Orderliness predicts academic performance: behavioural analysis on campus lifestyle

Yi Cao¹, Jian Gao¹, Defu Lian², Zhihai Rong¹, Jiatu Shi², Qing Wang³, Yifan Wu¹, Huaxiu Yao⁴ and Tao Zhou¹,²

¹CompleX Lab, Web Sciences Center, ²Big Data Research Center, and ³Key Laboratory for NeuroInformation of Ministry of Education, School of Life Science and Technology, Center for Information in Medicine, University of Electronic Science and Technology of China, Chengdu 617731, People’s Republic of China
⁴College of Information Science and Technology, Pennsylvania State University, PA 16802, USA

Quantitative understanding of relationships between students’ behavioural patterns and academic performances is a significant step towards personalized education. In contrast to previous studies that were mainly based on questionnaire surveys, recent literature suggests that unobtrusive digital data bring us unprecedented opportunities to study students’ lifestyles in the campus. In this paper, we collect behavioural records from undergraduate students’ \( N = 18,960 \) smart cards and propose two high-level behavioural characters, orderliness and diligence. The former is a novel entropy-based metric that measures the regularity of campus daily life, which is estimated here based on temporal records of taking showers and having meals. Empirical analyses on such large-scale unobtrusive behavioural data demonstrate that academic performance (GPA) is significantly correlated with orderliness. Furthermore, we show that orderliness is an important feature to predict academic performance, which improves the prediction accuracy even in the presence of students’ diligence. Based on these analyses, education administrators could quantitatively understand the major factors leading to excellent or poor performance, detect undesirable abnormal behaviours in time and thus implement effective interventions to better guide students’ campus lives at an early stage when necessary.

1. Introduction

A major challenge in education management is to uncover underlying ingredients that affect students’ academic performance, which is significant in working out teaching programmes, facilitating personalized education, detecting harmful abnormal behaviours and intervening students’ mentation, sentiments and behaviours when it is very necessary. For example, it has been demonstrated that physical status (e.g. height and weight) [1–5], intelligence quotient (IQ) [6,7] and even DNA [8–10] are correlated with educational achievement. Accordingly, we can design personalized teaching and caring programmes for different individuals. Since we cannot change a student’s height or DNA via education, more studies concentrate on the aspects of psychology and behaviour, with a belief that learning problems resulting from psychological and behavioural issues can be at least partially intervened. For example, early interventions according to the predictions on course scores or course failures have been discussed recently for K12 education [11–14].

Extensive experiments about relationships between personality and academic performance have been reported in the literature, suggesting that agreeableness, openness and conscientiousness, among the big five personality traits, are significantly correlated with tertiary academic performance, say GPA and course performance [15–17]. In particular, the correlation between...
conscientiousness and GPA is the strongest (at about 0.2) [15–17]. Behaviours are also associated with academic performance. Class attendance has long been known as an important determinant of academic performance [18–22], and additional studying hours are positively correlated with GPA [23–25]. In addition to studying behaviours, some experimental evidences indicate that students with healthy lifestyles and good sleep habits have higher GPAs on average [26–29].

Under the traditional research framework, a large portion of datasets come from questionnaires and self-reports, which are usually of very small sizes (most sample sizes scale from dozens to hundreds), see meta-analysis reviews [15,16,18,23,26]) and suffer from social desirability bias [30,31], resulting in the difficulties to draw valid and solid conclusions. Thanks to the fast development of modern information technology, we have unprecedented opportunities to collect real-time records of students’ living and studying activities in an unobtrusive way, through smartphones [32], online courses [33], campus WiFi [34] and so on. Analyses on these data revealed many unreported correlations between behavioural features and academic performance. For example, watching more of the video and pausing more than once are two strong indicators for better course performance in MOOCs [33], campus WiFi [34] and so on. Analyses on these data revealed many unreported correlations between behavioural features and academic performance. For example, watching more of the video and pausing more than once are two strong indicators for better course performance in MOOCs [33], campus WiFi [34] and so on.

