Ogata, H. et al. (Eds.) (2015). Proceedings of the 23rd International Conference on Computers in Education. China: Asia-Pacific Society for Computers in Education

Analysis on Learning Efficiency in the Context of Mobile e-Learning

Shengyuan SHU & Xiaoming JU*
Software Engineering Institute, East China Normal University, China
*xmju@sei.ecnu.edu.cn

Abstract: Mobile internet is now rapidly developing, it expands traditional Internet and benefits people's life in every aspect, including e-learning. With mobile e-learning, a learner can get online courses not only stick to a PC but could also be on a toilet, in a bed, or everywhere. This paper abstracted those ways out into 3 basic situations, combined with learning time periods to learning situations, then proposed a model to illustrate how the learning situations affects learning performance and introduced a method of iterative regression to evaluate the learning efficiencies of each situation with the learning data of 200 subject learners recorded in the Sophia Learning Management System (SLMS). The results demonstrate that different learning situations have unequal learning efficiencies, learning with a PC has higher learning efficiency, and sitting learning with a mobile also has a little higher rate than lying. Different courses have different learning efficiencies. It's helpful to compare learning efficiencies among courses so that learners could get recommendations of scheduling efficient learning.

Keywords: e-Learning, Learning Analytics, Learning Efficiency, Regression

1. Introduction

According to StatCounter Global Stats (2014), by the time Aug 2014, mobile internet usage increased to 35.3% (mobile 28.5%, tablet 6.8%) from 21.9% in Aug 2013 while desktop devices access decreased to 64.6%, which means that mobile internet access makes up one-third of the whole internet access. Mobile internet is now rapidly developing, providing more services for people, including e-learning. One of various benefits of e-learning is that the online educational data can be gathered and analyzed with analysis techniques. This process is called Learning Analytics (LA). Learning Analytics has been defined as “the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs” (U. S. Department of Education, Office of Educational Technology, 2012). The purpose of LA is to statistic and analyze the learner profiles and behavioral data and to learn the patterns for improving learning efficiency. Techniques like social network analysis or predictive modelling are rapidly used for this. With the result generated by the techniques, conclusions for improving learning could be made.

Compared to traditional ways, mobile internet grant e-learning more flexibility and diversity, learners can get courses at any time in any occasion even sitting on toilet or lying in bed just with a mobile device. Diverse situation generates diverse data, this paper grouped the ways to 3 basic situations: learning through a PC, lying learning through a mobile device and sitting learning through a mobile device. combining the 3 basic situations with when does the learner learn, the authors proposed a model to illustrate the relationship between those factors and learning efficiency, analyzed how these factors affect learning performance and evaluated the learning efficiencies with a method of iterative regression, and compared different courses to find the variance in learning efficiencies.

The rest of this paper is organized as follows. In section 2, the related works are described. In section 3, the learning situations partition is presented. In section 4, the model and the method for evaluating learning efficiencies are presented. In section 5, the final conclusion and the future work are presented.

2. Related Works

To learn the success factors of e-learning, there has been lots of researches accomplished, here shows some of them related to this paper. Hassan M. Selim (2007) proposed the confirmatory factor model for calculating the criticality level of e-learning critical success factors (CSFs). The 53 e-learning critical success factors were grouped into 4 categories, i.e. instructor, student, information technology, and university support, e.g. “The instructor is enthusiastic about teaching” (in category “instructor”) and “The student enjoy using personal computers” (in category “student”). They can be further categorized into 8 kinds, each includes several CSFs, and its level of criticality was measured by its validity coefficient. Wannasiri Bhuasiri et al. (2012) also revealed 6 dimensions and 20 critical success factors that affects learning performance for e-learning systems, recommended implementing e-learning systems. Rabeb Mbarek and Dr. Ferid Zaddem (2013) extended an e-learning effectiveness model by adding the factor social presence to other studied factors like computer self-efficacy, perceived usefulness, perceived ease of use, and interaction between trainer and trainees, the model is to identify the influence of those factors to e-learning effectiveness. Haiping Zhu et al. (2014) analyzed learning behaviors and nonintellectual factors such as emotion, submit time of assignments, login time, and learning style, to find out the influence to learning performance. Those researches learned e-learning success factors, further, in the context of mobile e-learning, there could be more unstudied influences to be found out, and that’s what this paper presents.

