DOI: https://doi.org/10.13161/kibim.2016.6.4.035

기계적 학습의 알고리즘을 이용하여 아파트 공사에서 반복 공정의 효과 비교에 관한 연구

Identifying the Effects of Repeated Tasks in an Apartment Construction Project Using Machine Learning Algorithm

김현주[1)]
Kim, Hyunjoo[1)]

Received December 6, 2016; Received December 20, 2016 / Accepted December 21, 2016

ABSTRACT: Learning effect is an observation that the more times a task is performed, the less time is required to produce the same amount of outcomes. The construction industry heavily relies on repeated tasks where the learning effect is an important measure to be used. However, most construction durations are calculated and applied in real projects without considering the learning effects in each of the repeated activities. This paper applied the learning effect to the repeated activities in a small sized apartment construction project. The result showed that there was about 10 percent of difference in duration (one approach of the total duration with learning effects in 41 days while the other without learning effect in 36.5 days). To make the comparison between the two approaches, a large number of BIM based computer simulations were generated and useful patterns were recognized using machine learning algorithm named Decision Tree (See5). Machine learning is a data-driven approach for pattern recognition based on observational evidence.

KEYWORDS: Machine Learning, Computer Simulation, Learning Effect, Decision Tree

키워드: 기계적 학습, 컴퓨터 시뮤레이션, 학습 효과, 의사결정 나무

1. Introduction

This paper intends to improve the current practice of managing a project by applying the learning effect to projects. Learning effect will be taken into account in this research to produce schedules with higher accuracy and identify the patterns within the schedule. Learning effect is the reality that individuals or a crew become more efficient executing a construction task when performing that same task repeatedly (Jarkas & Horner, 2011). For example, a construction crew responsible for installing floor tiles in a new building may be able to complete a certain amount of square footage the first day, but the second day they will be more familiar with the site and the productivity of installing floor tiles will increase. This trend would continue until the amount will level off and the tile installation crew will reach an optimum level of installation quantity. This demonstrates the learning effect from repetitive construction and when incorporating this into estimation of activity duration the results can be improved (Roy, 1992).

Without knowing the effect of learning effect people do not know the effect of activity splitting. Activity splitting is a part of construction when an activity is started, then stopped, and restarted again on a future date. Many activities do not have the capability to be stopped and started as described, and others could efficiently do so, but there are some activities that would benefit from having this characteristic. Activity splitting can keep several activities progressing in parallel (Gordon & Tulip, 1997) and can be used to obtain an enhanced schedule and increase resource utilization (Hariga & El-Sayegh, 2011).

This paper utilizes object-oriented technology in Building

1)정회원, 서울시립대학교 글로벌건설학과 (hkim01@uos.ac.kr) (교신저자)

Information Modeling (BIM). BIM is not simply a 3D model with geometry, but a process that has been used for a wide range of purposes to improve performance of a facility through the entire life cycle (Lu, Peng, Shen, & Li, 2013). This paper produces numerous schedule results and oftentimes patterns cannot be readily recognized from multiple simulations of schedules, thus machine learning tools will be applied to analyze the construction schedules.

The construction industry heavily relies on repeated tasks where the learning effect is an important measure to be used. However, most construction durations are calculated and applied in real projects without considering the learning effects in each of the repeated activities. This paper applied the learning effect to the repeated activities in a small sized apartment construction project. The result showed that there was about 4.5 days of difference in construction duration. To make the comparison between the two approaches, a large number of BIM based computer simulations were generated and useful patterns were recognized using machine learning algorithm named Decision Tree (See5). Machine learning is a data-driven approach for pattern recognition based on observational evidence (Worden & Manson, 2007)

2. Literature Review

The learning curve effect was first defined mathematically by T.P. Wright in 1936 in relation to airplane production (Wright, 1936). He found a 20 percent reduction in the amount of man-hours required to assemble an airplane each time the unit production volume doubled. Cunningham (1980) continued in this research to recognize how learning effect applies to many other products produced in US industry. Some examples of this are the Model-T Ford production from 1910 to 1926, steel production from 1920 to 1955, and disk memory drives from 1975 to 1978. He found that during that time period, the disk memory drives had a learning curve slope of 76 percent, which indicates that each time the unit volume is doubled the production cost is 76 percent of what it was previously (Cunningham, 1980).

