

## **A Case Study for 3D Modelling Process Analysis based on BIM Log File**

Nan Zhang<sup>1</sup>, Yichen Tian<sup>1</sup>, and Mohamed Al-Hussein<sup>2</sup>

<sup>1</sup> *MSc Student, Department of Civil & Environmental Engineering, University of Alberta*

<sup>2</sup> *Professor, Department of Civil & Environmental Engineering Hole School of Construction Engineering, University of Alberta*

*\*Corresponding author's e-mail: nz@ualberta.ca*

### **ABSTRACT**

Building Information Modeling (BIM) design becomes popular with the growing needs of enhancing construction efficiency. During the design process, 3D Modeling is the foundation of generating and analyzing building information. The modeling tool, Autodesk Revit, automatically creates log files, which can record all design activities. This paper investigates and analyzes the operation data from unstructured temporal Revit log files. These data were recorded when executing the general and external command to build a 3D model of a typical residential two-story building. The objectives of this analysis were to figure out the influence elements of producing a good framing model and to give some suggestions for improving modeling efficiency. In order to extract the useful specific commands, the information in the text journal file was transferred to an excel file, before reading by a data mining tool, RapidMiner. It generated a pie chart to show the repetition amount of modeling activities. Analyzing the reasons of repeating some commands frequently, this paper contributes the possibility to reduce the useless repetition for building a good 3D model.

