A Dynamic Indicator to Model Students' Digital Behavior

Oriane Dermy, Anne Boyer and Azim Roussanaly
Loria, Université de Lorraine, Nancy, France

Keywords: Students' Digital Behavior, Time-interval Pattern Mining, Sequential Pattern Mining, Change Mining, Pattern Histories, COVID-19.

Abstract: During the first French Covid19 lockdown, students had to switch to a fully online learning mode. Therefore, understanding students' digital behavior becomes crucial for analysts serving public institution policy. In particular, they want to determine and interpret the evolution of students' digital behavior. This paper aims to offer them indicators. We propose to study generic student logs corresponding to standard digital workspace services. Therefore, this paper contributes to the scientific question: *Can we give an easy-to-interpret and visual indicator to model students' behavior changes from poor and generic data?* We first verify that we can extract epidemic-specific temporal patterns on these logs using Contrast Mining. These patterns represent students' behaviors and pace. Then, we propose a new method called Temporal Pattern Histories (TPH), representing the evolution of the temporal patterns' over time. It is a dynamic representation of students' digital behavior. Using this method, we present graphically abrupt changes during the Cov19 lockdown, and we give some hypotheses about these results. This case study proves the relevance of TPH to detect and analyze students' behavioral changes in an interpretive way. This approach has the advantage of representing the global evolution of students' behavior without giving students specific information.

1 INTRODUCTION

From March 2 to August 31, 2020, the French government reported that 25 million people were affected by COVID-19 (Cov19), with approximately 119,500 people being hospitalized. To cope with this epidemic, the government imposed a lockdown from March 17 to May 20, 2020.

Students had to adapt their learning methods to attend distance learning courses abruptly. This online education might have changed students' digital behaviors. Many research studies have then used questionnaires and surveys to analyze and understand how students and teachers perceive these changes. However, only a few research works focus on detecting and understanding these changes, which is what analysts need. Analysts are educational actors working on what actions (e-service development, specific support to local lockdown, etc.) take and how to measure their impact on the students' behavior, as presented in Fig 1. They need indicators to detect and analyze *students' digital behavior* changes in their Digital Working Environment. Such indicators must be easily understandable, interpretive, and respect the anonymity of the students. As Cov19 was not predictable, no specific data has been collected to determine such indicators. It requires working on existing data, mainly generic (as we want to have a global representation of the impact of Cov19 on students' digital behavior) and poor. Thus, this paper aims at answering the scientific question: *Can we give an easy-to-interpret and visual indicator to model students' behavior changes from poor and generic data?*

This paper contributes to this research question with a model,called **Temporal Pattern Histories** (**TPH**), which represents the evolution of the support of temporal patterns over time. The contributions of this paper are (1) Validation of the use of generic and "poor" data to identify students' behavioral changes; (2) Creation of the TPH graphical Data Mining (DM) method; and (3) Application to the Cov19 case study to represent its impact on students' digital behavior.

This paper is organized as follows. Sec. 2 presents the Cov19 research on students' digital behavior. Sec. 3 describes the studied dataset. Then, Sec. 4 develops the first approach with a static DM, followed by the second approach with dynamic DM (Sec. 5). Finally, Sec. 6 and Sec. 7 present discussions about the results, conclusions, and perspectives.

Dermy, O., Boyer, A. and Roussanaly, A.
A Dynamic Indicator to Model Students' Digital Behavior.
DOI: 10.5220/0011039400003182
In *Proceedings of the 14th International Conference on Computer Supported Education (CSEDU 2022) - Volume 2*, pages 163-170
ISBN: 978-989-758-562-3; ISSN: 2184-5026

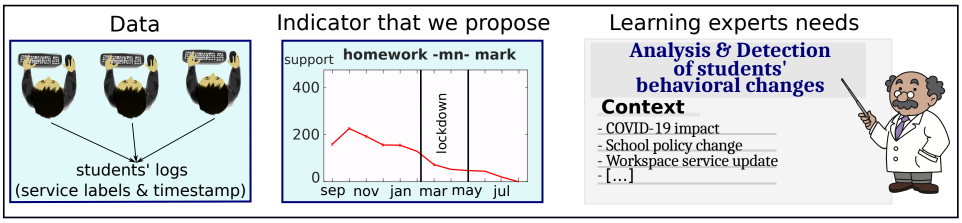

Figure 1: Scientific contribution: A new indicator helping analysts serving the public education policy.