To quantitatively understand the relationships between university students’ behavioural patterns and academic performance as well as the predictive power of the patterns of students’ further academic performance, through campus smart cards, we have collected digital records of undergraduate students’ \( N = 18\,960 \), all of them are pseudonymous) daily activities in the University of Electronic Science and Technology of China (UESTC) from September 2009 to July 2015 (see Data collection in Material and methods for detailed description). The data resolution was reduced before analysis to protect the individuals’ privacy (see Privacy protection in Material and methods). According to the methodology (figure 1) used in this study, we have extracted two high-level behavioural characters from the records, including orderliness (evaluated by the purchase records for showers \( n = 3\,151\,783 \)) and meals \( n = 19\,015\,773 \), which quantifies daily-life regularity and diligence (evaluated by the entry–exit records in the library \( n = 3\,412\,587 \) and fetching water records in teaching buildings \( n = 2\,279\,592 \)), which estimates how long time is spent on studies. Empirical results suggested the significant correlation between academic performance (GPA) and orderliness. Further, we found that orderliness as an important feature improves the prediction accuracy for academic performance even in the presence of students’ diligence. Our work helps education administrators quantitatively understand the major behavioural factors that affect academic performance and provides a promising methodology towards quantitative and personalized education management.

2. Results

2.1. Orderliness

Intuitively, a regular lifestyle would stand us in good stead for college study. In particular, teachers and administrators in most Asian countries, (e.g. Japan, Korea, Singapore, China, etc.) ask students to be self-disciplined both in and out of class [36], and a significantly positive relationship between disciplinary climate and school performance has been revealed [37,38]. Moreover, previous studies based on questionnaires showed that to improve the regularity of class attendance [18,39] and to cultivate regular studying habits [23] will enhance academic performance. However, these studies have not distinguished orderliness in living patterns from diligence in study, since more regular studying habits will result in longer studying time. To our knowledge, a clear and quantitative relationship between orderliness in living patterns and academic performance of college students has not yet been unfolded in the literature. Fortunately, with the large-scale behavioural data, especially the extracurricular behaviours records, we are able to quantitatively measure the orderliness of a student’s campus lifestyle.

According to the dataset, taking a specific behaviour, say taking showers, as an example, if the starting times of taking showers of student \( A \) always fall into the range [21:00, 21:30] while student \( B \) may take a shower at any time, we could say student \( A \) has a higher orderliness than student \( B \) for showers. Next, we turn to the mathematical issue of quantifying the orderliness of a student. Again, considering a specific...
behaviour (e.g. taking showers, having meals, etc.) of an arbitrary student, within total $n$ recorded actions happening at time stamps $\{t_1-d_1, t_2-d_2, \ldots, t_n-d_n\}$, where $t_i \in [00:01, 24:00]$ denotes the precise time with resolution in minutes, and $d_i \in [1$ September 2009, 20 July 2015$]$ records the date. All actions are arranged in order of occurrence, namely, the $i$-th action happens before the $j$-th action if $i < j$. A typical example could be $[21:12—20$ March 2012, 22:02—22 March 2012, $\ldots, 12:10—09$ April 2014$]$. In the analysis of orderliness, we only concentrate on the precise time within a day, say $\{t_1, t_2, \ldots, t_n\}$. We first divide 1 day into 48 time bins, each of which spans 30 min and is encoded from 1 to 48 (specifically, 0:01—0:30 is the 1st bin, 0:31—1:00 is the 2nd bin, $\ldots$). Then, the time series $\{t_1, t_2, \ldots, t_n\}$ can be mapped into a discrete sequence $\{t'_1, t'_2, \ldots, t'_k\}$ where $t'_i \in \{1, 2, \ldots, 48\}$. For example, if a student’s starting times of five consecutive showers are $\{21:05, 21:33, 21:13, 21:48, 21:40\}$, the corresponding binned sequence is $\mathcal{E} = \{43, 44, 43, 44, 44\}$.

