]. Two years later, in 2015, these works were also improved and confirmed by Martin et al [9]. However, it is remarkable that the year before, Moore had yet published an interesting work including performance metrics associated to student satisfaction and interaction [10], which settles down the principles for our study. The two last studies relevant to our objectives were published in 2016, both providing insights on the way the
students tackles their strategies within the subject, alike the way they interact with each other and the instructor [11], [12]. In the end, the idea underneath it is that the broader the student’s network is, the easier to specialise. Factors that might explain this are splitting tasks, sharing knowledge, or giving support each other, which may lead to lower difficulties. However, the research done so far speculates plainly with this, assuming it as a work statement rather than a verified fact. Actually very little research has been done about whether it works at any time, any kind of student or any sort of learning process.

2 METHODOLOGY

The researching team followed a method combining two techniques, namely social networks and linear regression. In a first stage, we could only outline the relationships among the groups of students, using their co-authoring as a measure for their networks. Secondly, once calculated several metrics across the students’ networks, we could run a regressive model for identifying the measures that best predicts a given student’s grades.

2.1 Social networks

Since the goal of this work was to state whether participating in a network could help a student to get a higher grade, we had first to have an accurate insight of how these networks looked like. For that, we gathered data coming from five years (2014-2018) of a single course about Human Resources Management, each of them split into three groups, one of them taught in English and the other two in Spanish every year. Like Year of study, group and final grade achieved, these students were registered according to the relation they had had with another students when asked to co-author assignments. Thus we could draw up a map (see Figure 1) where these relations could be seen and measured, gauging empirically the students’ proximity to each other. As a result we came to display a 994 nodes network, and between them we could register 1004 edges.

This allowed us to assign several relevant measures to any given student, depending on her position in the network. The main descriptive networks indicators calculated for any student were:

- Identification: Id, Label, language, grade
- Centrality: isometric_z_value, degree, weighted degree, eccentricity, closeness centrality, betweenness centrality, eigencentrality
- Dispersion: Authority, hub, modularity class, clustering

The whole network diameter was 7, whereas its modularity rise up to 0.292. Moreover, the graph density was 0.001.
2.2 Regression model

After getting the network analysis, we transferred this information to a SPSS 25 matrix, for running a linear regression multivariable model. 43 students were considered outliers (Using Cook’s Distance), eventually remaining 901 as useful data.

The model tested linear relationships having fixed the grade attained in the course as the explained variable, and the rest of variables worked as independent parameters. At this point we need to say that the group they participated (each of them with a different lecturer) and the language used for teaching were included in the model as dummy variables, so that we could control whether they had significant effects on the student’s grades.

The strategy for model’s creation was “forward”, setting up the confidence level at 95%. If p-value was less than 0.05, then effects were considered, whereas p-values higher than 0.1 discarded the respective effect. Partitioned data was not considered.

3 RESULTS

In the end, all the data boiled down in a model where a number of parameters turned up ineffective for predicting the student’s final score, whilst several of them remained useful for that purpose. The residuals distribution remained normal-shaped.

3.1 Results about language, group or year

These variables probed to be dummy variables, as we could not detect correlation between grades and language, year or group. Thus, it is not likely that utter variations in the students’ grades were due to their presence in a single group or year.

We concluded, at this feature, that being part of a morning of evening group was not a key aspect relevant to determine the future grade to be achieved by any give student. Nor the language of the teaching, neither the year of the course. There was no evidence, indeed, that the changes coming from different teachers or timetables had effect on the students’ performance. Same can be stated regarding the idiomatic constraints of the courses.

3.2 Results about networking parameters

On the contrary, turned out that several calculated parameters from the networks outlined were associated with the students’ performance (measured using their grades).

3.2.1 Uncorrelated variables

Several variables coming from network analysis turned out to be unfeasible, for they did not present a clear and significant relation with the grades attained by the student. In the process for creating the model the following effects were discarded, due that they surpassed the 0.1 p-value:

- Centrality: isometric_z_value, weighted degree, closeness centrality, betweeness centrality
- Dispersion: Authority, hub, modularity class

Those variables probed to be not reliable predictors of the students’ grades, hence we removed them from the analysis on the process of building a sound and parsimonious model.

