Chen, W. et al. (Eds.) (2017). Proceedings of the 25th International Conference on Computers in Education. New Zealand: Asia-Pacific Society for Computers in Education

Capturing Changes and Variations from Teachers' Time Series Usage Data

Longwei ZHENG[a*], Yuanyuan FENG[b], Rui SHI[a] & Xiaoqing GU[a]
[a]*Faculty of Education, East China Normal University, China*
[b]*School of Foreign Languages, East China Normal University, China*
*lwzheng@dec.ecnu.edu.cn

Abstract: The type of time series data is very common in educational settings, such data set usually provide usage trajectories involved complex contexts and unknown noisy. In this paper, we describe a statistical analysis method to detect a switch point over teacher's comment data set which is collected from a student evaluation system. We find that it is indeed possible to employ a statistical model and Bayesian inference to detect teacher who increased, decreased or no-changed their behavior. We also illustrated the potential of the application of change detection to conduct data exploration, which can be relative to understand teacher technology adoption and usage transition.

Keywords: change detection, teacher's technology adoption, statistical modeling, MCMC

1. Introduction

In educational settings, time series data is often applied for possible dropout, outcome predicting, learning pathway discovering, and engagement detecting (Arnold & Pistilli, 2012; Blikstein, 2011; Clow, 2013; Park et al., 2017). With this type of data set, statistical change detection techniques can be utilized to investigate segmentation or human behaviors, such as detecting special change time on individual's clickstream which is a widely studied topic in web interacting research (Carlin, 1992; Park et al., 2017). As in most cases of change detection study, the main contribution of change detection is to provide a feature to group users into similar transition groups or search a signal to partition a time series into two distinct observations (Eckley, 2011).

The purpose of this study is to investigate teacher's time series usage of a student evaluating system by employing a simple change detection technique which is introduced by Davidson-Pilon (2015). The comments which created in a student evaluating system will be computationally detected significant switch point of number of comments. We hope to identify teachers who increase, decrease or show no change in comment dataset, so that we can get more insights of teacher's technology adoption and usage by discovering the relations to other aspects of technology usage between before switch point and after.

2. Methodology

2.1 Data

To explore this issue, we collected teachers' comment data from a student evaluating system which supports teacher to code the students' performance by clicking on a tablet application, and then records the code data into system to automatically generate a weekly review for each student. The users of this platform are teachers at 8 primary schools of Shanghai. We obtained 991,979 comment records for 2,224 students in 59 classes which has almost same class size over a 4-month semester. Each comment records the praise or criticism that a teacher evaluates a student at a certain time. Typically, a teacher would click a button that might indicate 'You have all the right answers in arithmetic quizzes today! Good job!' if he or she wants to commend a student's performance in class. As the teachers' contexts (such as grade, subject, students' contexts, etc.) are various, the comment usage and comment content can vary from one group of teachers to another.

2.2 Modeling of Switch Point

A Poisson random variable is appropriate model for the type of count data of teacher comment records. We assumed that time is a discrete with *T* discrete days and *t = 1, 2, …, T* being as index running from the first day to the last day of period in which an individual teacher *i* is active in the system. Therefore, day t's comment count can be denoted by C_{it},

$$C_{it} \sim Poisson(\lambda_i) \tag{1}$$

The expected value of the Poisson distribution is equal to its parameter λ_i, we assume that the λ_i on day τ_i which is independent from a uniform distribution, the parameter λ_i suddenly jumps to a higher value or declines rapidly to a lower value. Such a sudden transition can be denoted as two λ_i parameters, one is assigned for the period before τ_i, and the other one is assigned for the rest period. In addition, the exponential distribution is employed to model λ_i, so the parameter of exponential distribution α_i, which reflects the prior belief for λ_i, and is included in the switch point model:

$$\begin{cases} \lambda_{i1} \sim Exp(\alpha_i) & if\ t < \tau_i \\ \lambda_{i2} \sim Exp(\alpha_i) & if\ t \geq \tau_i \end{cases} \tag{2}$$

Thus the expected value of the comment count can be identified as:

$$\frac{1}{N}\sum_{t=0}^{N} C_{it} \approx E[\lambda_i|\alpha_i] = \frac{1}{\alpha_i} \tag{3}$$

Figure 1 shows a teacher's daily comment count data in the completed observing period. We can find a significant decrease from early half period to latter half period, so we assume that this teacher changed his or her behavior on day τ_i by switching λ_{i1}, the parameter of comment count C_{it}, to a higher value λ_{i2}.