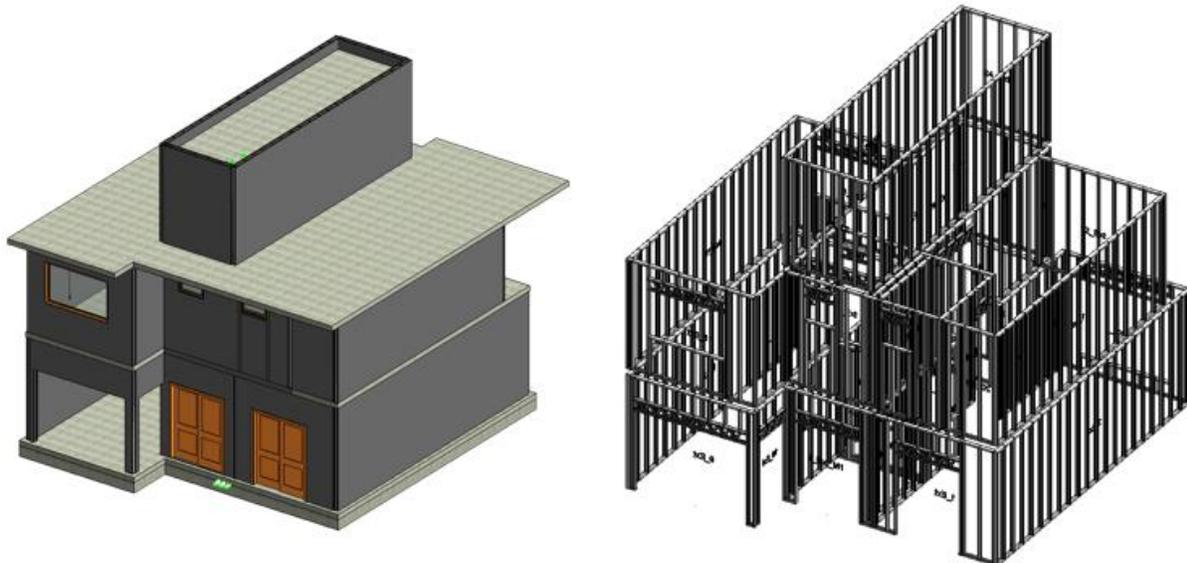
### **KEYWORDS**

Building information modeling; Revit; Log file; Design efficiency improvement

### **INTRODUCTION**

Building information modeling (BIM) is not only a software application but also a process of designing and documenting building information (Aranda-Mena et al., 2009). 3D model is the direct and visible information for architects, engineers and designers. Using BIM for framing and estimating is practical in residential constructions based on the BIM model (Franco et al., 2015). However, the model designed by the architect is hard to adapt to the needs of construction engineers. Taking this case study for an example, the 3D model given by an architect included floors, walls, roofs, windows, and doors. In construction company, the engineers also needed steel framings (Figure 1), which structured the walls and floors. When framing the walls, it took more extra time to prepare the model for running an automatic framing add-on in Revit. Thus, it is necessary to have an effective performance monitoring system (Pilehchian et al., 2015). For Revit, the 3D modeling tool, the design journal files record all the commands of users used. But getting meaningful information in the unstructured and temporal log files is a challenge.

Log files has been used for a wide variety of purposes: estimate the knowledge or expertise level of users (Hammer et al., 1979), record and store performance information efficiently and robustly (Finlayson et al., 1987), make observations to generate usage patterns and operation routines (Guzdial et al., 1993 and 1994), reduce the influence of software system failures (Fronza et al., 2013). Based on the research of using log files, the activity record and analysis is considered as the main information source for estimating users' operation. Thus, the first step is to collect users' commands in log files.

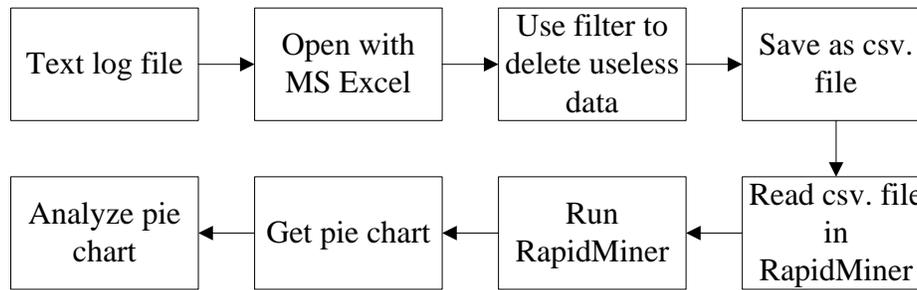


**Figure 1.** This is the 3D model overview, including the architecture and framings in the case study.

This study presents a data mining method for counting the repetition of specific commands. The next section shows the research methodology used to transfer data from text file to Comma Separated Values (CSV) file and to clean useless data. The analysis results and discussions are before the conclusion which is presented in the last section.

## METHODOLOGY

There were 195,351 lines of data in the design journal file of the case study. Thus, it is hard to extract meaningful information in the original text file. In this paper, the analysis process was shown in a flowchart (Figure 2), including eight activities. First, the log file was opened by MS Excel, before using the filter to delete useless data. Then, the excel file was saved as a csv file, which was read and mined in RapidMiner. After generating a pie chart, the commands that the designer used most were clear.



**Figure 2.** This is a flowchart for analyzing the data in BIM design log files.

### Data Collection

The data that were extracted in the temporal log file included the general commands and the command descriptions. The general command always began with “Jrn.Command” (Yarmohammadi et al., 2017), which were followed by “Internal”, “Ribbon”, “AccelKey” and “KeyboardShortcut”. After these keywords, there were specific commands which explained what exacted operations the designer finished. Each kind of general command can be followed by different specific commands. In other words, all specific commands were classified by four types beginning with four kinds of keywords. The example was shown in Table 1.

**Table 1.** This is the example of general commands and specific commands in log files.

	General Command	Specific Command
1	Jrn.Command “Internal”	“Open an existing project”
2	Jrn.Command “Ribbon”	“Create a reference plane”
3	Jrn.Command “AccelKey”	“Cancel the current operation”
4	Jrn.Command “AccelKey”	“Delete the selection”
5	Jrn.Command “KeyboardShortcut”	“Temporary hide selected elements (in the current view)”