Thomas (1986) went on to consider this occurrence in the construction industry and recognized that productivity improved when performing repetitive construction tasks. Eight reasons were identified for this occurrence: (1) Increased worker familiarization; (2) improved equipment and crew coordination; (3) improved job organization; (4) better engineering support; (5) better day-to-day management and supervision; (6) development of more efficient techniques and methods; (7) development of more efficient material supply systems; and (8) stabilized design leading to fewer modifications and rework (Thomas, Mathews, & Ward, 1986).

Learning effect of a particular production can be defined by a percentage, which establishes the slope of the learning curve. A greater learning occurs when the learning effect percentage is lower, and no learning takes place when the rate is 100% (Thomas et al., 1986). Thomas researched five mathematical models with the goal of identifying a reliable model for predicting future performance. The models used were the straight-line power model; the Stanford "B" model; the cubic power model; the piecewise (or stepwise) model; and the exponential model. He used these models with the time data for erecting and setting 466 precast concrete floor planks and concluded that straight-line model is the most simple to use, but not always reliable; and that nonlinear models better represent the effect of disruptions or delays upon the learning rate, but ultimately he realized that additional research was needed to identify a reliable learning curve prediction model (Thomas et al., 1986).

In the construction industry, the learning effect has not been broadly studied by many researchers. Therefore, this research would take on the effect of learning effects in construction activities and find out how much difference it would make in terms of the schedule duration.

3. Methodology

The main factors included in this research include an object-oriented BIM model with a collection of building components that makeup the construction activities required to construct the facility.

As previously described, the learning effect will be taken into account when consider the duration for each of the activities. While applying the learning effect to the construction activities, the methodology presented includes consideration for activity splitting to identify significant patterns from resulting durations. Figure 1 shows the general methodology that is

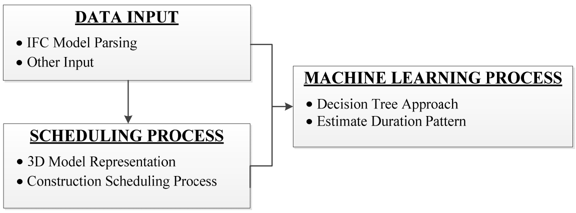

Figure 1. General methodology

applied in this research for an integration of object-oriented BIM, learning effect, and activity splitting for estimating durations for a construction project.

The general methodology is composed of 3 mains steps which include data input, scheduling process, and machine learning process. The first step of data input is primarily composed of collecting data from a 3D object-oriented BIM model using IFC modeling technology. This data collected from an IFC model is used in both the scheduling process and the machine learning process. The second step of the scheduling process first represents a 3D model of the facility then incorporates learning effect and activity splitting in the construction duration analysis. After the multiple schedules are produced from the scheduling process step, the machine learning process is commenced using the Decision Tree analysis approach to identify significant patterns from the multiple durations.

The second step of the overall process is scheduling process in BIM based computer simulation, which is presented in Figure 2. This step includes a BIM model representation and construction scheduling process. The BIM model representation includes draw 3D slab shape, draw 3D wall shape, draw 3D window shape, and draw 3D door shape. The drawing is represented in a 3D modeling software for visual verification of the facility from which the durations will be analyzed. The scheduling process goes through a procedure to identifying if certain activities can be split or not. The first step is to sequence activities then read the first activity from the scheduled list. Then determine if that activity can be split or not. If it can, list all the split parts of the activity, retrieve the extended duration data for each part form the learning database, revise the duration of each part based on the data, and schedule the parts of the activity. The update the unscheduled list and adjust the table of resource and time. Then determine if there are any