2.3 Bayesian Inference and Markov Chain Monte Carlo

In order to estimate parameters of switch point model, we employed Bayesian inference, that is a method of statistical inference in which Bayes' theorem is used to update the probability for a hypothesis as more evidence or information becomes available (Box & Tiao, 2011). Markov chain Monte Carlo (MCMC) is a general method for simulation of stochastic processes having probability densities known up to constant of proportionality. MCMC is often applied to Bayesian inference that measure and update the believability of events after considering new evidence (Green, 1995). The algorithm can be expressed as an approximation solution to the posterior: the way that the algorithm moves in the general direction towards the regions where posterior distribution exist, and collect samples based on the the position's adherence to the data and prior distributions. Once it reaches the posterior distribution, the samples can be collected as they likely belong to the posterior distribution (Davidson-Pilon, 2015).

In this study, we utilized Bayesian inference to get the posterior distributions of three parameters (λ_1, λ_2, τ) of change detection model by performing MCMC. Then we get the expectation from each parameter distribution and assign to every teacher. As shown in Figure 1, a teacher is detected that he or she has an obviously higher count of comments after day 33, and two λs of two distinct periods are also estimated with their expectations.

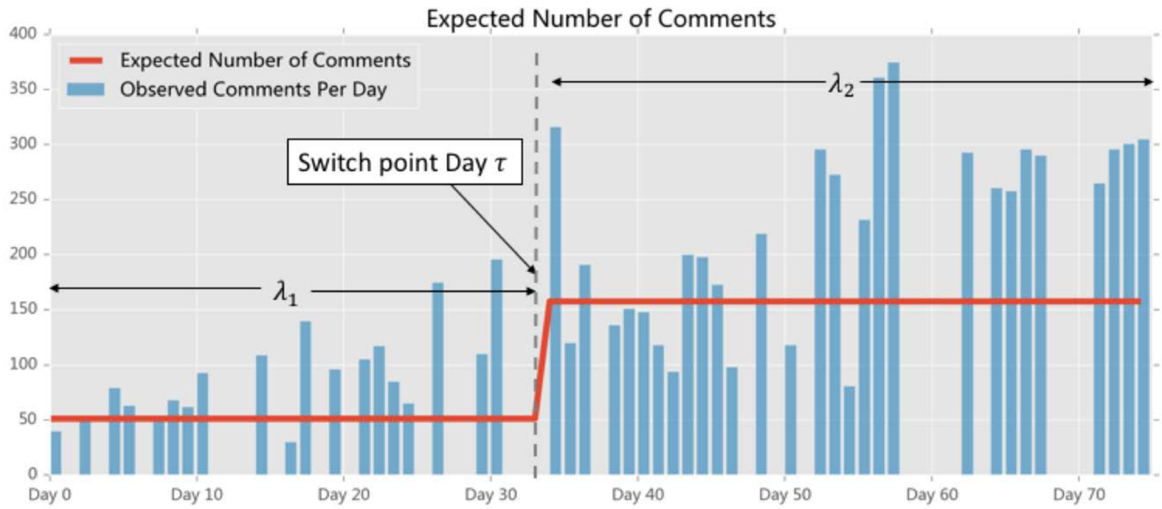

Figure 1. Modeling of switch point and parameter estimating result.

3. Result

3.1 Change Detection

We omit the data of teachers who has a very limited usage (count of active days < 20). We also restricted the switch day to be ranged from the 10th day from top to the 10th day from bottom, since switch point detection at the beginning or end of tends to be unreliable due to small sample sizes or be affected by semester timelines, and also not so meaningful in terms of interpreting (Park et al., 2017). Therefore, the usage which is detected switched at this two period will be considered as no-change.

Finally, we applied the change detection methodology to comment data performed by 145 teachers. The numbers of teachers detected as belonging to increase group and decrease group are shown in Table 1. The result shows that more teachers (42.07%) increased comments in the observing period. There is also a significant difference between two group's distributions. If we order the teachers by expected number of the first four weeks (λ), we can find that the teacher who has intensive usage has more possibility to increase comment rate, and the teacher who has more moderate or limited usage tend to decrease comment rate. Therefore, this change detection result, to a certain degree, can be considered as a feature when predicting teacher's technology adoption in early days.

Table 1: Change detection result summary.

<table>
<tr><th></th><th colspan="2">Number of Increased Teacher</th><th colspan="2">Number of Decreased Teacher</th><th colspan="2">Number of No-change Teacher</th><th colspan="2">Number of Analyzed Teacher</th></tr>
<tr><td rowspan="2">All</td><td colspan="2">61</td><td colspan="2">46</td><td colspan="2">38</td><td colspan="2" rowspan="2">145</td></tr>
<tr><td colspan="2">42.07%</td><td colspan="2">31.72%</td><td colspan="2">26.21%</td></tr>
<tr><td>Top Half by ordering λ</td><td>41</td><td>67.21%</td><td>15</td><td>32.61%</td><td>17</td><td>44.74%</td><td>73</td><td>50.34%</td></tr>
<tr><td>Bottom Half by ordering λ</td><td>20</td><td>32.79%</td><td>31</td><td>67.39%</td><td>21</td><td>55.26%</td><td>72</td><td>49.66%</td></tr>
</table>

Figure 2 provides a day-by-day summary for the number of detected teacher changes per day. There is an obvious pattern in this visualization. The bar plot shows that there are many changes detected (more than half teachers from increased group increased their comments) from 45th day to 60th day, which is the followed two weeks after midterm. This, to a certain degree, can agree with the intuition that some teachers adopt this system to help them evaluate student after a half-semester probationary using period.