3. Data Extracting and Pretreatment

3.1 Learning Situation Distinguish

To examine how learning situation affects learning efficiency, it's necessary to define the learning situations. According to what a learner learns through, the 2 situations learning with PC and learning with mobile can be defined. Learning with a mobile is a flexible way to learn, but also can be summarized to 2 kinds: lying and sitting, combined with the PC occasion. So there are now 3 basic situations: lying learning through mobile, sitting through mobile, and learning through PC. It's easy to distinguish learning through PC and learning through mobile because the former is to visit the Learning Managing System (LMS) site and the latter is to use the LMS Application in the mobile. While to distinguish the lying situation and sitting situation through a mobile, an accelerometer which a smart phone should have one could be used for body position and posture sensing (Foerster, Smeja, & Fahrenberg 1999). Accelerometers calculate the direction of the gravity, so that the orientation of the mobile phone could be determined. According to this, a posture recognition program could be built in the LMS application to recognize the learner's posture.

3.2 Posture Recognition Program

When the accelerometer identifies change on acceleration (forces including gravity are essentially accelerations), the program will receive a sensor event including 3 directions of axis of acceleration, i.e. x, y, and z, as shown in Figure 1.

Figure 1. Axis of Accelerometer.

The values of x, y, z are represented by the values of components of the gravity in opposite axis directions. E.g. while the screen surface is facing upward, it will be *(0, 0, 10)*. Thus the orientation of a mobile device can be represented by *(x, y, z)*.

3.2.1 Sitting Learning With a Mobile

While a learner is sitting or standing using a mobile device, it could have one of the 3 orientations:

- The mobile device lies on a plane with the screen orients upward, the vector should be *(0, 0, 10)*.
- The learner holds the mobile in hand, screen of the mobile orients the horizontal direction, the vector should be *(0, 10, 0)*.
- The learner holds the mobile in hand, screen of the mobile orients oblique upward, the vector should be between *(0, 0, 10)* and *(0, 10, 0)*.

Summarizing the above situations, it can conclude that when the learner is sitting or standing using a mobile, the vector *(x, y, z)* should met the following conditions:

$$\begin{cases} \exists e(x = 0 + e) \\ \exists e(y \geq 0 + e) \\ \exists e(z \geq 0 + e) \\ \exists e(\sqrt{y^2 + z^2} = 10 + e) \\ e \in [-E, E] \end{cases}$$

The *E* represents acceptable maximum error for that a mobile device should always tilt a bit, e.g. while *E* equals to 2.93 (i.e. $\left(1 - \frac{\sqrt{2}}{2}\right) \times 10$), the device could at most tilt 45 degrees from aligned situations mentioned before.

While the device is accelerating that is to say the resultant acceleration is significantly greater than 10, the above conditions will never met.

3.2.2 Lying Learning With a Mobile

While a learner is lying using a mobile device, it could have one of the below orientations:

- The flat lying learner holds the mobile over his face, the screen orients downward, the vector should be *(0, 0, −10)*.
- The side lying learner holds the mobile to the left of his body, the left side of the device orients downward, the vector should be *(10, 0, 0)*.
- The side lying learner holds the mobile to the right of his body, the right side of the device orients downward, the vector should be *(−10, 0, 0)*.
- The lying learner holds the mobile above his face, the screen orients oblique downward, the vector could be between *(0, 0, −10)* and *(10, 0, 0)*, or between *(0, 0, −10)* and *(−10, 0, 0)*.

It can be concluded that when the learner is lying using a mobile, the vector *(x, y, z)* should met the conditions:

$$\begin{cases} \exists e(y = 0 + e) \\ \exists e(z \leq 0 + e) \\ \exists e(\sqrt{x^2 + z^2} = 10 + e) \\ e \in [-E, E] \end{cases}$$

The *E* represents acceptable maximum error. In the SLMS it was set as 2 for precision.