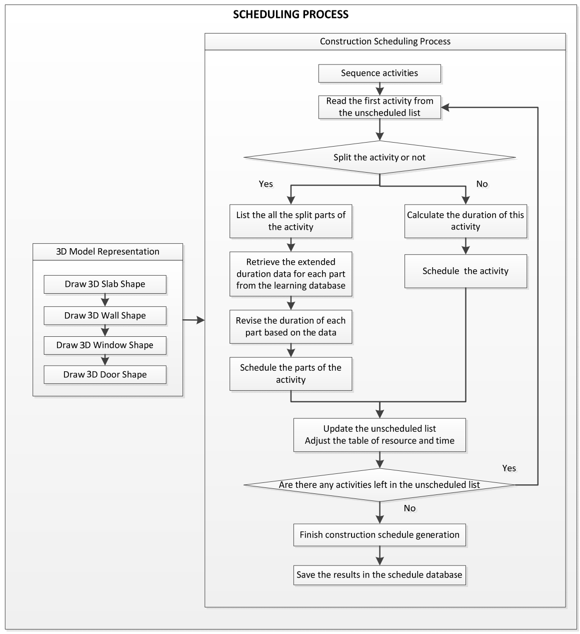

Figure 2. Scheduling process in BIM based computer simulations

activities left in the unscheduled list. If there are, continue back to the first step of reading the first activity from the unscheduled list. If the activity cannot be split, then calculate the duration of the activity and schedule activity then update the unscheduled list and adjust the table of resource and time. At this point continue the loop until all the activities are considered for splitting. Once there are no remaining activities on the unscheduled list the construction schedule generation is finished and the results are saved in the schedule database.

The third step of the process is complete while the machine learning process is finished. The purpose of the machine learning process is to identify patterns from the schedules which were produced, and this process can be seen in Figure 3.

4. Case Study

A BIM model of a four-story building was used for this case study. The goal was to determine the impact of activity splitting and learning effect on the total duration of construction for the project. AutoDesk Revit 2016 was used for this case study to create a BIM model, which is shown in Figure 4. The BIM model used is a four-story apartment building that

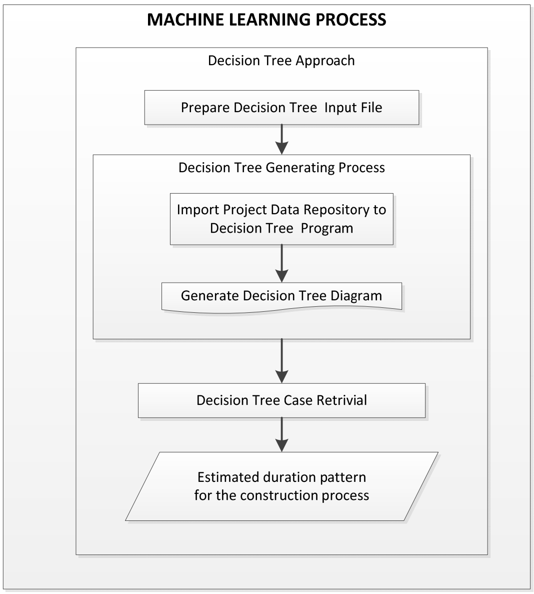

Figure 3. Machine learning process in identifying useful patterns from the repeated activities

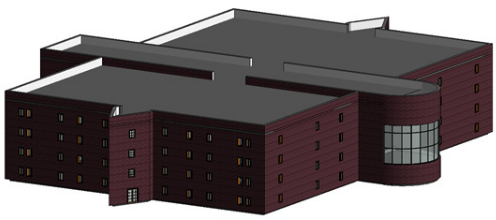

Figure 4. BIM model for 4-story apartment building

includes 15 construction components. The construction components collected were selected for this project as a representative of a complete project to demonstrate the effect of activity splitting and learning effect.

Table 1 shows the 15 activities that were used for this project. Based on the nature of the activity, several of the activities cannot be split. Some examples of this are foundation, framing, and roofing. This research does not address all the activities that would be included in typical apartment building, but the focus is on the major activities. The focus is to present a prototype that could be applied to a larger number of activities and different building types.

Decision tree analysis was used on this project to determine the total duration based on multiple simulations. 181 cases were created for the four-story apartment building with variations in the schedule, which included splitting and non-splitting and different number of activities split. Figure 5 shows the total duration in days for each of the cases.