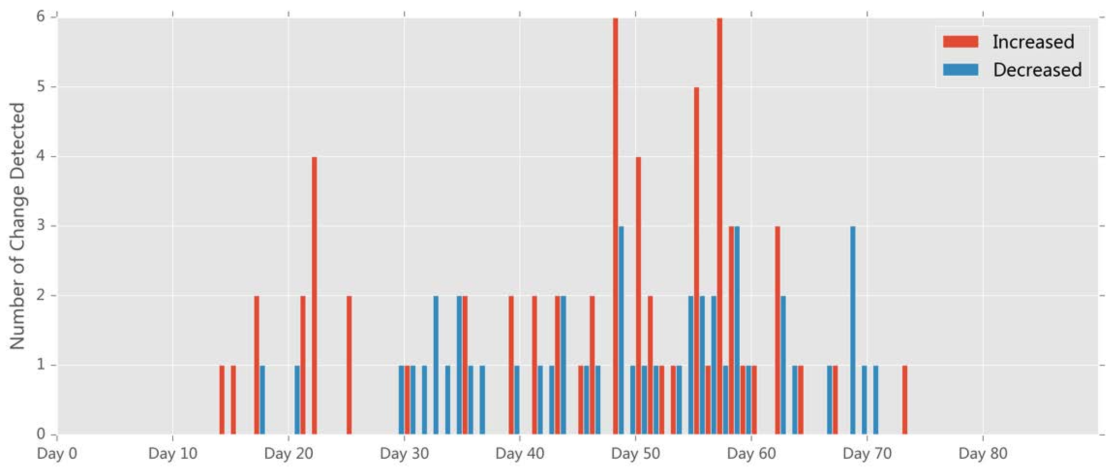

Figure 2. Number of teachers who increased or decreased in each day.

3.2 Discourse Variation

In this section, we applied switch point detection result to partition time series to two distinct data sets. Discourse variation analysis will be conducted by comparing the data set before switch point and the after.

Firstly, we utilized Shannon Entropy to calculate variety of the comments that was given by teacher. Shannon entropy is a mathematical construct to model variety from information theory (Shannon, 2001).

As shown in Table 2, there is significant difference of comment variety transition between two switched point detected groups, it makes sense that teachers from increased group have more chance to give more comment for more students. However, if we partition each group to two sub-groups by median number of types of comments that are given by every teacher (N_{ct}, the median of population is 11), we can find two distinct different results. Again it makes sense that the teachers who are from increased group and prefer to use specific words have more probability to increase variety (the percentage of the Entropy increased is 66.67%). We can also see the teachers who are from comment decreased group and prefer diverse comments have more probability to increase entropy (the percentage of the Entropy increased is 70.00%). It indicates that these teachers might tend to make specific comments to specific students and avoid to give high-coverage comments after their switch points, such behavior change will lead to a comment decrease detection.

Table 2: The change summary of comments variety.

	All Switch Point Detected					$N_{ct} \le 11$					$N_{ct} > 11$				
	Total	Entropy Increased		Entropy Decreased		Total	Entropy Increased		Entropy Decreased		Total	Entropy Increased		Entropy Decreased	
Comment Increased Group	61	37	60.66%	24	39.34%	48	32	66.67%	16	33.33%	13	5	38.46%	8	61.54%
Comment Decreased Group	46	22	47.83%	24	52.17%	26	11	42.31%	15	57.69%	20	14	70.00%	6	30.00%

Secondly, we compared the sentiment transition between two groups by measuring the comment's sentiment. The sentiment measuring process consists of following steps: 1. Segment each comment to a set of words tagged with part-of-speech (such as noun, name, verb and etc.); 2. Build sentiment corpuses by labeling positive or negative for segmentation result; 3. Assign degree weights to adverbs by employing a dictionary of degree of Chinese adverbs (Dong & Dong, 1999), such as the superlative word ('most', 'best', 'extremely'), comparative words ('more', 'a bit', 'much'), negative words ('not', 'no', 'but'); 4. Compute the sentiment score of each comment by multiply degree weight and set sign according to its sentiment corpus.

The sentiment analysis result shows a significant difference of quantity between the positive comments (958,442) and the negative comments (33,537). It illustrates that teacher are inclined to make more encouraging evaluations or positive praises to students. But we also find that most teachers began to adopt more negative comments over time. Thus we select the teachers who has a sentiment score difference which was statistically significant between two periods.

