Journal of ICT, Design, Engineering and Technological Science (JITDETS)
VOL. 2, NO. 1, pp. 16-22, 2018
DOI: https://doi.org/10.33150/JITDETS-2.1.3

ORIGINAL CONTRIBUTION

Evaluation for e-Learning Website of Physics by Browsing Path Analysis and Cluster Analysis of Access Log

Tomoshige Kudo [1*], Keita Nishioka [2], Akira Nakamura [3]

[1, 2, 3] Kanazawa Institute of Technology, Academic Foundations Programs, Nonoichi, Japan

***Abstract*— KIT Physics Navigation, a self-adaptive e-learning website of physics covering study contents for high school and university students, was published on the web in March 2016. It was built on the concept that "one web page should contain one topic". For the first time, the access log analysis was performed on this website by examining how visitors browsed the webpages and deepened their understandings. It is noted that this analysis was carried out by using only the access logs acquired from the visitors who had browsed a webpage entitled "Uniformly accelerated linear motion" at least one time to extract the browsing path of the visitors who had an interest in the topic of the webpage. As a result, it was found that most of the visitors deepened their understandings of physics in stages by browsing from the web pages about fundamental topics to those about advanced topics. Furthermore, cluster analysis, which is widely known as the unsupervised learning method of machine learning, was performed on this website. Here, Ward's Method was applied, and the variables were the number of visits and the visit duration. The result showed that the web pages about the following topics, "Derivation of uniformly accelerated linear motion from graph" and "Derivation of uniformly accelerated linear motion by using integration", was classified as the group which had a large number of visits and long visit duration by the dendrogram. In the future, the websites need further improvements based on the results of these analyses.**

***Index Terms*— Cluster Analysis, Access Log, e-Learning**

Received: 21 April 2018; **Accepted:** 30 May 2018; **Published:** 23 June 2018

I. INTRODUCTION

By Conceive, Design, Implement, Operate (CDIO) initiative [1], it is suggested that future engineers should foster the fundamental knowledge of mathematics, science, etc. and fundamental expertise. However, there is no comprehensive e-learning website of Science, Technology, Engineering, and Mathematics (STEM) which is built on the concept that "one webpage should contain one topic" and visualizes the connections among these topics. Therefore, the self-adaptive e-learning websites of mathematics [2], physics [3], and engineering [4] have been developed to establish the STEM learning environment [5].

At first, KIT Mathematics Navigation, a self-adaptive e-learning website of mathematics, was published on the web in 2004 [2], and it consists of more than two thousand webpages so far. Then, in March 2016, KIT Physics Navigation, a self-adaptive e-learning website of physics, was published on the web [3]. This website covers the study contents not only for high school students but also for university students based on the concept that "one webpage should contain one topic" [6]. There are more than seventy webpages about mechanics in study contents for high school students so far. The webpages about electromagnetics, wave mechanics, and thermodynamics are currently in production.

Moreover, KIT Mathematics Navigation and KIT Physics Navigation can be translated into any language by Google translation. At present, they have a lot of access to laboratories and universities. For example, in 2018 only, they had more than three million accesses in total. This number includes many accesses from mobile users. To make STEM's e-learning website more suitable for learners, access log analysis is needed. It is one of our main goals to make a contribution to society by establishing comprehensive learning environment for the students in high school and university, engineers, and working people.

II. LITERATURE REVIEW

Various e-learning websites have been developed in recent years [7]. In learning strategy, e-learning websites can be categorized into two types; instructive learning and constructive learning [7, 8, 9]. For example, the traditional step by step learning by using text books and Massive Open Online Courses (MOOC) [10, 11, 12] is categorized into instructive learning. On the other hand, keyword research websites such as Wikipedia, Wolfram MathWorld [13], Google, Yahoo and Bing provide the suitable e-learning material for the need of individual learners are categorized into constructive learning. If learners do not understand the content in the text book, they often use such search engines to comprehend it. Instructive learning and constructive learning complement each other.

Meanwhile, many educators have been conducting an analysis of access log data generated by the operating system like Learning Management System (LMS) to improve the learning environments [14, 15]. How-

*Corresponding author:Tomoshige Kudo
†Email: kudo@neptune.kanazawa-it.ac.jp

ever, there are few research studies for the analyses of the access log data [16, 17, 18].

KIT Mathematics Navigation [5] and KIT Physics Navigation belong to constructive learning, and are public-facing e-learning websites. Since the main characteristic of our e-learning website is the concept that "one webpage should contain one topic", it is easy to perform the analysis of the access log data in detail to improve the learning environment. KIT Physics Navigation was made by considering the connection between the concepts of mathematics and physics [19] and the connection between the contents for high school and university [3]. Moreover, it was made to visualise the knowledge structure [20].