2 LITERATURE REVIEW

The first research papers on the impact of the Cov19 epidemic focus on subjective data, interviewing and analyzing what students and experts think about online learning, either verbally or through surveys. Thanks to these interviews, researchers identify limitations, critical challenges, or factors that facilitate the use of e-learning (Abbasi et al., 2020; Almaiah et al., 2020; Radha et al., 2020). In (Photopoulos et al., 2021), the author goes further to discuss the economic and political impact of e-learning development.

DM methods began to be used in (Reigal et al., 2020) with K-Means and clustering to identify changes in users' habits of a psycho-social assessment platform.

Then, some researchers quantitatively analyzed e-learning usage from more objective data, with a glance at the dynamic evolution of students' behavior. In (Favale et al., 2020) the authors study students' learning and campus network usage during an Italian lockdown. They showed that teams or private chat messages, calls, and meetings increase noticeably during the lockdown. In (Ebner et al., 2020), they analyze the number of online activities, clicks, new publications, and hosts and participants at video conferences. All these activities are studied over predefined sets of periods. Both papers also use dynamic indicators of students' behavior without DM methods.

Other works are closer to our paper, as they use objective data and DM methods. In (Ilieva et al., 2021), data comes from different countries and corresponds to different pandemic periods. They use many methods (classical statistics, machine learning) to cluster students and give statistics about their behavior. In (Lau et al., 2021), researchers adapted the school policy to the Cov19 period, using a revised bloom's taxonomy with flipped classrooms, virtual classroom activities. Then, they analyzed the efficiency of this policy change using machine learning and students' attendance, engagement, and global scores data.

The specificity of our work is that: a) it uses generic, objective and poor data; b) it models dynamic of students' behavior; c) it uses DM methods; and d) our new method doesn't require predefined datasets. Thus, this research work completes previous ones by going further with a dynamic representation of students' behavior using generic and objective data and DM methods.

3 DATASET DESCRIPTION

The experiments below are carried out using real data corresponding to students' logs when using their digital workspace. We had the opportunity to mine data that has been collected from different digital workspaces, initially to monitor e-services. It represents their usage from different middle and high schools all over France. These digital workspaces perform much the same functionality as LMS but only support face-to-face courses and not fully online ones. As a result, collected data are less rich than in LMS.

This paper focuses on data collected during the 2018-19 and 2019-20 school years (from September to August) and on academic institutions located in two distant areas in France. The two chosen areas have different granularities. They model students' behavior independently of the area and granularity. Thus, we will model students' behavior and not educational school policies. The first area is a "*departement*", a territorial division in France, and the second a city.

We represent the evolution of students' digital behaviors monthly to make abrupt changes more visible and normalize results without altering them with temporary disruptions. Data are pre-processed and recorded as temporal or sequential sequences of students' activities. There is one sequence of activities per month and student. A temporal sequence:

$$s_{i,m} = \langle (t_0, E_0), (t_1, E_1), \ldots, (t_n, E_n) \rangle_{i,m}$$

represents the activities of a student i who clicked on n identified workspace services during a month m, where (t_i, E_i) are the identified service type E_i and its timestamp (*i.e.,* when the student clicks on it). The different services are:

«**upload**; collaborative work (**collab**); **mark**; **mail**; **absence**; homework (**hwk**); pedagogical Itinerary (**pedagIt**); school life (**SchooL**); time management (**timeM**) ».

The first approach intends to visualize the emerging / vanishing patterns specific to the lockdown. This study would allow analysts to understand students' behavioral changes. This experiment focuses on the city and we create 4 sub-datasets (before Cov19, lockdown, and same periods previous year):
- $\mathbf{D_{bef}}$: Before Cov19, 01/11-24/12 2019 (53days) - $\mathbf{D_{cov}}$: lockdown, 17/03-10/05 2020 (54 days)
- $\mathbf{D_{bef,18}}$: 01/11/18 - 24/12/18 period (53 days) - $\mathbf{D_{cov,18}}$: 17/03/19 - 10/05/19 period (54 days) We collected 20.000 students' logs per dataset.

The second experiment consists of visualizing and analyzing the history of *s/ti*-patterns (*i.e.,* sequential and temporal patterns respectively). We worked on 10.000 logs per month, when enough logs were available, for the 4 datasets:

- $\mathbf{D_{c18}}$, city, 2018, 111.992
- $\mathbf{D_{c19}}$, city, 2019, 120.000
- $\mathbf{D_{d18}}$, departement, 2018, 120.000
- $\mathbf{D_{d19}}$, departement, 2019, 120.000

3.1 Algorithms Selection

Pattern mining (PM) methods are numerous and solve different challenges, as presented in (Han et al., 2012). We choose methods that mine sequential databases (our data type).