As shown in Table3, more teachers decreased sentiment significantly after switch point, and there is no significant difference between the comment increased group and comment decreased group. However, if each group is partitioned to two sub-groups by median number of types of comments that are given by every teacher, again we can find that the teachers who prefer diverse comments from either group have more probability to increase comment sentiment. As mentioned in entropy analysis above, this is to be expected with less negative comment behavior since these teachers' decrease for less-frequently-used comment in latter period. On the other hand, the teachers who prefer same comments are more likely to adopt more negative words after switch point.

Table 3: The change summary of comment sentiment

<table>
<tr><td rowspan="2"></td><td colspan="5">All Switch Point Detected</td><td colspan="5">$N_{ct} \leq 11$</td><td colspan="5">$N_{ct} > 11$</td></tr>
<tr><td>Total</td><td colspan="2">Sentiment Increased</td><td colspan="2">Sentiment Decreased</td><td>Total</td><td colspan="2">Sentiment Increased</td><td colspan="2">Sentiment Decreased</td><td>Total</td><td colspan="2">Sentiment Increased</td><td colspan="2">Sentiment Decreased</td></tr>
<tr><td>Comment Increased Group</td><td>49</td><td>23</td><td>43.06%</td><td>26</td><td>46.94%</td><td>36</td><td>14</td><td>38.89%</td><td>22</td><td>61.11%</td><td>13</td><td>9</td><td>69.23%</td><td>4</td><td>30.77%</td></tr>
<tr><td>Comment Decreased Group</td><td>33</td><td>17</td><td>51.52%</td><td>16</td><td>49.48%</td><td>18</td><td>7</td><td>38.89%</td><td>11</td><td>61.11%</td><td>15</td><td>10</td><td>66.66%</td><td>5</td><td>33.34%</td></tr>
<tr><td rowspan="2">Total</td><td rowspan="2">82</td><td colspan="2">40</td><td colspan="2">42</td><td rowspan="2">54</td><td colspan="2">21</td><td colspan="2">33</td><td rowspan="2">28</td><td colspan="2">19</td><td colspan="2">9</td></tr>
<tr><td colspan="2">48.78%</td><td colspan="2">51.22%</td><td colspan="2">38.89%</td><td colspan="2">61.11%</td><td colspan="2">67.86%</td><td colspan="2">32.14%</td></tr>
</table>

4. Discussion

This paper introduced an approach for change detections based on teachers' comment records from a student evaluation system. It is indeed possible to employ a statistical switch point model and Bayesian inference to detect teacher who increased or decreased their action rate, even the count data of user action generally pose its noisy nature. We also illustrated the potential of the application of change detection to distill features, which can be relative to teacher technology adoption. In our study, this approach to exploring teacher technology use behaviors from searching transition signals have been very revealing. They help us to understand that the teacher who has intensive usage in early days has more possibility to adopt technology in future. Further, modeling technology use in terms of detected change result combined with other dimensions, such as variety and sentiment, may be a useful diffuse framework for capturing the usage pattern variation and characterizing how usage transition happens. In our study, the feature of comments diversity may be helpful to explore why the changes of variety and sentiment appear in different groups. In the future, our work involves the discoveries of more unknown behaviors within technology usage of teacher by multi change point detections.

References

Arnold, K. E., & Pistilli, M. D. (2012). Course signals at Purdue: Using learning analytics to increase student success. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 267-270). ACM.

Blikstein, P. (2011). Using learning analytics to assess students' behavior in open-ended programming tasks. In *Proceedings of the 1st international conference on learning analytics and knowledge* (pp. 110-116). ACM.

Box, G. E., & Tiao, G. C. (2011). *Bayesian inference in statistical analysis* (Vol. 40). John Wiley & Sons.

Carlin, B. P., Gelfand, A. E., & Smith, A. F. (1992). Hierarchical Bayesian analysis of changepoint problems. *Applied statistics*, 389-405.

Clow, D. (2013). MOOCs and the funnel of participation. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 185-189). ACM.

Davidson-Pilon, C. (2015). *Bayesian Methods for Hackers: Probabilistic Programming and Bayesian Inference*. Addison-Wesley Professional.

Dong, Z., & Dong, Q. (1999). *Introduction to HowNet.* Retrieved from http://keenage.com/Theory%20and%20practice%20of%20HowNet/04.pdf.

Eckley, I. A., Fearnhead, P., & Killick, R. (2011). Analysis of changepoint models. *Bayesian Time Series Models*, 205-224.

Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 711-732.

Park, J., Denaro, K., Rodriguez, F., Smyth, P., & Warschauer, M. (2017, March). Detecting changes in student behavior from clickstream data. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 21-30). ACM.

Shannon, C. E. (2001). A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, *5*(1), 3-55.