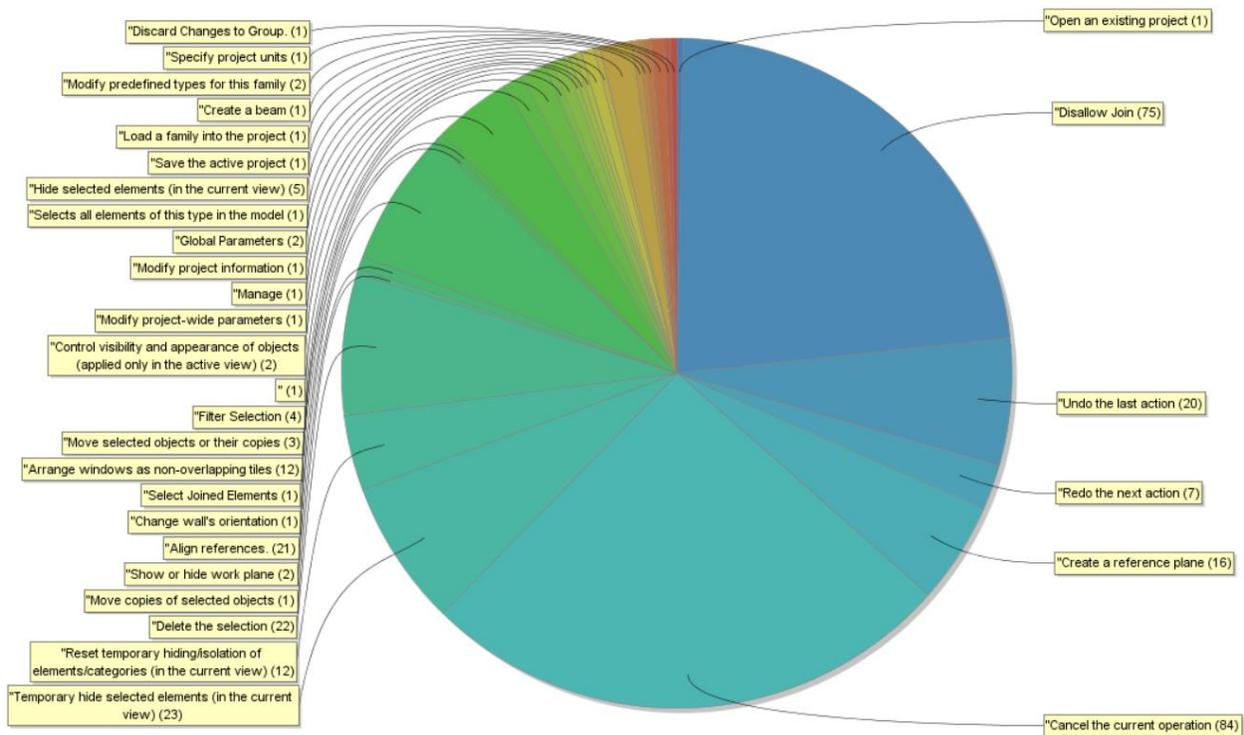
After transferring the data from text file to excel file, there were almost 200 thousand rows. Among these cells, the data starting with general commands or “Jrn.Command”, were meaningful in this case study. As a result, using the filter on the first column to filtrate the commands was inevitable. The useful information was copied to a new sheet with only 326 lines left. These data were what the designer needed in this case study.

### Data Analysis

For data mining and analysis, a data science software platform, named as RapidMiner (Hofmann et al., 2013), was applied. In this data mining tool, the process included one operator, “Read CSV”. Running the process, the plotter in the results window can be set as a pie chart for displaying data. There were 29 specific commands in the statistical graph (Figure 3). The designer repeated two commands frequently, “Cancel the current operation” and “Disallow Join”, with 84 and 75 times respectively. The next five most repetitions were “Align reference”, “Delete the selection”, “Temporary hide selected elements (in the current view)”, “Undo the last action” and “Create a reference plane”, which were all around 20 times. The data of these seven specific commands used frequently can be ranged in a table (Table 2) and calculated the ratios.