In this paper, we apply the \textit{actual entropy} $[40, 41]$ to measure the orderliness of any sequence $\mathcal{E}$ (see Material and methods for details). The actual entropy is considered as a metric for orderliness: the smaller the entropy, the higher the orderliness. The advantages of using actual entropy instead of some other well-known metrics, such as information entropy $[42]$ and Simpson’s diversity index $[43]$, are presented in electronic supplementary material, S1.

Among various daily activities on campus, we calculate orderliness based on two behaviours: taking showers in dormitories and having meals in cafeterias. The reasons to choose these two behaviours are fivefold: (i) they are both high-frequency behaviours so that we have a large number of records; (ii) the data are unobtrusive and thus can objectively reflect students’ lifestyles without experimental bias; (iii) they are not directly related to diligence; (iv) they are less affected by the specific course schedules since any schedule will leave time for meals and showers; (v) most university students in China live and study on campus, and thus the used datasets have sufficient coverage to validate the results. We show the distributions $p(S)$ of actual entropies of students on taking showers in dormitories (figure 2a) and having meals in cafeterias (figure 2b), respectively. The broad distributions guarantee the discriminations of students with different orderliness. We compare two typical students (figure 2c), respectively, with very high orderliness (at the 5th percentile of the distribution $p(S)$, named as student H) and very low orderliness (at the 95th percentile of the distribution $p(S)$, named as student L). As clearly shown by the behavioural clock, student H takes most showers around

![Figure 2](http://rsif.royalsocietypublishing.org/content/journal/rsif/2018/15/20180210.largefig2.jpg)
21:00 while student L may take showers at any time in a day only except for a very short period before dawn, from about 2:30 to about 5:00. We observe a similar discrepancy between two students, respectively, with very high and very low orderliness on having meals (figure 2d). In a word, students with higher orderliness have more concentrated behaviours over time while students with lower orderliness have much more dispersed temporal activities.

In addition to orderliness, we have also considered another high-level behavioural character called diligence, which estimates the effort a student makes in his/her academic studies. Considering the difficulties in quantifying diligence due to the lack of ground truth, we roughly estimate diligence based on two behaviours: entering/exiting the library and fetching water in teaching buildings. Specifically, we use a student’s cumulative occurrences of entering/exiting the library and fetching water as a rough estimate of his/her diligence (see electronic supplementary material, S2 for details). Empirical analysis also demonstrates that the corresponding distributions are broad enough to distinguish students with different diligence (see electronic supplementary material, figure S1).

2.2. Analysis

Intuitively, students with higher orderliness are probably more self-disciplined since orderliness is an intrinsic personality trait that not only affects meals and showers but also acts on studying behaviours. Hence, we would like to explore whether orderliness is correlated with academic performance, say GPA. The orderliness is simply defined as \( O_{E} = -S_{E} \) and both orderliness and GPA are firstly regularized by Z-score [44] (see Material and methods). The relationships between regularized GPA and regularized orderliness (meal and shower) indicate significantly positive correlations (see figure 3). Considering that the relationships between behavioural features and GPA are not simply linear (see electronic supplementary material, figure S3), we apply the well-known Spearman rank correlation coefficient [45] to quantify the correlation strength (see Material and methods). Spearman’s rank correlation coefficient \( r \) lies in the range \([-1, 1]\), and the larger the absolute value is, the higher the correlation is. Spearman’s rank correlation coefficients for meal \( r = 0.182; p < 0.0001 \) and shower \( r = 0.157; p < 0.0001 \) both suggest the statistical significance.