3.1.1 Sequential Pattern Mining (SPM)

SPM has been extensively used in Educational Data Mining (Anjum and Badugu, 2020). With the same objective than us, the authors in (Gutierrez-Santos et al., 2010; Poon et al., 2017) want to identify frequent patterns of students' activities. This field of research aims to discover frequent *s*-patterns (Agrawal and Srikant, 1995). A sequential database is composed of a set of sequences s="$E_1E_2 \dots E_x$", with $E_i \in E$ being the set of ordered events. An *s*-pattern is a sub-sequence, say $p =$ "$E_iE_jE_k$" contained in at least k sequences (k is the *minimum support*[1]). In this paper, we select the *PrefixSpan* method, which is one of the most efficient and commonly used algorithms, based on the *pattern-growth* method (Pei et al., 2001).

However, we also want to use the timestamping of logs to improve patterns.

3.1.2 Temporal Pattern Mining (TI-PM)

Other approaches add to classical *s*-patterns contextual information (Wang et al., 2018; Dong et al.,). In our case, we use logs' timestamps. Many methods add temporal information to discovered patterns, as reviewed in (Dermy and Brun, 2020). The authors' conclusions show that to model students' behavior and pace, we can complete *s*-patterns with *ti*-patterns using the TI-PM method. This takes into account gap values between events and groups them in predetermined time intervals. A *ti*-pattern is defined as: $\alpha =$"$E_i\tau_1E_j\tau_2 \dots E_k$", where $E_i \in E$ is the set of events for $1 \leq i \leq l$ and $\tau_i \in TI$ the set of time-intervals. The sequence α is a time-interval pattern if $support(\alpha) >= \delta$. Two algorithms have been developed by (Chen et al., 2003) to mine *ti*-patterns. For our study, we select the *I-PrefixSpan* algorithm.

TI-PM and SPM have the disadvantage of statically modeling students' behavior. We decide to model students' behavior dynamically to represent the evolution of *s/ti*-patterns as a function of time, thanks to *Contrast* and *change DM*.

4 CONTRAST DM APPROACH

Contrast DM methods aim to find "contrast patterns" describing significant differences in patterns found between datasets. Datasets differ temporally, locally, or through different contrasting conditions (*e.g.,* user groups). Many algorithms exist for Decision Trees, Clustering, or PM (Boettcher, 2011).
This section aims to answer the research question:
Can We Automatically Extract from Poor and Generic Data Students' Digital Behavior Specific to Cov19?

Proposal: Dataset Comparison using Contrast TI-Pattern Mining.
To provide an answer, we statically compare the results of DM performed on the pre-lockdown dataset ($\mathbf{D_{bef}}$) with the one recovered during the lockdown ($\mathbf{D_{cov}}$). To differentiate changes related to Cov19 with the ones related to the school period, we compare the same periods during the previous year ($\mathbf{D_{bef,18}}$ and $\mathbf{D_{cov,18}}$). The comparison focuses on *ti/s*-pattern changes detected in students' logs. We also check if Cov19 impacts students' pace by calculating, for

[1]The support is the sequence number in which the pattern appears.

each time interval, the proportion of frequent patterns which contain this time interval.

These proportions are compared across datasets. The following experiment intends to discover students' behaviors specific to the lockdown: patterns from the $\mathbf{D_{cov}}$ dataset should have more differences than between the other datasets.

4.1 Experiment

Experiment: Research of Students' Digital Behaviors Specific to the Lockdown.
We perform the following experiment, with a minimum support variable equal to 30 and using the *PrefixSpan* and *TI-PrefixSpan* algorithms (Gao et al., 2008; Fournier-Viger et al., 2016). This experiment statically compares *ti*-patterns of each dataset (*c.f.,* Table 1). For the sake of clarity, this Table 1 presents only some representative patterns (most frequent/longer patterns), but changes between other patterns for all datasets follow the same trend. Results are presented as follows: each cell comprises the two most frequent or longer patterns, followed by ":" and its support (*e.g.,* "mark mn mark:2014"). Long and repetitive patterns (rows 4 and 5) are abbreviated with dots (*e.g.,* "mail mn mail mn ...mn mail:38 (s7)", where "(s7)" is the pattern-size). Students' pace is approximated in the 6^{th} row (proportion of time intervals in the *ti*-patterns). The last row presents statistics about each service page.