Table 1. Construction activities for four-story apartment building

Activity #	Duration (days)		Description
	Per floor	Total	
1	–	4	Foundation
2	4x3	12	Framing
3	–	4	Roofing
4	4x4	16	Rough Plumbing
5	4x3	12	Rough Electrical
6	4x3	12	Face Brick
7	4x2	8	Windows Exterior
8	4x1	4	Insulation
9	4x14	16	Drywall
10	4x1	4	Painting
11	4x3	12	Cabinets & Trim
12	4x1	4	Floor Tile
13	4x2	8	Finish Plumbing
14	4x1	4	Finish Electrical
15	4x2	8	Carpets

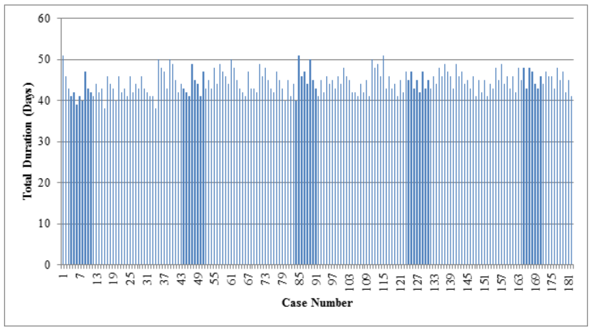

Figure 5. Total duration per case

As the figure shows, there is a variation between each of the cases and it is difficult to identify any trends. These cases presented are based on splitting activities and not splitting activities as well as learning effect and non-learning effect.

The 181 cases used in this research are varied from the base schedule, which is shown in Figure 5. The base schedule was built from the 15 activities shown in Table 1 and you can see from the schedule that each of the activities has sub-activities which relate to the construction of that particular activity on different floors of the apartment building. For example, in Figure 6 the total duration for Framing is 12 days, but that is composed of 4 framing activities: Framing 1, Framing 2, Framing 3, and Framing 4. Each of those framing activities relates to that particular floor in the building. Figure 6 shows what the schedule would

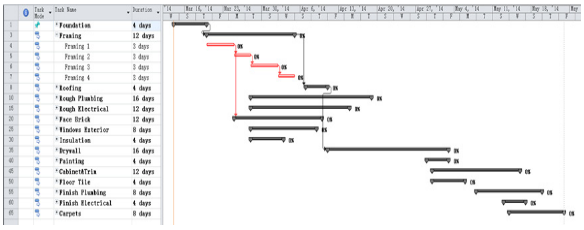

Figure 6. Base project schedule without learning effect (51 days)

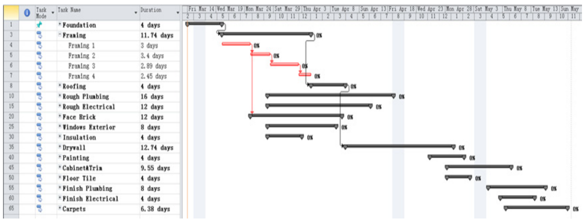

Figure 7. Base project schedule with learning effect (46 days)

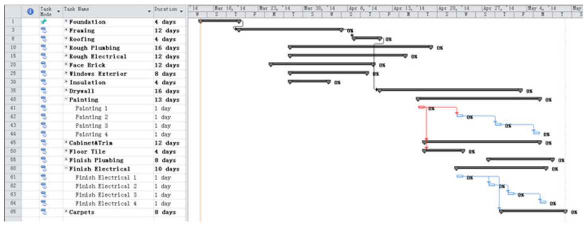

Figure 8. Project schedule – painting and finish electrical split without learning effect (41 days)

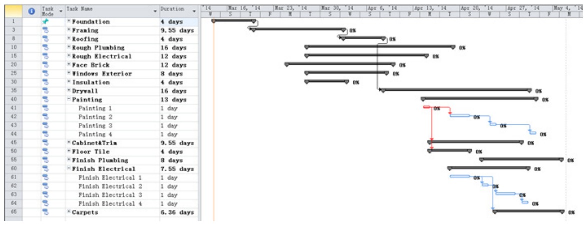

Figure 9. Project schedule – painting and finish electrical split with learning effect (36.5 days)

be if neither activity splitting nor learning effect were considered. Figure 7 shows the project schedule with learning effect applied to the activities.

Figures 6 and 7 show what the total duration would be if activity splitting was not considered and learning effect was considered for one of the cases. The result shown in these figures was that the base project schedule without learning effect is 51 totals days and the base project schedule with learning effect is 46 days. From this we can see that learning effect can shorten the total duration of the project.