The significant correlation implies that orderliness can be considered as a feature class to predict students’ academic performance. Diligence is also significantly correlated with academic performance (see electronic supplementary material, figure S2) and thus is considered to be another feature class in the prediction model. We apply a well-known supervised learning to rank algorithm named RankNet [46] (see Material and methods) to predict the ranks of students’ semester grades. We train RankNet based on the extracted orderliness and diligence values in one of the first four semesters and predict students’ ranks of grades in the next semester. We use the AUC value [47] to evaluate the prediction accuracy, which, in this case, is equal to the percentage of student pairs whose relative ranks can be consistently predicted with the ground truth. The AUC value ranges from 0 to 1 with 0.5 being the random chance, therefore to which extent the AUC value exceeds 0.5 can be considered as the predictive power. We calculate the AUC values under different feature combinations (table 1). It is noticed that both orderliness and diligence are effective for predicting academic performance in all testing semesters, and the introduction of orderliness can remarkably improve the prediction accuracy even at the presence of diligence. At the same time, we have checked that orderliness and diligence are not significantly correlated (see electronic supplementary material, figure S4). That is to say, orderliness has its independent effects on academic performance. In particular, orderliness is for the first time, to our knowledge, proposed as an important behavioural character that is significantly correlated with a student’s academic performance.
The present report is relevant to education management. On the one hand, understanding the explicit relationship between behavioural patterns and academic performance could help education administrators to guide students to behave like excellent ones and then they may become excellent later on. On the other hand, we can detect undesirable abnormal behaviours in time and thus implement effective interventions at an early stage. The behavioural pattern of students who are addicted to the Internet may be largely different from those without Internet addiction. For example, previous studies have shown that adolescents with Internet addiction have higher irregular bedtimes and dietary behaviour [53], and there is a significant and negative correlation between Internet addiction and academic performance [54,55]. Therefore, identifying Internet addicts at an early stage is critical for effective interventions.

Yet, the current findings are not beyond their limitations on data and method. First of all, some factors that have large effects on GPA could not be captured by our methods such as psychological factors, talent and luck during the exam. Secondly, we do not have the full scope of data that could be used to estimate orderliness (such as bedtimes) and diligence (such as duration of self-studying). Thirdly, our method may underestimate the diligence of some students with different living habits, for example, some students may mainly drink bottled water instead of fetched water, even though they are also taking classes and studying in the teaching buildings. Some students with low orderliness and diligence may exhibit a high academic performance (see electronic supplementary material, figure S5). Therefore, we will collect more relevant data in future works. In addition, we could not establish the causal link between behavioural features and academic performance based on the current data. We expect to reveal causality relations by designing a controlled experiment.

Another interesting yet challenging issue for future study is the generality of our findings across different cultures and educational atmospheres. For example, East Asia creates a higher level of disciplined atmosphere than other cultures, and student academic performance is significantly positively correlated with the disciplinary climate [37,38]. Although in China orderliness is positively correlated with academic performance, whether orderliness is a quality that is predictive across all cultures still remains an open question. Moreover, most undergraduate students in universities in China live in campus dormitories and most of their activities take place within the campus. However, students in other countries may live off-campus or spend a considerable portion of time doing part-time jobs. Accordingly, the ties between collectable behavioural data and academic performance in other countries may be weaker than those in China.

In summary, we hope the reported approaches in this paper, together with some other works [32,35,50–52] in the same direction, will induce methodological and ideational shifts in pedagogy, eventually resulting in quantitative and personalized education management in the future.

### 4. Material and methods

#### 4.1. Data collection

In most Chinese universities, every student owns a campus smart card with real-name registration. The smart card can be used for student identification and serves as the unique payment medium for semester 2 and predict the ranks of examination performance in semester 3.