Results: Yes, We Can Extract Students' Digital Behavior Specific to the Covid Period.
For each dataset, we first notice that for each frequent temporal pattern "$E_1\tau_1E_2$", its symmetric "$E_2\tau_1E_1$" is also frequent with similar support. It is visible in Table 1 by patterns presented in row 3, and it is also valid for other patterns. For example, the symmetric of "*mail mn Ped.It:495*" exists with similar support: "*Ped.It mn mail:402*". Thus, the service order doesn't seem important for students even if the pace (here I_x) is. During Cov19, students' digital behavior is suddenly changing:

1. The most frequent pattern goes from "*mark mn mark*" (3 other datasets) to "*mail mn mail*", and the support of "*mark mn mark*" became 210, which is much smaller (not shown there). The lockdown might cause these changes since teachers communicated by mail. Moreover, analysts interpret and confirm that teachers decided to give no marks to students during the lockdown. The second more important pattern is "*mail mn mail mn mail*", which confirms that sequence of mails is really frequent. It could suggest to analysts that i) students need to check their mails often to follow instruction mails sent by teachers; ii) students were a lot stressed and clicked many times on mails. This second hypothesis is rejected because students' pace (row 6) didn't accelerate.

2. The most frequent patterns without duplicates go from "*mark mn hwk*" and its symmetric to "*mail mn Ped.It*" and "*hwk mn mail*". This result confirms that students have few marks during this lockdown. It might also be because, when students began to work in a fully online mode, teachers give instruction to use the Pedagogic Itinerary service by mails with some additional information and because students return their homework by mails.

3. The two longest patterns without duplicates highlight again that the *Pedagogical Itinerary* service is a lot used during Cov19.

4. Regarding students' pace (row 6), the comparison of the Cov19 period with others shows that the proportion of "*second*" interval decreases and the "*hour*" one increases. About the hour interval, it might suggest that students check services between each online course (one or more hours). The decrease of the "second" interval may be because schools' Internet connections are sometimes saturated, forcing students to click on the same pages several times.

For each result, we gave some reasons that could explain students' behavioral change. However, they require validation by analysts. To highlight the increase of pattern change during Cov19, we study the percentage of "similar s/ti-patterns" between $\mathbf{D_{bef,18}}$ and $\mathbf{D_{cov,18}}$ and that between $\mathbf{D_{bef}}$ and $\mathbf{D_{cov}}$.
We consider that two patterns (p1 and p2) from two different datasets are similar if (p1 = p2) and

$$\frac{|support\,(p1) - support\,(p2)\,|}{\frac{support(p1)+support(p2)}{2}} < 0.5$$

Table 2 shows that the percentage of "similar" *ti*-patterns is littler between $\mathbf{D_{bef}}$ and $\mathbf{D_{cov}}$ than between $\mathbf{D_{bef,18}}$ and $\mathbf{D_{cov,18}}$. Thus, the change rate related to Cov19 is more significant than we could expect from 2018.

5 CHANGE DM APPROACH

The previous Contrast Mining approach only allows comparing statically several databases that must be pre-defined. Thus they can neither detect qualitative leap about students' behavior[2] nor identify the dynamics of behavioral changes since dynamic analysis

[2]A qualitative leap is when there is a change in behavior significant enough to consider that students have changed their strategy.

Table 1: Contrast *ti*-pattern mining.