The access log analysis for the number of visits and the visit duration had been already performed in KIT Mathematics Navigation [21]. The access log analysis for the difference between mobile users and PC users had been also performed [21]. In order to examine visitor's browsing path and their way of learning, the comprehensive website which was built based on the concept that "one webpage should contain one topic" is necessary.

In this report, the browsing path analysis of the access logs was performed to examine how visitors browsed the webpages in KIT Physics Navigation and deepened their understandings of physics. Furthermore, the cluster analysis of the access logs was conducted by using Ward's Method. The webpages were classified by the dendrogram in the cluster analysis. Here, the analysis of the access logs was performed by using logs of the visitors who had browsed a webpage entitled "Uniformly accelerated linear motion" at least one time to extract the browsing path of the visitors who had an interest in this topic.

III. METHODOLOGY

The methodologies of access log analysis are: 1) Browsing path analysis and 2) Cluster analysis.

These methodologies are explained in detail as follows.

A. Browsing Path Analysis

The webpage entitled "Uniformly accelerated linear motion" is shown in Fig. 1.

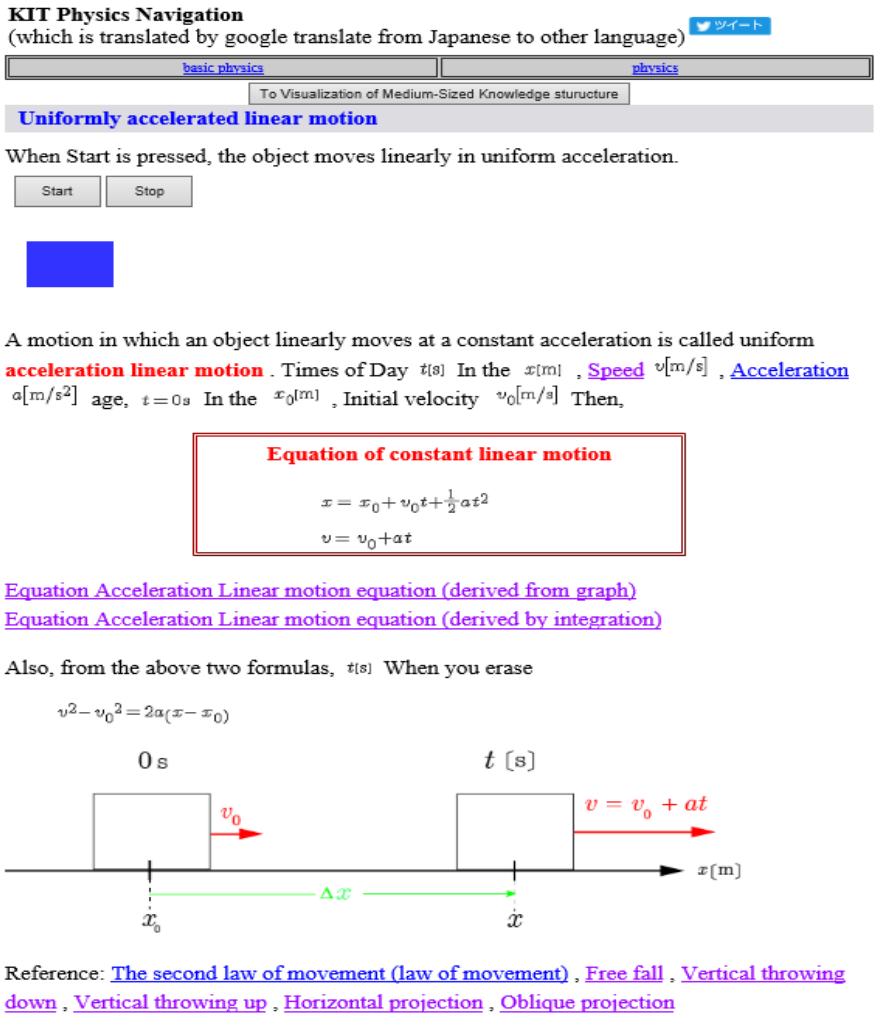

Fig. 1. A webpage entitled "Uniformly accelerated linear motion"

On this webpage, only the topic of "Uniformly accelerated linear motion" was explained by using the motion of an object, figure, and equations as the visitors can understand it smoothly. It was designed to make visitors interested in the motion of the object first. Specifically, when they push the start button, a blue object will move automatically with uniformly accelerated linear motion. Since this simulation will be operated by JavaScript, the visitors can see the motion by using any device. The equations of uniformly accelerated linear motion and its' figure were set below the simulation. For the visitors who are accessing our webpages from their mobiles, the webpage was developed in an easy-to-read layout.