3.3 Partition of Time Periods

200 subject learners were asked to learn *Computer Architecture* through SLMS, and their learning data has been recorded. Figure 2 shows 40 of all the recorded learning time periods distribution.

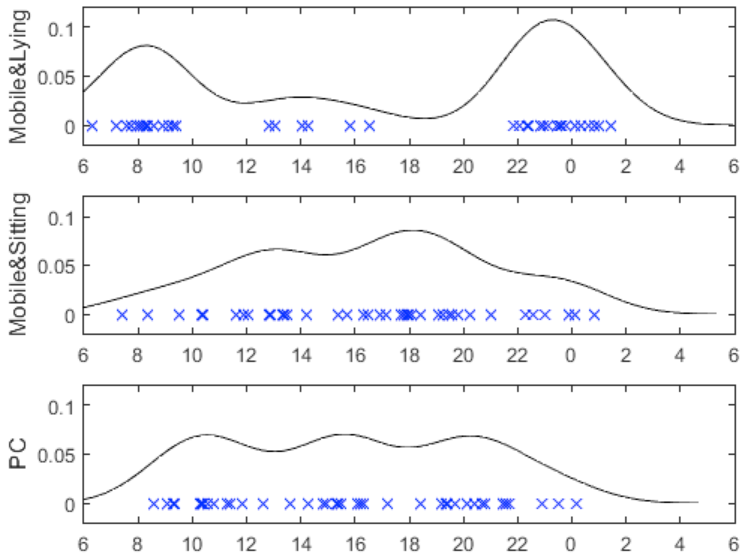

Figure 2. Learning Time Periods Distribution.

Each spot represents a period of learning, and the curves represent probability densities of each basic situation. The X-axis indicates when the period of learning happens, it was measured by the average time of starting learning and stopping learning. The Y-axis indicates probability density for the curves.

As the figure shows, learners' learning time periods dispersed throughout all the time of a day. To classify them, a cluster analysis should be performed, with k-means algorithm (MacQueen, 1967) which was applied and studied in diverse disciplines. Although the k-means algorithm was presented for nearly 50 years, it's still considered as one of the most popular clustering algorithms (Anil, 2010).

K-means algorithm's process is to minimize the sum of distances of each spot to the center spot of the class which the spot belongs to. It can be presented by the following expression:

$$J = \sum_{n=1}^{N}\sum_{k=1}^{K} r_{nk}\left\|x_n - \mu_k\right\|^2$$

While spot n is in class k, $r_{nk} = 1$, otherwise $r_{nk} = 0$.

In this case, the distance represents the time difference. While time of a day is circular, e.g. the time difference between 23:00 and 1:00 is 2 hours but not 22 hours, so the expression can be adjusted as follows:

$$J = \sum_{n=1}^{N}\sum_{k=1}^{K} r_{nk} \times \min^2\left(\left\|x_n - \mu_k\right\|, 24 - \left\|x_n - \mu_k\right\|\right)$$

The probability density curve shown in Figure 2 implies that those periods of learning could be classified into 3, 2, and 3 clusters respectively for each basic situation. Figure 3 shows the k-means clustering results.

The different styles of the marks represents different clusters, as Figure 3 indicates, the 3 basic situations can be divided into the 8 kinds of situations shown in Table 1.

With the clustered samples as the training data, further collected data could be classified into one of those 8 situations with k-Nearest Neighbor algorithm (Fix, & Hodges, 1951) which is considered as one of the top 10 algorithms in data mining (Wu, et al. 2008).