City datastets		Dbefore18 01/11/18-24/12/18	Dbefore19 01/11/19-24/12/19	Dcov18 17/03/19–10/05/19	Dcov19 17/03/20–10/05/20
2 most frequent		mark mn mark:2014 hwk mn hwk:766	mark mn mark:1252 mail mn mail:684	mark mn mark:1476 mail mn mail:679	**mail mn mail:1782** **mail mn mail mn mail:678**
2 most frequent without duplicates		mark mn hwk:530 hwk mn mark:467	mark mn hwk:418 hwk mn mark:340	hwk mn hwk:651 mark mn hwk:417	**mail mn Ped.It.:495** **hwk mn mail:412**
2 longuests (and most frequent)		mark mn mark [...] [mn mark:49 mark mn mark wk [...] mn mark:37	mark mn mark wk [...] mn mark:36 (s6) mark mn mark wk [...] mn mark:35 (s6)	mark mn mark [...] [mn mark:46 (s6) mark mn mark [...] [mn mark:53 (s5)	**mail mn mail [...]** **[mn mail:38 (s7)** **mail mn mail [...]** **[mn mail:63 (s6)**
2 longuests without duplicates		mark mn hwk mn mark:99 hwk mn mark mn hwk:69	mark mn hwk mn mark:89 mark mn mail mn mark:58	mark mn hwk mn mark:69 mark mn mail mn mark:61	**mail mn Ped.It. mn mail:112** **mail mn hwk mn mail:99**
Proportion of time intervals (%)	sec	7.1%	6.2%	**9.4%**	6.8% ↘
	mn	59.5%	60.1%	56.4%	54.4%
	hour	8.0%	8.6%	**7.2%**	12.3% ↗
	day	5.9%	5.9%	5.6%	5.8%
	wk	12.9%	12.7%	9.0%	9.7%
	mth	6.5%	6.5%	12.4%	11.0%
Proportion of service page usage (% [nb])	upload	1.7% [111]	0.5% [33]	2.0% [120]	0.9% [49]
	collab	5.2% [331]	7.1% [438]	7.4% [448]	7.8% [439]
	mark	65.3% [4193]	44.5% [2715]	**54.1% [3282]**	13.8% [771] ↘
	abs	6.2% [395]	3.8% [234]	**6.4% [388]**	0.8% [46] ↘
	schooL	2.6% [165]	9.5% [580]	2.5% [153]	5.2% [292]
	timeM	0.6% [38]	0.2% [14]	0.4% [24]	0.3% [19]
	hwk	31.9% [2048]	30.6% [1871]	30.0% [1816]	27.2% [1524]
	mail	26.3% [1690]	31.4% [1918]	**31.4% [1902]**	64.4% [3605] ↗
	ressOL	0.0% [1]	5.7% [349]	0.0% [3]	2.6% [144]
	PedaIt	3.1% [201]	27.1% [1655]	**22.5% [1364]**	31.7% [1773] ↗

One dataset per column. Rows 2 to 5 summarize the most representative patterns. Row 6 gives the proportion of time intervals in all discovered patterns, followed by the dataset proportion of page usage per year (row 7). Minimum support = 30.

Table 2: Percentage of *ti*/*s*-patterns contained both in periods *before* & *during* Cov19 during School year.

common	$\mathbf{D_{bef18} - D_{cov18}}$	$\mathbf{D_{bef} - D_{cov}}$
ti-patterns	45.3%	3.4%
s-patterns	90.3%	16.6%

requires observing an event over time. Therefore, we now integrate temporal information into our algorithm. This section aims to answer the research question: ***How to Represent the Dynamic Evolution of the Students' Digital Behavior in an Interpretive Way?***

Proposal: Using Pattern Change Histories.
Change Mining algorithms represent the evolution of models made by different DM algorithms. (Boettcher, 2011) presents an overview of Change Mining methods. In our case, we are interested in analyzing the temporal evolution of patterns' supports, called "pattern histories" (Wang, 2011; Chen et al., 2004). This paper uses *ti*-pattern histories, which have never been done before.

We want to represent the evolution of the most frequent patterns that stay frequent during some months. Such patterns are called "change patterns". To discover them, we compute frequent *s/ti*-patterns per month and select the frequent ones for at least 3 months. Finally, we plot TPH (graphical representations of the support evolution of each change pattern). On these graphs, we represent covid-19 year ($\mathbf{D_{c19}}$ or $\mathbf{D_{d19}}$) and previous one ($\mathbf{D_{c18}}$ or $\mathbf{D_{d18}}$) to compare them. For this experiment, we use a relative minimum support of 20 and the *(TI)-PrefixSpan* algorithms to mine *s/ti*-patterns.

Experiment: Representation of the Dynamic Students' Digital Behavior.
To represent the dynamic evolution of students' digital behavior in an interpretative way, we look at the 10 most frequent *s*-patterns and *ti*-patterns for each month and dataset. Then, we select the ones that are change patterns. Finally, we represent the *s/ti*-pattern change histories (resp. SPH and TPH) where support histories are computed each month.

Results: SPH and TPH Give an Easy-to-Interpret Representation.
Fig. 2 presents TPH appearing in all datasets: $\{\mathbf{D_{c18}},\mathbf{D_{c19}},\mathbf{D_{d18}},\mathbf{D_{d19}}\}$. For each represented TPH, we have the corresponding SPH, with curves that follow the same dynamic (to facilitate the readability, we don't represent them). Moreover, whatever the dataset, frequent time-interval change patterns are generally symmetrical.