The structure of the hyperlinks was explained as follows:

a. On top of the webpage, hyperlinks for the following topics were set
–Table of contents for Basic Physics
–Table of contents for Advanced Physics

b. Within a sentence, hyperlinks for the topics of "Velocity" and "Acceleration" were set.

c. Under the equations, hyperlinks for the following topics: "Derivation of uniformly accelerated linear motion from graph" and "Derivation of uniformly accelerated linear motion by using integration" was set.

d. On the bottom of the webpage, hyperlinks for the related topics and the applied contents were set.

Here, a lot of browsing paths of the visitors who moved directly to other webpages which had no links on the webpage entitled "Uniformly accelerated linear motion" were found. This phenomenon happened for the following reason. When visitors pushed the back button on the webpage entitled "Uniformly accelerated linear motion", our system is designed not to count it as the access log. And if they go back to the table of contents, and go to other webpages, the system regards this action as they directly moved to other webpages from the webpage entitled "Uniformly accelerated linear motion".

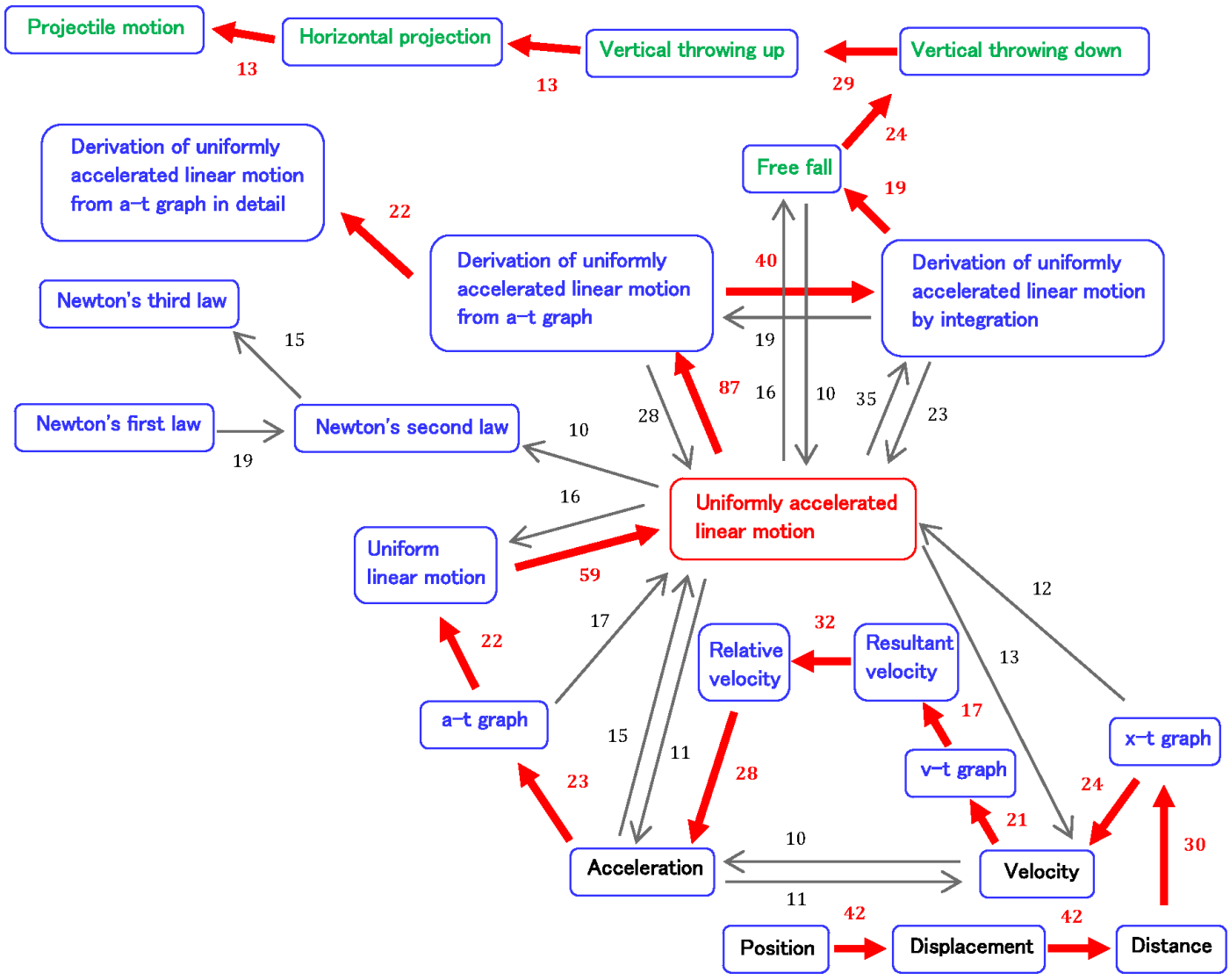

Fig. 2. The browsing paths of the visitors

Fig. 2 shows the visitors' browsing paths acquired from the access logs. The directions of arrows indicate the directions the visitors moved to, and the numbers of times the visitors moved are written beside the arrows. These numbers were generated from only the access logs of the visitors who browsed the webpage entitled "Uniformly accelerated linear motion" at least one time. In fact, the actual numbers are much bigger than the numbers indicated in Fig. 2. The access log analysis was performed for 10 months from January to October in 2018, and the logs which were considered as web crawlers were eliminated. The arrows are drawn when the numbers of times the visitors moved are more than 10 times. Red arrows mean visitors' main browsing path.