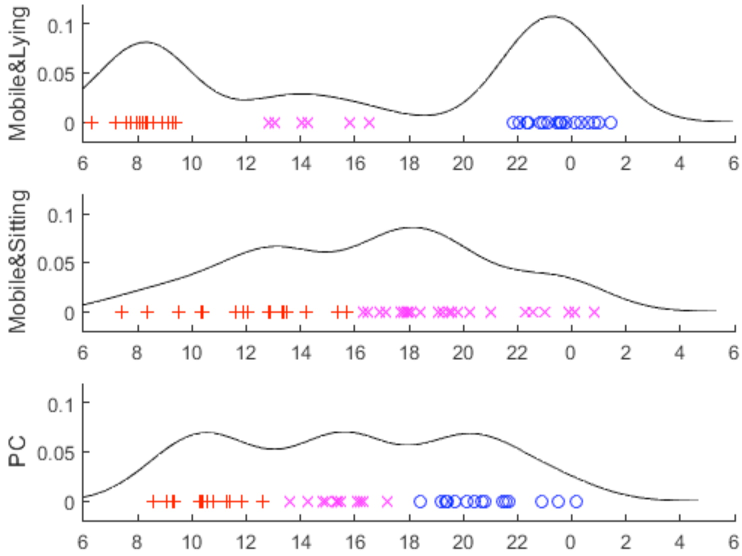

Figure 3. Learning Time Periods Clustering.

Table 1: Situations.

Situation	Indicator
mobile, lying, forenoon	T_1
mobile, lying, afternoon	T_2
mobile, lying, evening	T_3
mobile, sitting, daytime	T_4
mobile, sitting, evening	T_5
PC, forenoon	T_6
PC, afternoon	T_7
PC, evening	T_8

4. Analysis on Learning Efficiency

4.1 Relationship Between Learning Efficiency and Score

The learning efficiencies of learning in different situations of the 8 above are unequal. Learning in some situations may be faster while some other may be slower. To evaluate the efficiencies of situations it's practicable to find clues from how they affects learners score of this course, it can be abstracted to the following expressions:

$$T_e = \sum_{i=1}^{n} b_i T_i \middle| n = 8$$

$$S = f(T_e) \middle| S \in \{A^+, A, B, C, D\}$$

For each subjective learner of the 200 taken part in the learning of *Computer Architecture*, the experiment recorded his/her learning time duration happened in each situation respectively, and his/her score of this course. In this case, $n = 8$ for there are 8 situations. The score recorded is indicated by one of the 5 ranks $\{A+, A, B, C, D\}$.

Accumulate each learner's daily learning time duration in ith situation as T_i, each situation has a learning efficiency, which is represented by b_i. The bigger it is, the more efficient learning in that situation is. Multiply the efficiency b_i by the time duration learning in corresponding situation T_i, there comes the equivalent learning time duration in *i*th situation, then sum up these equivalent time durations, here comes the total equivalent learning time duration T_e which affects the score a learner can get. The total equivalent learning time duration of a learner has a non-linear relationship *f()* to the score he could perform.

4.2 Method for Evaluating the Learning Efficiencies

To figure out the learning efficiencies b_i and the mapping relation *f()* between the score *S* and the equivalent learning time duration T_e, the following method can be used, it is iterative regression.

Before the iterations start, assume that $b_{i(0)}|i \in [1,n] = 1$, that is where the iteration starts. So in the first iteration, there comes the expression $T_{e(1)} = \sum_{i=1}^{n} b_{i(0)} T_i = \sum_{i=1}^{n} T_i$, in which $T_{e(1)}$ is the equivalent learning time duration based on the assumption that all the situations have a same learning efficiency, it updates in further iterations. Gather all students' equivalent learning time durations of a same score and calculate the mean equivalent learning time duration $t_{mj(1)} = mean(\{T_{e(1)}|EQT(j,T_{e(1)})\})|j \in \{A+,A,B,C,D\}$ where $EQT(j,T_e)$ means that a *j* scored learner has an equivalent learning time duration of T_e. Then there comes the inverse function $f^{-1}{}_{(1)}(j) = t_{mj(1)}|j \in \{A+,A,B,C,D\}$, in where the *j* is a discrete enumeration which represents the score of a learner. With the function $f^{-1}{}_{(1)}()$ there comes a possibly more accurate total equivalent learning time duration $t_{mj(1)}|j \in \{A+,A,B,C,D\}$ of a *j* scored learner. Then regress the learning time durations in each situation $T_i|i \in [1,n]$ and the total equivalent learning time durations $t_{mj}(1)|j \in \{A+,A,B,C,D\}$ for the linear relationship between them, then there comes more accurate learning efficiencies $b_{i(1)}|i \in [1,n]$ of each situation. With $b_{i(1)}$, a second iteration can be performed. But before that, there is one thing should be clear that the factor truly affects is the ratio among b_is but not the exact values of b_is. Iteration by iteration b_is may get smaller and smaller, so it is necessary to fix the total value of b_is to a certain value (can be *n*), holding the ratio among each b_is. Here did $b_i := b_i \times \frac{n}{\sum_{I=0}^{n} b_I}$ so that the sum of b_is can be fixed to *n*.