For example, the first graph presents the histories of "*homework -mn- mail*", but there also exist

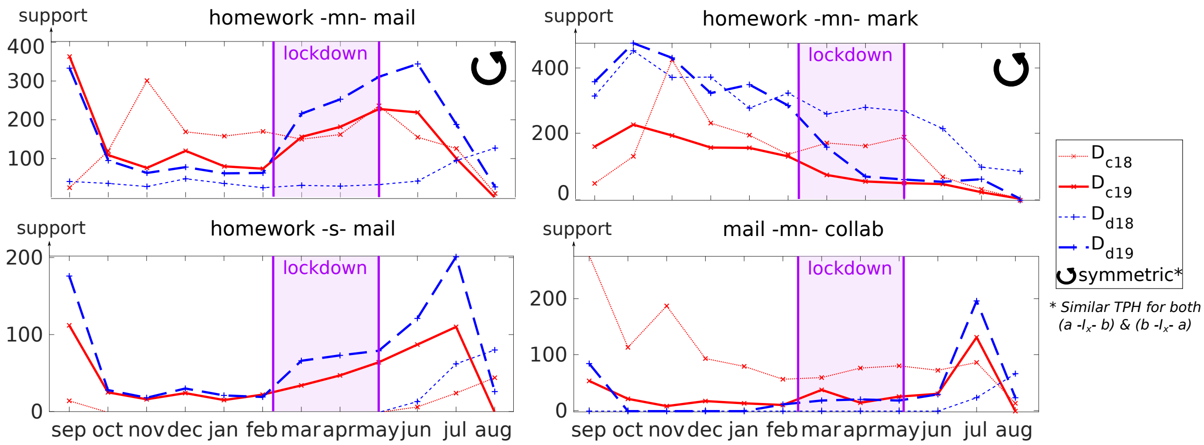

Figure 2: Display of TPH that change over time similarly in the two study locations. Note: corresponding SPH (homework-mail, homework-mark, or mail-collab) are similar (not displayed here).

graphs "*mail -mn- homework*", "*mail-homework*", and "*homework-mail*", with a similar support dynamic.

Those graphs show that students often check the *homework*, *mail*, and *mark* services successively without a specific order, but with a particular pace: a few minutes delay before changing services. In the *departement* area, we even have SPH of size 3, based on the "*mark-homework*" pattern and its symmetric, which follow the same support's trend (not presented here). This shows how concerned the students were about this sequence. Independently of the lockdown, curves show some students' behavioral change between 2018 and 2019. For example, in the city on 11/2018, there was a spike in the use of "*homework -mn- mail*" and "*homework -mn- mark*" patterns that don't appear in 2019. However, these changes are lighter than the ones during the lockdown.

The Comparison of the Four Datasets Suggests That during the Lockdown:
(1) Students brutally stopped using *mark* and *homework* services successively, within the hour. This may be because, during the lockdown, students weren't graded, as analysts told us.

(2) Students tend to use *homework* and *mail* services more often (Fig. 2, left), and they move more and more quickly from one service to another (bottom left).

(3) The pattern "*mail -mn- collab*" (bottom-right) has a spike of activities in July 2020 for both areas. This might correspond to a final collaborative project linked to Cov19. Specific to the *departement* area, 2019 School year, this pattern appears in February, just before the lockdown, with a little support (around 25). This same pattern remains frequent about the city area during the whole 2018 School year, which seems to correspond to a learning method based on projects. This behavior is still present during the 2019 School year, with a lower frequency.

The interpretation of these results is not the focus of this paper. Our goal is to provide this new indicator to analysts. This indicator can also be helpful for teachers' dashboards.

Patterns Specific to Students of the City (*c.f.*, Fig. 3), Give Other Information:
(1) They often check *mail* (or *pedagogical Itinerary*) with *mark* services successively, without temporal regularity. Since the lockdown, this behavior is less frequent than during the previous year. *pedagogical Itinerary - mark* pattern ends up disappearing. Again, this might be because students have no marks during the outbreak.

(2) During the 2019 School year, they often perform the "*mail -mn- pedagIt*" pattern, with this specific order and with a regular gap. This behavior might result from a policy of one or more educational institutions in the city. The lockdown reinforces this behavior since the support of this *ti*-pattern increases from 100 (February) to around 275 (May).

However, we gave some clues to explain students' behavioral changes. These results validate the ability of TPH graphs to visually detect significant students' behavioral changes that correspond to the lockdown period. Some of these behavioral changes are common to students located in the two analyzed areas of France, and others are location-specific. Finally, we saw that the lockdown also impact students' pace. Thus, we can confirm that SPH and TPH can model the dynamic of students' behavioral changes in an interpretative way.