The goal of this research was not just to see how learning effect could shorten project duration, but to the effect of the combination of learning effect and activity splitting on the construction schedule. Figures 8 and 9 present construction schedules with activity splitting included. Figure 8 shows what the schedule would be if the activities of painting and finish electrical were split into separate activities based on the level. For example, the first floor of the building would be painted, but there would be a break and then the second floor would be painted, and so on until the fourth floor. In these examples this was applied to both painting and finish electrical. Figure 8 shows the project schedule with these two activities split without learning effect included and the total duration is 41 days. Figure 9 shows these two activities split with learning effect included and the total duration is 36.5 days. This shows that learning effect shorts the estimated total duration of the project. In this case it shortened the total duration by 4.5 days. The goal of this research was not just to look at 2 cases, but to consider a large number of cases. These cases include a combination the number of activities split with and without learning effect.

The decision tree software See5.0 was used in this case study as a data analysis tool to determine trends from activity splitting and learning effect in construction scheduling. The See5.0 sofware is based on the C4.5 decision tree algorithm. Figure 10 shows the results from decision tree analysis of 169 construction schedule cases. There were several factors used in the decision tree analysis: learning effect, activity splitting, number of activities split, and total duration. The total construction duration was the target factor in the decision tree analysis, which was broken into three categories: Short, Medium, and Long. These three categories represented the total duration of the project based on changing the different factors.

As Figure 10 shows, there were cases that had learning effect and some that did not. Learning effect was the first factor and from that factor there was “branches” of factors in the decision tree analysis. What this ultimately shows is the total duration based on the values of each of the factors. From these results we can see which factors influnce a short, medium, or long project duration. In the decision tree

```
Decision tree:

Learning Effect? = No:
:...Windows Exterior = WindowsExterior: Medium (14/1)
:   Windows Exterior = No:
:   :...Floor Tile = FloorTile: Medium (8/1)
:       Floor Tile = No:
:       :...Number of Splitted <= 1: Medium (11/4)
:           Number of Splitted > 1:
:           :...Rough Electrical = RoughElectrical:
:               :...Drywall = Drywall: Long (2)
:               :   Drywall = No: Medium (15/3)
:               Rough Electrical = No:
:               :...Painting = Painting: Medium (6/2)
:                   Painting = No: Long (28/12)
Learning Effect? = Yes:
:...Number of Splitted <= 1: Short (13/3)
    Number of Splitted > 1:
    :...Carpets = Carpets: Medium (8)
        Carpets = No:
        :...Insulation = Insulation: Medium (12/2)
            Insulation = No:
            :...Floor Tile = FloorTile: Short (7/2)
                Floor Tile = No:
                :...Windows Exterior = WindowsExterior: Short (10/4)
                    Windows Exterior = No:
                    :...Face Brick = FaceBrick: Short (11/5)
                        Face Brick = No: Medium (23/7)
```

Figure 10. See5.0 decision tree output

analysis it was assumed that the following construction activities could be split: rough plumbing, rough electrical, face brick, windows exterior, insulation, drywall, painting, cabinet trim, floor tile, and finish plumbing. The decision tree analysis used activity splitting for a combination of these activities to determine the total duration. Also, for each of the activities there was a simulation including learning effect to determine the total duration with that factor considered.

5. Conclusions

This paper utilized BIM technology in improving learning effects in the industry. This research demonstrated that BIM is a process that has been used for a wide range of purposes to improve performance of a facility through the entire life cycle. This paper produced numerous schedule results and showed that oftentimes patterns cannot be readily recognized from multiple simulations of schedules, thus machine learning tools was applied to analyze the construction schedules. Machine learning tools were applied to the multiple schedule simulation results to identify the patterns presented from learning effect algorithms and activity to identify the benefit of using such characteristics of constructions scheduling for a more realistic schedule.

The result showed that there was about 10 percent of difference in duration (one approach of the total duration with learning effects in 41 days while the other without learning effect in 36.5 days). To make the comparison between the two approaches, a large number of BIM based computer simulations were generated and useful patterns were recognized using machine learning algorithm named Decision Tree (See5).

This research applied the learning effects in an apartment construction project. However, it is realized that there needs to be a more fundamental study to be followed to measure the practical effects of the repeated activities in various kinds of different construction projects.

References

Couto, J., & Teixeira, J. (2005). Using linear model for learning curve effect on highrise floor construction. *Construction Management and Economics, 23*(4), pp. 355-364.