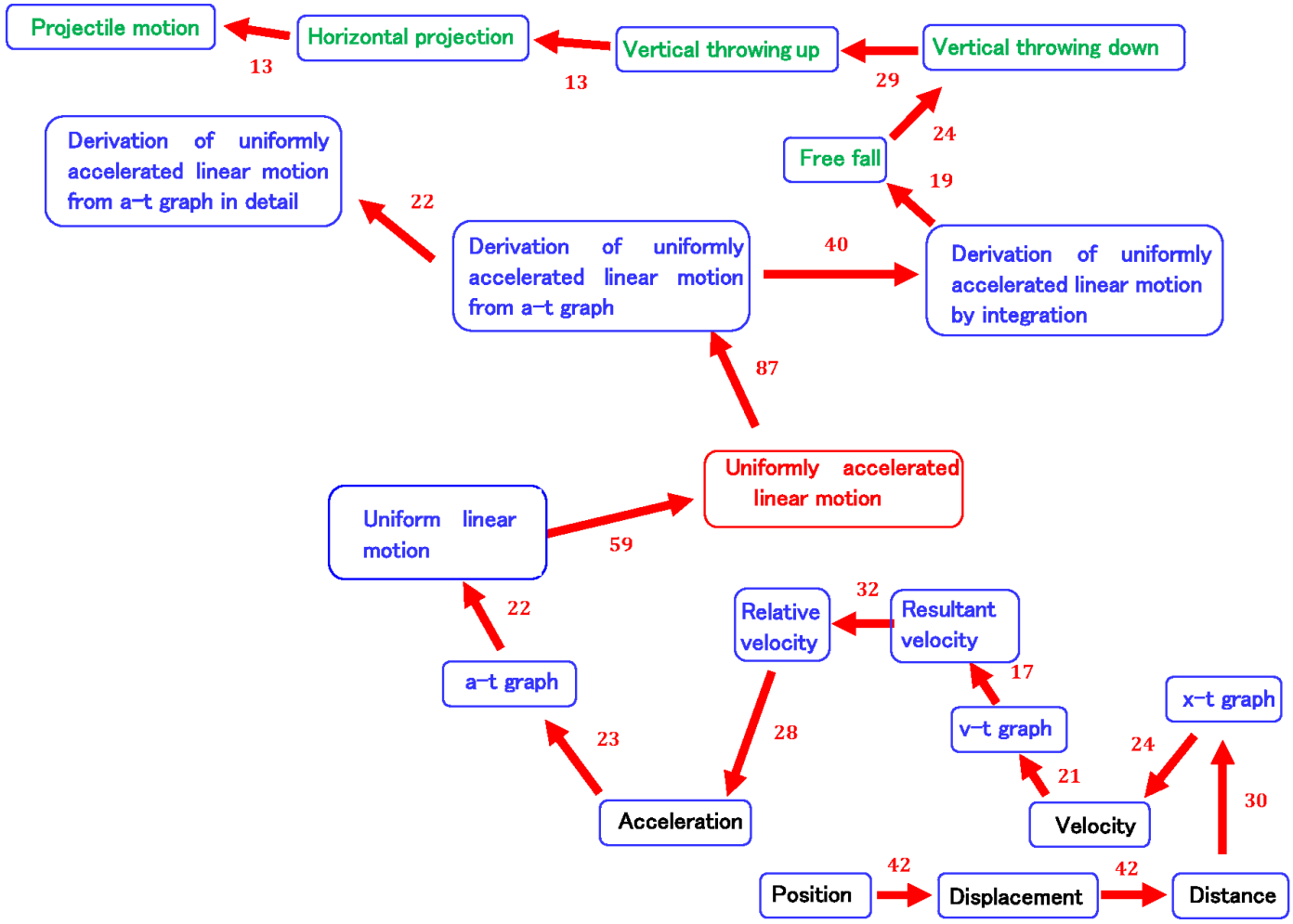

Fig. 3. Visitors' main browsing path

Fig. 3 shows visitors' main browsing path. Only the arrows which are most frequently exited from one webpage to another webpage were indicated. The main visitor started to move from the webpage entitled "Position" in a fundamental knowledge section, and passed through the webpage entitled "Uniformly accelerated linear motion", and then reached a webpage entitled "Projectile" in an applied knowledge section. Judging by this, lots of the visitors deepened their understandings of physics in stages with a start browsing the webpage in fundamental knowledge section to the webpage in an applied knowledge section.

B. Cluster Analysis

Fig. 4 shows the dendrogram of cluster analysis. Cluster analysis which is well known as unsupervised learning method of machine learning was performed here. And Ward's Method and Squared Euclidean distance were applied to it. The variables were the number of visits and the visit duration. The analysis was conducted on only the access logs acquired from the visitors who had browsed the webpage entitled "Uniformly accelerated linear motion" at least one time. The visit duration was calculated by the average value of the time in which the above-mentioned visitors browsed the webpage. The access logs of the visitors who hadn't browsed the webpage were excluded consecutively. For example, the access logs of the visitors who had browsed only one webpage were excluded, because it was impossible to calculate the time they had spent on the webpage. The access logs of the visitors who had stayed more than 30 minutes on one webpage were excluded too, since there is possibility that they just left it open without browsing it.