<table>
<thead>
<tr>
<th>SEMs</th>
<th>SEM 2</th>
<th>SEM 3</th>
<th>SEM 4</th>
<th>SEM 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.618</td>
<td>0.617</td>
<td>0.611</td>
<td>0.597</td>
</tr>
<tr>
<td>D</td>
<td>0.630</td>
<td>0.655</td>
<td>0.663</td>
<td>0.668</td>
</tr>
<tr>
<td>0 + D</td>
<td>0.668</td>
<td>0.681</td>
<td>0.685</td>
<td>0.683</td>
</tr>
</tbody>
</table>
for many consumptions in the campus. In addition, almost all
Chinese undergraduate students live on campus in dormitories
until graduation. In the case of UESTC, the university provides
campus dormitories to all undergraduate students and in principle
does not allow students to live off-campus. Therefore, smart cards
record a large volume of behavioural data in terms of students’
living and studying activities. For the 18 960 anonymous students
under consideration (they cover almost the whole population of
undergraduate students in UESTC, except for very few students
who live off-campus for health reasons or have less than 15 actions
in one or more types of behaviours under consideration), the
data cover the period from the beginning of their first year to
the end of their third year. The data used in this paper contain
four kinds of daily behaviours within the campus. Specifically, there
are 3 151 783 records for taking showers in dormitories, 19 015 773
records for having meals in cafeterias, 3 412 587 records for entering/
exiting the library and 2 279 592 records for fetching water in
teaching buildings, respectively. In addition, some other consump-
tion and entry–exit behaviours are also recorded, including
purchasing daily necessities in campus supermarkets, doing the
laundry, having coffees in cafes, taking school buses, entering/
exiting the dormitories and so on. GPAs of undergraduate students
in each semester are also collected.

4.2. Privacy protection
In the data collection and analyses, we deal with privacy issues
very carefully and tried to avoid infringement of student privacy.
The students are already pseudonymous in the raw data. More-
ever, considering that outside information can be used to link
the data back to an individual if the individual’s spatio-temporal
patterns are unique enough [56,57], we tried to reduce the resol-
tion of the data. For instance, all the information about dates
was removed, the precise happening times of behaviours were
divided into 48 bins. From the data, we only know a student
started to take the shower sometime between 21:00 and 21:30
on some day, while there are about 1000 possible shower rooms,
as well as over 15 cafeterias, over 10 teaching buildings and so on.
After the raw data were processed, it would be reasonably hard
to re-identify individuals by the method reported by Montjoye
et al. [57].

4.3. Actual entropy
We take the actual entropy [40,41] to measure the orderliness of
any sequence $E$. Formally, the actual entropy is defined as
\[ S_E = \left( \frac{1}{n} \sum_{i=1}^{n} A_i \right)^{-1} \ln n, \]
where $A_i$ represents the length of the shortest subsequence starting
from $i$ of $E$, which never appeared previously. If such a subse-
dence does not exist, we set $A_i = n - i + 2$ [41]. Following this
definition, given the binned sequence $E = (43,44,43,44,44)$, we
have $A_1 = 1$, $A_2 = 1$, $A_3 = 3$, $A_4 = 2$, $A_5 = 2$, and thus
$S_E = 0.894$. In this paper, the actual entropy is considered as a
measurement for orderliness: the smaller the entropy, the higher the orderliness.

4.4. Data regularization
The distributions of orderliness and GPA are spread around differ-
ent value scopes. To eliminate the potential effect on correlation
analysis, we use the Z-score [44] to regularize the data, namely,
\[ O'_E = \frac{O_E - \mu_O}{\sigma_O} = \frac{\mu_S - S_E}{\sigma_S}, \]
where $O'_E$ is the regularized orderliness for the student with binned
sequence $E$, $\mu_O$ and $\sigma_O$ are the mean and standard deviation
of orderliness $O$ for all considered students, and $\mu_S$ and $\sigma_S$ are
the mean and standard deviation of actual entropy $S$ for all considered
students. Indeed, orderliness is simply defined as $O_E = -S_E$ under
a monotone and one-to-one relationship. Obviously, $\mu_O = -\mu_S$
and $\sigma_O = \sigma_S$. As a result, the predictability of orderliness and
entropy is the same. Analogously, the regularized GPA for an arbi-
trary student $i$ is defined as
\[ G'_i = \frac{G_i - \mu_G}{\sigma_G}, \]
where $G_i$ is the GPA of student $i$, and $\mu_G$ and $\sigma_G$ are the mean
and standard deviation of $G$ for all considered students.