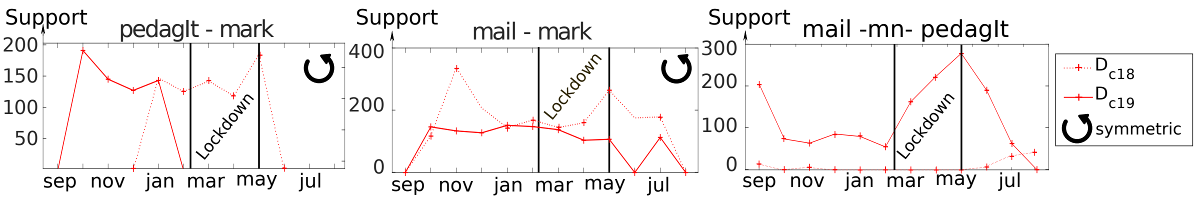

Figure 3: Some TPH and SPH specific to city datasets. Dot lines represent the baseline (2018-19 School year), not impacted by the outbreak. A change history aI_xb is said "Symmetric" if the change history bI_xa is similar.

6 DISCUSSION

The first experiment validates the relevance of data to detect an abrupt change in students' digital behavior caused by Cov19. It shows that students use the mail service a lot more during the lockdown and are more likely to follow an hour space gap between explored services, probably because students use e-services between classes. However, the most classic gap between services remains the second gap. Results highlight that the order in which the students use the different services seems irrelevant: s/ti-patterns are "symmetrical" (Sec.4.1). To conclude this experiment, we show there are more ti/s-pattern changes during the lockdown than in the previous year (baseline). It would be beneficial to evaluate the amount of pattern change between datasets based on contrast-sets specific measures in future works, as proposed in (Magalhães and Azevedo, 2015).

The method of the 2^{nd} experiment allows to dynamically represent students' behavioral changes, visually and understandably, intending to facilitate the analysis of the evolution of students' digital behavior. This method highlights the lockdown impact on students' digital behavior, allowing analysts to interpret it. The results highlight a qualitative leap in students' behavior during the lockdown, even without a-priori knowledge. So, this approach enables analysts to detect important students' e-learning changes. Moreover, since this graphical method allows for overlapping different pattern histories, we can easily compare the dynamic of students' behaviors, which is efficient to perform analysis. Hence, results show that some patterns are more and more followed by students during the lockdown (*e.g.,* "homework-s-mail"), while others stop being followed suddenly by students (*e.g.,* "homework-mn-mark").

7 CONCLUSION AND FUTURE WORKS

This paper proposes two approaches allowing analysts to detect and analyze students' behavioral changes. They have been experimented on the Cov19 case study. Considering generic, objective, and poor data, we succeed in detecting a global trend and visually representing the dynamic of students' behavior in an easy-to-interpret way, thanks to the new approach: TPH. Thus, we answer the global research question and extend the Cov19 state-of-the-art research.

The 1^{st} approach makes a temporal pattern comparison between datasets. It is a static comparison (pre-defined datasets), useful for learning experts to compare a specific period with others. In this 1^{st} study, we explore students' behavioral changes during the lockdown with other temporal periods, thanks to Contrast Mining. We discovered the emergence and vanishing of some temporal patterns and the modification of students' pace, and this, only with e-services data information.

The 2^{nd} approach allows analysts to detect and interpret the dynamic of students' behavioral changes, thanks to TPH. This new method successfully represents clearly, thanks to graph representations, the dynamic of students' behavior. On these graphs, the Cov19 impact was visible even for non-experts. These methods succeed in representing the general trend of a group of students rather than targeting a specific student.

In the future, we will apply these methods to support teachers. We also want to work with analysts to analyze deeper the Cov19 impact on students' digital behavior and to have clues to improve and create dynamical indicators. We will finally search methods that automatically detect significant students' behavior change periods.

REFERENCES

Abbasi, S., yoob, T., Malik, A., and Memon, S. I. (2020). Perceptions of students regarding e-learning during covid-19 at a private medical college. *Pakistan J. of Med. Sci.*

Agrawal, R. and Srikant, R. (1995). Mining sequential patterns. In Yu, P. S. and Chen, A. L. P., editors, *Proc. 11th Int. Conf. on Data Eng.*, Taipei. IEEE Comp. Soc.