Formally, in *p*th iteration:

$$
\begin{aligned}
T_{e(p)} &= \sum_{i=1}^{n} b_{i(p-1)} T_i \big| i \in [1,n] \\
t_{mj(p)} &= mean\left(\left\{T_{e(p)} \big| EQT\left(j, T_{e(p)}\right)\right\}\right) \\
T_{m(p)} &= t_{mj(p)} \big| SC(j) \\
b_{i(p)} &= regress\left(T_{m(p)}, \left\{T_i \big| i \in [1,n]\right\}\right) \\
b_{i(p)} &:= b_{i(p)} \times \frac{n}{\sum_{I=0}^{n} b_{I(p)}} \Big| i \in [1,n]
\end{aligned}
$$

EQT(j, T_e): a j scored learner has an equivalent learning time duration of T_e.

SC(j): the learner has a score of *j*.

Once the b_i converges to a certain value, it can be decided as the learning efficiency of *i*th situation.

4.3 Calculated Results

After running the algorithm for 10 iterations with the data of 200 learners, the values of b_i in each iteration comes in Figure 4.

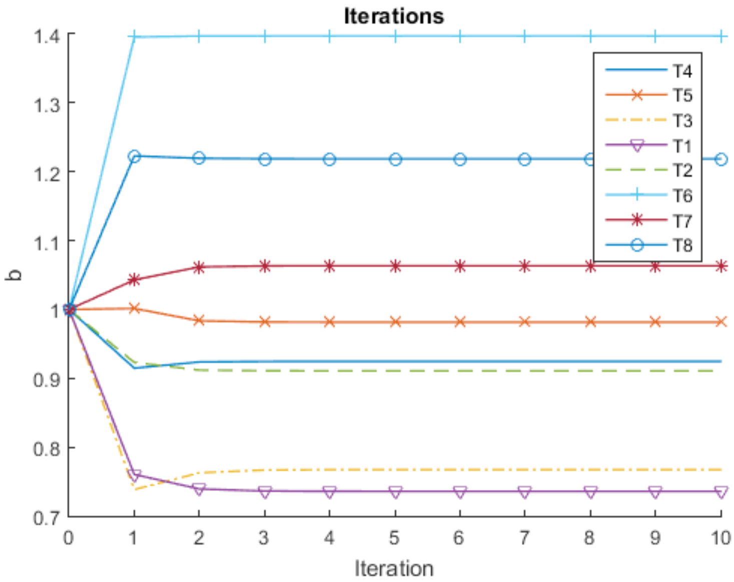

Figure 4. *b*s in Iterations.

The 8 lines represent b_is, the X-axis indicates iterations, and the Y-axis indicates the value of them. As Figure 4 shows, in 5th iteration, the learning efficiencies $b_{i(5)}$ have converge to stable values as follows:

$$\begin{cases} b_1 = 0.7359 \\ b_2 = 0.9108 \\ b_3 = 0.7674 \\ b_4 = 0.9246 \\ b_5 = 0.9816 \\ b_6 = 1.3974 \\ b_7 = 1.0634 \\ b_8 = 1.2189 \end{cases}$$

This result reveals learning efficiencies in each situation, 6th situation i.e. learning through PC in a morning is the most efficient learning situation, followed by 8th situation i.e. learning through PC at night. Generally, learning through PC has a higher efficiency than through mobile, for mobile situation, sitting has a little higher rate than lying.