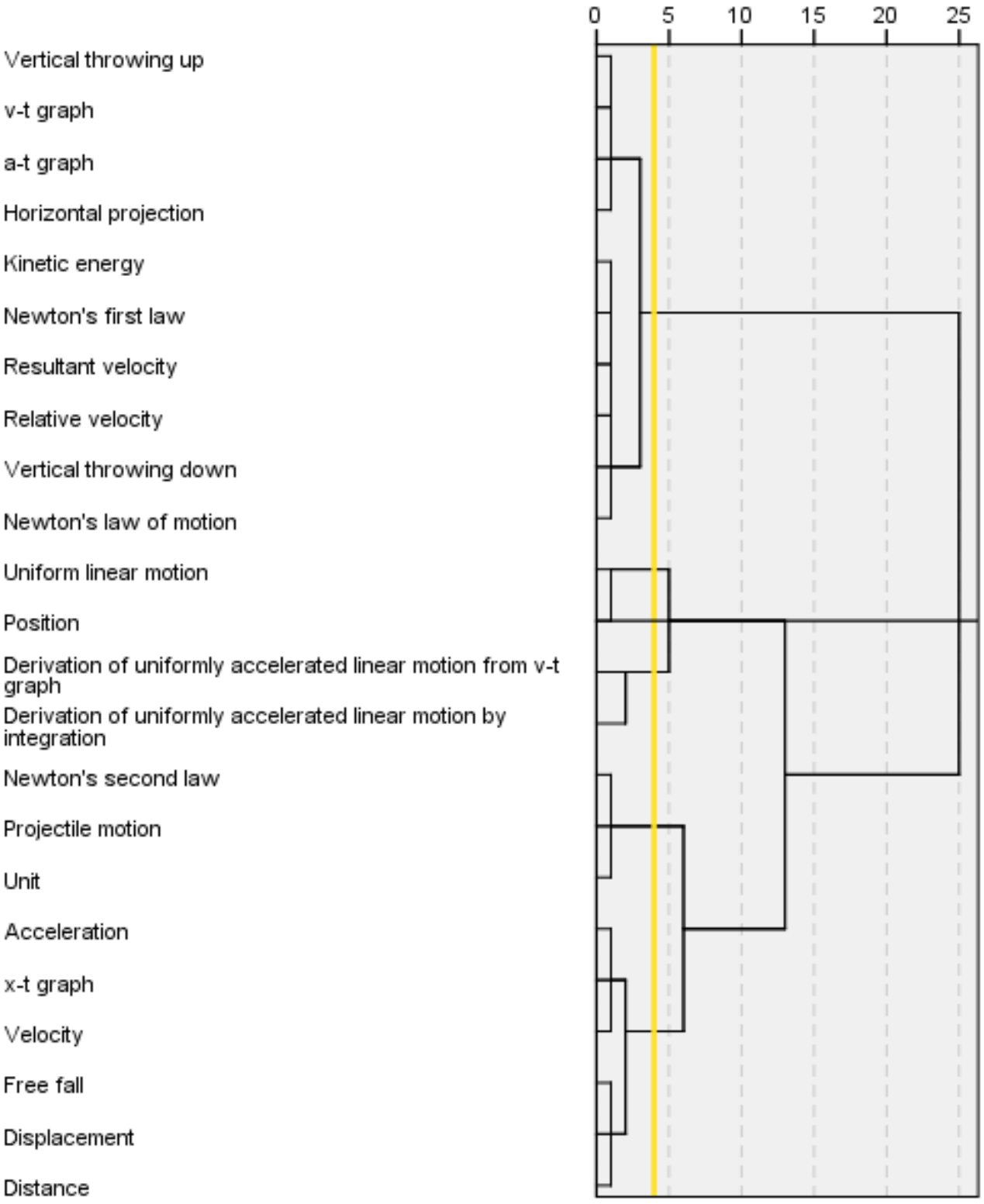

Fig. 4. Dendrogram of clustering

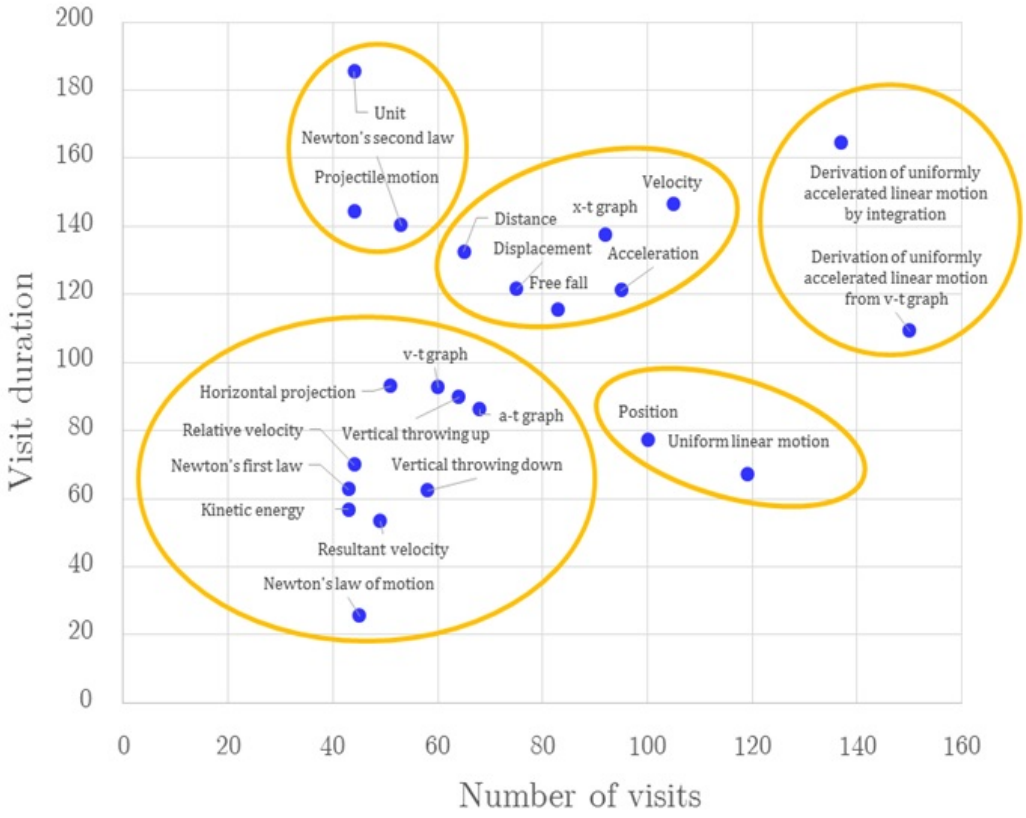

Fig. 5. Map of cluster analysis

Fig. 5 shows a map of cluster analysis. The horizontal axis indicates the number of visits and the vertical axis shows the visit duration. The topics of physics were separated by the orange line in Fig. 4, and were circled in Fig. 5. As a result, it was found that the webpages entitled "Derivation of uniformly accelerated linear motion from graph" and "Derivation of uniformly accelerated linear motion by using integration" were classified by the dendrogram as the group which had a large number of visits and the long visit duration. The visitors had many interests and concerns for this group, and they spent a lot of time to understand the contents. On the other hand, short visit duration on a webpage seems to indicate that visitors considered the difficulty of the webpage was low, or that there was little explanation for them on the webpage, etc. Fig. 5 is one of the guidelines to evaluate whether the difficulty and the quantity of the content of the webpage are adequate or not for the visitors.