## RESULTS AND DISCUSSION

Among these most repetition commands, some were required by the BIM modeling tool, and others were specific in this case study. In Revit, some general operations, such as drawing a reference line, copying a selection, placing a wall etc., should end with one fixed activity, “Cancel the current operation”, by pressing the Esc at the keyboard. This is the reason why the designer repeated operation cancelation frequently. However, disallowing joins continually is not normal when building a 3D model. This action is for disconnected two walls. In this case study, the model was produced by an architect at first. For framing and estimating, the engineer in the construction company separated the walls by penal. Actually, this repetition can be avoided when the architect drew the walls at the beginning. The next five commands the designer used most were helped to modify the model, which was normal in the modeling process. Compared with the top two commands, the rest took small proportions in total operations.



**Figure 3.** This is the pie chart showing the proportions of repeated specific commands.

**Table 2.** This is the table of top seven repeated commands. It shows the repetition times and percentage of each specific operation in total general commands.

	Specific Command	Count	Percentage
1	“Cancel the current operation”	84	25.77%
2	“Disallow Join”	75	23.01%
3	“Temporary hide selected elements (in the current view)”	23	7.06%
4	“Delete the selection”	22	6.75%
5	“Align reference”	21	6.44%
6	“Undo the last action”	20	6.13%
7	“Create a reference plane”	16	4.91%

## CONCLUSION

A good model is crucial for framing and estimating, although it cannot meet all demands of both architects and construction engineers. To notify the specific design at the beginning will be helpful to build a well-managed model. In this case study, identifying the format that stored different information at the first step was an essential contribution. For extracting the data, MS Excel was a reasonable tool to clean and reorganized by deleting non-value information. In the next step, the data mining tool accepted and read the csv file which was produced from the excel file. At last, the pie chart from data mining results showed the significant difference of repeating commands. There is a limitation in this case study. The data quality is limited by the temporal log file. This file is generated and deleted by Revit implicit setting during the modeling process. In the future, the Revit log file can be used as a rich resource to get performance variations in different models and projects, if it is possible to keep or retrieve the log files.

## REFERENCES

- Aranda-Mena, G., Crawford, J., Chevez, A., & Froese, T. (2009). "Building information modelling demystified: does it make business sense to adopt BIM?" *International Journal of managing projects in business*, 2(3), 419-434.
- Finlayson, R., & Cheriton, D. (1987). "Log files: an extended file service exploiting write-once storage." *ACM*, Vol. 21, No. 5, pp. 139-148.
- Franco, J., Mahdi, F., & Abaza, H. (2015). "Using building information modeling (BIM) for estimating and scheduling, adoption barriers." *Universal Journal of Management*, 3(9), 376-384.
- Fronza, I., Sillitti, A., Succi, G., Terho, M., & Vlasenko, J. (2013). Failure prediction based on log files using random indexing and support vector machines. *Journal of Systems and Software*, 86(1), 2-11.
- Guzdial, M. (1993). "Deriving software usage patterns from log files." Georgia Institute of Technology.
- Guzdial, M., Santos, P. J., Badre, A., Hudson, S. E., & Gray, M. H. (1994). "Analyzing and visualizing log files: A computational science of usability." Georgia Institute of Technology.
- Hammer, J. M., & Rouse, W. B. (1979). "Analysis and modeling of freeform text editing behavior." *Proceedings of the 1979 International Conference on Cybernetics and Society*, pp. 659-664.
- Hofmann, M., & Klinkenberg, R. (Eds.). (2013). "RapidMiner: Data mining use cases and business analytics applications." CRC Press.
- Pilehchian, B., Staub-French, S., & Nepal, M. P. (2015). "A conceptual approach to track design changes within a multi-disciplinary building information modeling environment." *Canadian Journal of Civil Engineering*, 42(2), 139-152.
- Yarmohammadi, S., Pourabolghasem, R., & Castro-Lacouture, D. (2017). "Mining implicit 3D modeling patterns from unstructured temporal BIM log text data." *Automation in Construction*, 81, 17-24.