The standard deviations of equivalent learning time durations of learners of each same score are shown in Figure 5.

The X-axis indicates iterations, and the Y-axis indicates standard deviations. As the figure shows, with the iteration continues, standard deviations of equivalent learning time durations are getting smaller, which means these calculated equivalent learning time durations are gathering up, they are more accurate than that in last iteration.

With the stabilized b_is, calculated learners' equivalent learning time durations classified by the score are shown in Figure 6.

Each dot indicates each learner, its X value is his equivalent learning time duration, and each circlet indicates the mean of those equivalent learning time durations classified by score.

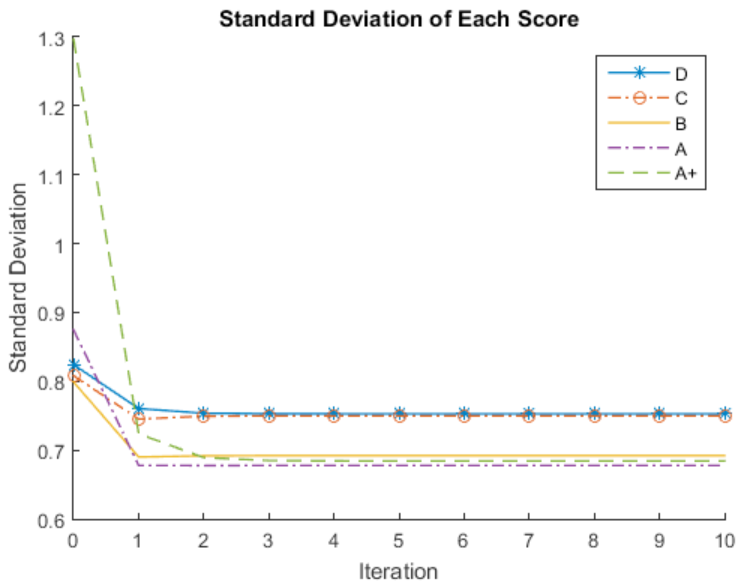

Figure 5. Standard Deviations of Equivalent Learning Time Durations.

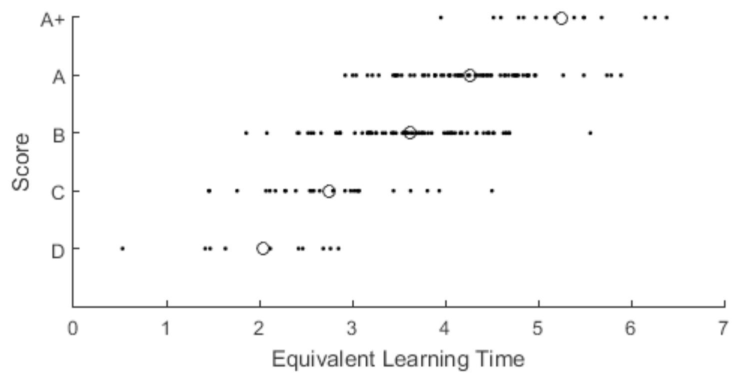

Figure 6. Equivalent Learning Time Durations Distribution.

4.4 *Difference Between Courses*

Former result reveals each situation's learning efficiency about the course *Computer Architecture*, but different course has different efficiency in each situation, following Table 2 shows the result of the course *Business English* gained using the above method with records of 200 subject learners, compared to *Computer Architecture*:

Table 2: Learning Efficiencies of Different Courses.

Situation	Computer Architecture	Business English
mobile, lying, forenoon	0.7359	0.8813
mobile, lying, afternoon	0.9108	0.8990
mobile, lying, evening	0.7674	0.8888
mobile, sitting, daytime	0.9246	0.9744
mobile, sitting, evening	0.9816	1.0312
PC, forenoon	1.3974	1.1898
PC, afternoon	1.0634	1.1233
PC, evening	1.2189	1.0123

Business English also has a higher learning efficiency through PC, but slighter than *Computer Architecture*. That makes *Business English*'s learning efficiency through PC relatively lower, mobile relatively higher. Due to the learning efficiencies, it's possible to recommend that in situation T_1, T_3, T_4, T_5, T_7, it's better to learn *Business English*; in situation T_2, T_6, T_8, learning *Computer Architecture* is better.