4.5. Spearman’s rank correlation
In the analysis of relationships between regularized orderliness
and regularized GPA, Spearman’s rank correlation coefficient [45]
is defined as
\[ r_S = 1 - \frac{6 \sum_{i=1}^{N} d_i^2}{N^2 - 1}, \]
where $N$ is the number of students under consideration, $d_i = r(O_i) - r(G_i)$, with $r(O_i)$ and $r(G_i)$ being the ranks for student $i$’s
orderliness and GPA, respectively. Spearman’s rank correlation
coefficient falls into the range $[-1, 1]$, and the larger the absolute value is, the higher the correlation is.

4.6. Prediction approach
Given a characteristic feature vector $x \in \mathbb{R}^d$ of each student, a
pair-wise learning to rank algorithm, RankNet [46], has been
exploited to predict students’ academic performance. RankNet
tries to learn a scoring function $f: \mathbb{R}^d \rightarrow \mathbb{R}$, so that the predicted
ranks according to $f$ are as consistent as possible with the
ground truth. In RankNet, such consistence is measured by
cross entropy between the actual probability and the predicted
probability. Based on the scoring function, the predicted prob-
bility that a student $i$ has a higher GPA than another student $j$
(denoted as $i \succ j$) is defined as $P(i \succ j) = \sigma(f(x_i) - f(x_j))$, where
$\sigma(z) = 1/(1 + e^{-z})$ is a sigmoid function. Here we consider a
simple regression function $f = w'x$, where $w$ is the vector of pa-
rameters. The cost function of RankNet is formulated as follows:
\[ \mathcal{L} = - \sum_{i \neq j} \log \sigma(f(x_i) - f(x_j)) + \lambda R(w), \]
where $\mathcal{L}$ is the regularized term to prevent over-fitting.
Given all students’ feature vectors and their ranks, we apply grad-
ient decent to minimize the cost function. The gradient of the
loss function with respect to parameter $w$ in $f$ is
\[ \frac{\partial \mathcal{L}}{\partial w} = \sum_{i \neq j} \left( \sigma(f(x_i) - f(x_j)) - 1 \right) \left( \frac{\partial f(x_i)}{\partial w} - \frac{\partial f(x_j)}{\partial w} \right) + \lambda \frac{\partial R(w)}{\partial w}. \]

Data accessibility. The dataset needed to evaluate the conclusions in the
paper has been uploaded as part of the supplementary material. The
original data of precise behavioural records, however, cannot be
released in order to preserve the privacy of individuals.

Authors’ contributions. Y.C., J.S., Q.W., Y.W. and H.Y. per-
drived the research. All authors analysed the data. T.Z. drafted
the manuscript. Y.C., J.S., Q.W., Y.W. and H.Y. performed the research. All authors analysed the data. T.Z. drafted the
manuscript. Y.C., J.S., Q.W., Y.W. and H.Y. performed the research. All authors gave final approval for publication.

Competing interests. We have no competing interests.

Funding. This work was partially supported by the National Natural
Science Foundation of China (61603074, 61472086, 61430314,
61802083). D.L. acknowledges the Fundamental Research Funds for
the Central Universities (no. ZYGX2016J087). T.Z. acknowledges the
Science Promotion Programme of UESTC (no. Y0311023901014006).

Acknowledgements. The authors acknowledge the anonymous reviewers for valuable comments and suggestions. The authors would like
to thank Hao Chen, Yan Wang from Nankai University, Qin Zhang,
Junming Huang and Jiansu Pu from UESTC for helpful discussions.
References