Almaiah, M. A., Al-Khasawneh, A., and Althunibat, A. (2020). Exploring the critical challenges and factors influencing the e-learning system usage during COVID-19 pandemic. *Educ. Inf. Technol.*

Anjum, N. and Badugu, S. (2020). A study of different techniques in EDM. In *ICETE*. Springer, online.

Boettcher, M. (2011). Contrast and change mining. *Wiley Int. Reviews: DM & KD*, 1(3):215–230.

Chen, C. F., Yang, S.-H., Falsafi, B., and Moshovos, A. (2004). Accurate and complexity-effective spatial pattern prediction. In *HPCA*, Madrid. IEEE.

Chen, Y.-L., Chiang, M.-C., and Ko, M.-T. (2003). Discovering time-interval sequential patterns in sequence databases. *Expert Systems with Applications*, 25(3):343–354.

Dermy, O. and Brun, A. (2020). Can we take advantage of time-interval pattern mining to model students activity? In *Proc. 13th Int. Conf. on EDM*, online. Int. EDM Soc.

Dong, W., Fan, W., Shi, L., Zhou, C., and Yan, X. A general framework to encode heterogeneous information sources for contextual pattern mining. In *Proc. ACM Int. CIKM*, Maui.

Ebner, M., Schön, S., Braun, C., Ebner, M., Grigoriadis, Y., Haas, M., Leitner, P., and Taraghi, B. (2020). Covid-19 epidemic as e-learning boost? *Future Internet*, 12(6):94.

Favale, T., Soro, F., Trevisan, M., Drago, I., and Mellia, M. (2020). Campus traffic and e-learning during covid-19 pandemic. *Computer Networks*.

Fournier-Viger, P., Lin, J. C.-W., Gomariz, A., Gueniche, T., Soltani, A., Deng, Z., and Lam, H. (2016). The spmf open-source data mining library version 2. In *Joint European conf. on ML and KD in DB*. Springer.

Gao, C., Wang, J., He, Y., and Zhou, L. (2008). Efficient mining of frequent sequence generators. In Huai, J., Chen, R., Hon, H., Liu, Y., Ma, W., Tomkins, A., and Zhang, X., editors, *Proc. of the 17th Int. Conf. on WWW*. ACM.

Gutierrez-Santos, S., Mavrikis, M., and Magoulas, G. (2010). Sequence detection for adaptive feedback generation in an exploratory environment for mathematical generalisation. In *AIMSA*, Varna. Springer.

Han, J., Kamber, M., and Pei, J. (2012). 7 - advanced pattern mining. In *Data Mining*, The Morgan Kaufmann Series in DMS. Boston, 3rd edition.

Ilieva, G., Yankova, T., Klisarova-Belcheva, S., and Ivanova, S. (2021). Effects of covid-19 pandemic on university students' learning. *Information*.

Lau, E., Chai, K. K., Goteng, G. L., and Wijeratne, V. (2021). A neural network modelling and prediction of students' progression in learning: A hybrid pedagogic method. In *CSEDU*.

Magalhães, A. and Azevedo, P. J. (2015). Contrast set mining in temporal databases. *Expert systems*.

Pei, J., Han, J., Mortazavi-Asl, B., Pinto, H., Chen, Q., Dayal, U., and Hsu, M.-C. (2001). Prefixspan: Mining SP efficiently by prefix-projected pattern growth. In *Proc. 17th Data Engineering*, Heidelberg. IEEE.

Photopoulos, P., Stavrakas, I., and Triantis, D. (2021). Post-covid-19 education: A case of technology driven change? In *CSEDU*.

Poon, L. K., Kong, S.-C., Wong, M. Y., and Yau, T. S. (2017). Mining sequential patterns of students' access on learning management system. In *DM & big data*, Fukuoka. Springer.

Radha, R., Mahalakshmi, K., Kumar, V. S., and Saravanakumar, A. (2020). E-learning during lockdown of covid-19 pandemic: A global perspective. *Int. jour. of control & automation*.

Reigal, R. E., Pastrana-Brincones, J. L., González-Ruiz, S. L., Hernández-Mendo, A., Morillo-Baro, J. P., and Morales-Sánchez, V. (2020). Use of data mining to determine usage patterns of an online evaluation platform during the covid-19 pandemic. *Front. in Psycho.*

Wang, J.-D. (2011). A novel approach to compute pattern history for trend analysis. In *8th Int. FSKD*, Shanghai. IEEE.

Wang, R., Ji, W., Liu, M., Wang, X., Weng, J., Deng, S., Gao, S., and Yuan, C. (2018). Review on mining data from multiple data sources. *Pattern Reco. Letters*.