Formally, for a learner who has a course selection list C, in situation i, the recommended course due to learning efficiency is

$$r \middle| b_{ir} = \max\left(\left\{b_{ic} \middle| c \in C\right\}\right)$$

The b_{ic} indicates course c's learning efficiency in situation i. With this expression, a recommendation system about when to learn what could be built.

5. Conclusion

With the growth of mobile internet, mobile e-learning is also rapidly developing. Learning Analytics in this environment is introduced in this paper. The influence of learning posture and time period defined as learning situation on learning efficiency was considered. Each situation has a different learning efficiency which affects the learning performance. Then the paper proposed a model to evaluate learning efficiencies of each situation, and used the method with collected data of the course *Computer Architecture* and *Business English* to get results. Those results demonstrates that learning through a PC has a higher learning efficiency than mobile, and sitting also has a little higher rate than lying. Difference in *Business English* is slighter than *Computer Architecture*. The results and the model with the method could help making suggestions about improving courses, especially those for mobile. Also, with the results, a recommendation system based on learning efficiencies could be made, which helps learners learn better. This paper simply considered 3 basic learning situations, which could not completely represent the real situations. So in the future, the authors would improve the posture recognition program to identify more realistic situations and make analysis on these more detailed data, as well as more other facts that may influence the learning efficiency.

Acknowledgements

We'd like to thank Xiaoming Ju for leading the whole project, Bingtao Hu for building and maintaining the server side and Chen Chen for helping collecting data. This research is supported by the Shanghai University Knowledge Service Platform (No.ZF1323) and Software/Hardware Co-design Engineering Research Center, MoE.

References

StatCounter Global Stats (2014). Mobile internet usage soars by 67%. Retrieved July 29, 2015, from http://gs.statcounter.com/press/mobile-internet-usage-soars-by-67-perc

U. S. Department of Education, Office of Educational Technology (2012). Enhancing Teaching and Learning through Educational Data Mining and Learning Analytics: An Issue Brief. Washington, D. C.

Selim, H. M. (2007). Critical success factors for e-learning acceptance: Confirmatory factor models. *Computers & Education*, 49, 396-413.

Bhuasiri, W., Xaymoungkhoun, O., Zo, H., Rho, J. J., & Ciganek, A. P. (2012). Critical Success Factors for E-Learning in Developing Countries: A Comparative Analysis between ICT Experts and Faculty. *Computer & Education*, 58(2), 843-855.

Mbarek, R., & Zaddem, F. (2013). The examination of factors affecting e-learning effectiveness. *International Journal of Innovation and Applied Studies*, 2(4), 423-435.

Zhu, H., Zhang, X., Wang, X., Chen, Y., & Zeng, B. (2014). A case study of learning action and emotion from a perspective of learning analytics. *IEEE 17th International Conference on Computational Science and Engineering*, 420-424.

Foerster, F., Smeja, M., & Fahrenberg, J. (1999). Detection of posture and motion by accelerometry: a validation in ambulatory monitoring. *Computers in Human Behavior*, 571–583.

MacQueen, J. (1967). Some Methods for Classification and Analysis of Multivariate Observations. *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, 281-297.

Anil, K. J. (2010). Data clustering: 50 years beyond K-Means. *Pattern Recognition Letters*, 31(8), 651-666.

Fix, E., & Hodges, J. L. (1951). Discriminatory analysis, nonparametric discrimination. USAF School of Aviation Medicine, Randolph Field, Tex., Project 21-49-004, Rept. 4, Contract AF41(128)-31, February 1951.

Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., Ng, A., Liu, B., Yu, P. S., Zhou, Z., Steinbach, M., Hand, D. J., & Steinberg, D. (2008). Top 10 Algorithms in Data Mining. *Knowledge and Information Systems*, 14, 1-37.