IV. RESULTS

The access log analysis was performed on the webpages in the e-learning website of physics, KIT Physics Navigation. The analysis was conducted on the browsing paths acquired from only the visitors who had browsed the webpage entitled "Uniformly accelerated linear motion" at least one time. By analysing the main browsing path which visitors took, it was found that they started to browse a webpage in fundamental knowledge section to applied knowledge section and gained a better understanding of physics in stages. Moreover, the cluster analysis was performed. Here the number of visits and the visit duration were variables. As a result of these analyses, two kinds of webpages entitled "derivations of uniformly accelerated linear motion" were classified as the group having a large number of visits and long visit duration.

V. DISCUSSION

Most of the visitors started to browse a webpage in the fundamental knowledge section to applied knowledge section. However, some of the visitors started to browse a webpage in applied knowledge section to fundamental knowledge section in Fig. 2. This browsing path means link back learning. Some of the visitors moved directly to other webpages which had no links on the webpage. KIT Physics Navigation is a self-adaptive e-learning website of physics belonging to constructive learning [5], and it can meets the different learning needs of individual learners.

Since two kinds of webpages regarding the derivations were classified as the group having a large number of visits and long visit duration by using the cluster analysis, the visitors who browsed the webpage entitled "Uniformly accelerated linear motion" at least one time have the most interest in the derivations. This result suggested the contents of the webpages regarding the derivations are very substantial. The webpages entitled "Position" and "Uniform linear motion" were classified as the group having a large number of visits and short visit duration, and thus it is thought that the contents of these webpages are easy and understandable for them.

The learners get knowledge step by step from the instructive learning like MOOC[10, 11]. On the other hand, the learners gain the knowledge they need from constructive learning like Wikipedia and Wolfram MathWorld [13]. There are few research studies for the analysis of the access log data for public-facing e-learning websites in the field of education. Since KIT Physics Navigation and KIT Mathematics Navigation are built on the concept that "one webpage should contain one topic", the learners' motives to visit a particular webpage are clear. Therefore, it is significant for us to perform an analysis of the access log data to improve learning environment and to meet the different learning needs of learners.

VI. CONCLUSION AND IMPLICATIONS

Based on the result of these analyses, it is important to modify the structure of hyperlinks and the contents of our webpages for the visitors so as they can obtain the knowledge smoothly, and understand the connection between the fundamental knowledge and the advanced knowledge more effectively. Moreover, it is needed to modify the webpages based on the evaluations using cluster analysis.

In this report, to extract the browsing paths of the visitors who had an interest in the webpage entitled "Uniformly accelerated linear motion", the analysis was conducted on the access logs acquired from only the visitors who had browsed this webpage at least one time. It is a future issue for us to perform the access log analysis of the visitors who had an interest in the topics other than this, and to conduct the access log analysis on all the visitors' logs. It is necessary for us to continuously improve our websites to suit visitors' needs based on the results of access log analysis.

Currently, the webpages have been constructed in the contents of mathematics, a part of the contents of physics and engineering [20]. It is important to exert further efforts to enrich the contents of the webpages to establish STEM's e-learning website. Furthermore, it is also necessary to modify the layouts of STEM's webpages for the students in high school and university, engineers, and working people to make it easy for them to acquire the knowledge smoothly by using their mobiles. In order to prevent the declining of the number of science students today, it is needed to keep making improvements on our self-adaptive e-learning websites covering STEM's knowledge and visualizing the connections between these knowledge based on the CDIO initiative [1].