3.2.2 Correlated variables

On the other hand, when it came to run the following model, we found that the only variables linked to each student’s grades were four: degree, eccentricity, clustering and eigencentrality. The first and third, respectively, showed a reverse effect on the grades, alas the second and fourth seemed to pull the students’ performance in the same sense as they rise.

Degree and clustering seem to play against the academic performance, as if the more the student is related to other students, and the more this relations are contained within a single group, the less the score attained after the evaluation is done. This effect, in regards of the inverse relation, might be explained based on the logic that, provided a certain optimal, convenient relation with a manageable number of mates for co-authoring when presenting assignments, the students who choose to widen too much their academic network and take their chances systematically with another new students,
are much more likely to get worse results on that. This effect might be even aggravated by the clustering effect. If this happens in a certain closed network, restricted in size and where the number of possible compatible mates is not high, students might perceive a lack of chances to find alternative suitable co-authors, lowering the grades because of a self-limitation of the students. They would utterly find themselves in a restricted landscape, where the perception of what is feasible should diminish as the network is more clustered.

On the other side, eccentricity (the maximum number of edges between a node and another one in the network) and eigencentrality (influence of a node in a network, or relation to other influential nodes) are both a measure of connections. It can be said, under the light of these results, that the more a student is connected, and the more who is connected to other also-well-connected students, the more likely to get good grades.

### Table 1. Regression model results

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>t</th>
<th>Sig.</th>
<th>Lower confidence interval 95%</th>
<th>Upper confidence interval 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>64,984</td>
<td>0,645</td>
<td>100,742</td>
<td>.000</td>
<td>63,718</td>
<td>66,250</td>
</tr>
<tr>
<td>Degree</td>
<td>-16,271</td>
<td>4,723</td>
<td>-3,445</td>
<td>.001</td>
<td>-25,540</td>
<td>-7,003</td>
</tr>
<tr>
<td>Eccentricity</td>
<td>8,347</td>
<td>2,783</td>
<td>2,999</td>
<td>.003</td>
<td>2,885</td>
<td>13,809</td>
</tr>
<tr>
<td>Clustering</td>
<td>-271,366</td>
<td>115,760</td>
<td>-2,344</td>
<td>.019</td>
<td>-498,530</td>
<td>-44,202</td>
</tr>
<tr>
<td>Eigencentrality</td>
<td>31,867</td>
<td>14,542</td>
<td>2,191</td>
<td>.029</td>
<td>3,330</td>
<td>60,404</td>
</tr>
</tbody>
</table>

On the previous table it can be seen the results for the model, and the statistics associated to significance and confidence. As it can be stated, once measured a given student’s degree, eccentricity, clustering and eigenvector, it could be easily predicted her expected grade, considering constant the other factor which might influence, in turn, the explained variable.

### 4 CONCLUSIONS

As the model shows, the wideness of the students’ connections (the number of different mates who interact in the process of co-authoring assignments) seems to be linked to their respective grades. In other words, the more students are connected to a given one, the higher the chances of getting good grades. This underpins former proposals coming from previous works, such like those from Moore [10] or Gray and DiLoreto [12].

However, we must remark that there are two different effects who plays an important and significant role, pressing down the marks. If a student presented a high degree in her network, or, else, showed herself to be participant in a clustered network, this might be worse for her.

Should it be the case, for any instructor or lecturer, to base a large part of the grade in co-authored assignments, then it would be good for the academicians to bear in mind those effects. Perhaps keeping an eye on the nature, dimension and structure of networks generated as the course gain momentum could be a good practice. More than avoiding undesired effects when the students produce their strategies for choosing who co-authoring with, it also might help them to balance their activity. Should the lecturers make them realize the importance of setting up a good collaborative network, perhaps the group could move from a subjective perception of the need of caring who they work with into a well-designed strategy for building a good and profitable set of connections.

### ACKNOWLEDGEMENTS

The authors wish to express their gratitude to the members of project PIE 17-152 (University of Malaga), for their contributions.
REFERENCES