Cunningham, J. A. (1980). Management: Using the learning curve as a management tool: The learning curve can help in preparing cost reduction programs, pricing forecasts, and product development goals. *Spectrum, IEEE, 17*(6), pp. 45-48. doi: 10.1109/MSPEC.1980.6330359

Damci, A., Arditi, D., & Polat, G. (2013). Resource Leveling in Line-of-Balance Scheduling. *Computer-Aided Civil and Infrastructure Engineering, 28*(9), pp. 679-692.

Fan, S.-L., Tserng, H. P., & Wang, M.-T. (2003). Development of an object-oriented scheduling model for construction projects. *Automation in Construction, 12*(3), pp. 283-302. doi: http://dx.doi.org/10.1016/S0926-5805(02)00092-4

Gordon, J., & Tulip, A. (1997). Resource scheduling. *International Journal of Project Management, 15*(6), pp. 359-370. doi: http://dx.doi.org/10.1016/S0263-7863(96)00090-7

Hallberg, D., & Tarandi, V. (2011). On the use of open bim and 4d visualisation in a predictive life cycle management system for construction works. *Electronic Journal of Information Technology in Construction, 16*, pp. 445-466.

Halpin, D. W. (1973). *An investigation of the use of simulation networks for modeling construction operations*. University of Illinois at Urbana-Champaign.

Halpin, D. W. (1990). MicroCYCLONE user's manual. *Division of Construction Engineering and Management, Purdue University, West Lafayette, Indiana*.

Hariga, M., & El-Sayegh, S. (2011). Cost Optimization Model for the Multiresource Leveling Problem with Allowed Activity Splitting. *Journal of Construction Engineering and Management, 137*(1), pp. 56-64. doi: doi:10.1061/

(ASCE)CO.1943-7862.0000251

Hijazi, A. M., AbouRizk, S. M., & Halpin, D. W. (1992). Modeling and simulating learning development in construction. *Journal of Construction Engineering and Management, 118*(4), pp. 685-700.

Hossein Hashemi Doulabi, S., Seifi, A., & Shariat, S. (2011). Efficient Hybrid Genetic Algorithm for Resource Leveling via Activity Splitting. *Journal of Construction Engineering and Management, 137*(2), pp. 137-146. doi: doi:10.1061/(ASCE)CO.1943-7862.0000261

Jarkas, A., & Horner, M. (2011). Revisiting the applicability of learning curve theory to formwork labour productivity. *Construction Management & Economics, 29*(5), pp. 483-493. doi: 10.1080/01446193.2011.562911

Kalk, A., & Douglas, S. A. (1980). *INSIGHT: Interactive simulation of construction operations using graphical techniques*: Stanford University, Department of Civil Engineering.

Lluch, J. F. (1981). Analysis of construction operations using microcomputers.

Lu, W., Peng, Y., Shen, Q., & Li, H. (2013). Generic Model for Measuring Benefits of BIM as a Learning Tool in Construction Tasks. *Journal of Construction Engineering and Management, 139*(2), pp. 195-203. doi: doi:10.1061/(ASCE)CO.1943-7862.0000585

Lucas, J., Bulbul, T., & Thabet, W. (2013). An object-oriented model to support healthcare facility information management. *Automation in Construction, 31*(0), pp. 281-291. doi: http://dx.doi.org/10.1016/j.autcon.2012.12.014

Lutz, J. D., Halpin, D. W., & Wilson, J. R. (1994). Simulation of learning development in repetitive construction. *Journal of Construction Engineering and Management, 120*(4), pp. 753-773.

Malaikrisanachalee, S., & Vathananukij, H. (2011). INTEGRATION OF JAVA-BASED BIM WITH SPATIAL DATABASE. *International Journal of Civil Engineering, 9*(1), pp. 17-22.

Panas, A., & Pantouvakis, J. (2014). Simulation-Based and Statistical Analysis of the Learning Effect in Floating Caisson Construction Operations. *Journal of Construction Engineering and Management, 140*(1), 04013033. doi: doi:10.1061/(ASCE)CO.1943-7862.0000782

Roy, P. (1992). Principles of Construction Management: Mc Graw Hill Book Company Europe.

Singh, V., Gu, N., & Wang, X. (2011). A theoretical framework of a BIM-based multi-disciplinary collaboration platform. *Automation in Construction, 20*(2), pp. 134-144. doi: http://dx.doi.org/10.1016/j.autcon.2010.09.011