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number JP16K01137. We appreciate the student staff who engaged in developing our e-learning website.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] E. F. Crawley, ``The CDIO syllabus: A statement of goals for undergraduate engineering education,'' Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA, MIT CDIO Report 1, 2001. [Online]. Available: https://bit.ly/34R10wa

[2] Kanazawa Institute of Technology. KIT mathematical navigation. [Online]. Available: https://bit.ly/2CFFheA

[3] Kanazawa Institute of Technology. KIT physical navigation. [Online]. Available: https://bit.ly/2NGqUNy

[4] Kanazawa Institute of Technology. KIT engineering navigation. [Online]. Available: https://bit.ly/2X9P7yH

[5] A. Nakamura, ``Self-adaptive e-learning website for mathematics,'' *International Journal of Information and Education Technology*, vol. 6, no. 12, pp. 961-965, 2016. doi: https://doi.org/10.7763%2Fijiet.2016.v6.825

[6] T. Kudo, A. Nishioka, and A. Nakamura, ``"KIT physics navigation" showing relationship between high school and university,'' in *Proceedings of the 14th International CDIO Conference,* Kanazawa Institute of Technology, Kanzawa, Japan, Kanazawa, Japan, 2018, p. 223–229.

[7] S. Negash, M. Whitman, A. Woszczynski, K. Hoganson, and H. Mattord, *Handbook of Distance Learning for Real-Time and Asynchronous Information Technology Education*. Hershey, NY: IGI Global, 2008.

[8] Y. Atif, R. Benlamri, and J. Berri, ``Learning objects based framework for self-adaptive learning,'' *Education and Information Technologies*, vol. 8, no. 4, pp. 345-368, 2003. doi: https://doi.org/10.1023/B:EAIT.0000008676.64018.af

[9] C. L. S. Tablatin, F. F. Patacsil, and P. V. Cenas, ``Design and development of an information technology fundamentals multimedia courseware for dynamic learning environment,'' *Journal of Advances in Technology and Engineering Research*, vol. 2, no. 6, 2016. doi: https://doi.org/10.20474/jater-2.6.5

[10] Y. X. Guo, X. B. Chen, and P. F. Li, ``The effect of massive open online course on the universities in the western region,'' *Advanced Materials Research*, vol. 926, pp. 4657-4660, 2014. doi: https://doi.org/10.4028/www.scientific.net/AMR.926-930.4657

[11] D. T. Seaton, Y. Bergner, I. Chuang, P. Mitros, and D. E. Pritchard, ``Who does what in a massive open online course?'' *Communications of the ACM*, vol. 57, no. 4, pp. 58-65, 2014. doi: https://doi.org/10.1145%2F2500876

[12] H. Hashim, S. Salam, and S. N. M. Mohamad, ``Investigating learning styles for adaptive Massaive Open Online Cource (MOOC) learning,'' *Journal of Advances in Humanities and Social Sciences*, vol. 3, no. 5, 2017. doi: https://doi.org/10.20474/jahss-3.5.4

[13] Wolfram MathWorld,The Web's Most Extensive Mathematics Resource. [Online]. Available: https://bit.ly/2XekPuL

[14] H. Zhang and K. Almeroth, ``Moodog: Tracking student activity in online course management systems,'' *Journal of Interactive Learning Research*, vol. 21, no. 3, pp. 407-429, 2010.

[15] M. K. Khribi, M. Jemni, and O. Nasraoui, ``Automatic recommendations for e-learning personalization based on web usage mining techniques and information retrieval,'' *Educational Technology & Society*, vol. 12, pp. 30-42, 2009.

[16] D. Shen, X. Wang, and H.-L. Chen, ``Managing web-based learning resources for K-12 education: Lessons learned from web analytics,'' in *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education,* Association for the Advancement of Computing in Education (AACE), Las Vegas, Nevada, 2008, pp. 470-475.

[17] D. Ellis, ``Google analytics as a tool in the development of e-learning artefacts: A case study,'' in *Australian Society for Computers in Learning in Tertiary Education Annual ASCILITE Conference,* Wellington, New Zealand, 2012, pp. 299-303.

[18] W. Fang, ``Using google analytics for improving library website content and design: A case study,'' 2007. [Online]. Available: https://bit.ly/33IBwBd

[19] T. K. K. Nishioka and A. Nakamura, ``Learning support website of physics with emphasis on connection with mathematics,'' in *9th International Multi-Conference on Complexity, Informatics and Cybernetics,* Orlando, FL, 2018, p. 155 – 157.

[20] A. Nakamura, T. Kudo, and K. Nishioka, ``Development of the visualizing system of knowledge structure based on stem e-learning website,'' in *Proceedings of the 9th International Conference on Language, Innovation, Culture & Education,* Bangkok, Thailand, 2018, pp. 55-61.

[21] A. Nakamura, ``Usage analysis of the free math learning site,'' in *Proceedings of SICSS,* Shanghai, China, 2012, p. 668–673.