Iyer, S. et al. (Eds.) (2022). Proceedings of the 30th International Conference on Computers in Education.
Asia-Pacific Society for Computers in Education

Modeling Off-task Behavior of Learners Using Minecraft

Maricel A. ESCLAMADO [a,b*]
[a]*Ateneo de Manila University*
[b]*University of Science and Technology of Southern Philippines*
*maricel.esclamado@obf.ateneo.edu

Abstract: This study presents an attempt to develop a detector of off-task behavior using the interaction logs of the learners in Minecraft. Text replays will be used to obtain ground truth labels of off-task behavior. Detectors will be built using a Latent Response Model, which predicts whether an action or event is off-task behavior and assess how frequently each of the students is off-task. The relationship between off-task behavior and students' learning outcome will also be examined. After developing the detector, a WHIMC module implementation will be conducted to apply intervention to keep the students from off-task behavior and to investigate how the learning outcomes will change after the intervention.

Keywords: Minecraft, WHIMC, Off-task Behavior, Intervention, Philippines

1. Introduction

Off-task behavior, in which a learner disengages from the learning environment or task to engage in an unrelated behavior, may affect students' learning (Baker, 2007). In a classroom setting, talking to other students about unrelated subjects and doing activities unrelated to the tasks are examples of off-task behaviors. Detecting this behavior in a classroom setting usually requires the use of sensors such as eye trackers and microphones (Baker 2007). Also, field observations such as the Baker Rodrigo Ocumpaugh Monitoring Protocol (BROMP) can also be used as a method to detect off-task behavior (Botelho, Baker, & Heffernan, 2019). Sometimes, sensors are not available and conducting field observations is not possible. Thus, interaction logs of the students within the learning environment are used to detect patterns of behavior (Kang, Liu, & Qu, 2017). However, detecting off-task behavior based solely on interaction logs is limited and is quite challenging since it varies in different learning environments (Carpenter et al, 2020).

This behavior has been associated with poor learning outcomes but may also aid students in regulating negative affect and returning to the learning task (Sabourin & Lester, 2014). With its impact on learning, there has been a growing interest to detect off-task behavior and examine how it affects learners' performance. Previous studies investigated off-task behavior in intelligent tutoring systems such as ASSISTments tutoring system (Botelho, Baker, & Heffernan, 2019; Pardos et al, 2014; Wixon et al, 2012), Cognitive tutor software (Baker, 2007) and in game-based learning environments such as Crystal Island (Carpenter et al, 2020; Sabourin & Lester, 2014).

The focus of this study is on off-task behavior in Minecraft, an open-ended sandbox video game. Open-ended games provide players autonomy on how to do the tasks in the learning environment, thus there is no prescribed learning sequence and analysis of behavior becomes significantly important (Käser, & Schwartz, 2020). With the open-ended nature of Minecraft, it is challenging to detect behavior patterns such as off-task behavior.

In this study, a detector of off-task behavior will be developed using the interaction logs of the learners in Minecraft. The relationship between off-task behavior and students' learning outcome will also be examined. This study would like to answer the following research questions:

RQ1: What features predict off-task behavior?
RQ2: How do off-task behaviors relate to the student's learning outcome?
RQ3: How will the learning outcomes of the students change after the intervention?

2. Review of Related Literature

2.1 Off-task Behavior in educational games

Games used for education can act as rich primers for engaged learning and were reported effective in improving learning outcomes (Zhonggen, 2019). The use of games in education attempt to form positive mood leading to increased interest in gameplay and better learning performance. The interaction logs of the learners within these games can be collected and analyzed to examine their behavior patterns and how it affects their learning (Kang, Liu, & Qu, 2017). One of the behaviors that may affect learning is off-task behavior, which is disengagement from the learning environment or tasks.

Previous studies investigated off-task behavior in different learning environments. A study (Botelho, Baker, & Heffernan, 2019) examined the data collected within the ASSISTments, a webbased tutoring system, to detect off-task behavior and affect and used quantitative field observations following the BROMP protocol for the ground truth labels. Both expert-engineered and machinedlearned features were generated for off-task behavior and affect using the log data and applied Naïve Bayes classifier, a REP tree classifier, and Long-Short Term Memory (LSTM). Another study (Pardos et al, 2014), that investigated off-task behavior in ASSISTments tutoring system, used an observation protocol to code affect and behavior. The off-task detector included the total number of attempts, time taken by a student to answer, if a student has a correct action, the average number of scaffolds, and the total number of incorrect actions in the past. They attempt to fit detectors using eight common classification algorithms including J48 decision trees, step regression, JRip, Naïve Bayes, K*, and REPTrees. The best detector of off-task behavior was found using the REP-Tree algorithm. Another study (Wixon et al, 2012), that investigated data from ASSISTments environment, presented a detector of the WTF "Without Thinking Fastidiously" behavior, in which the students were interacting with the software, but their actions appear to have no relationship to the intended learning task, which is related to off-task behavior. Text replays, a pre-specified chunk of student actions presented in text (Sao Pedro et al, 2010), were used to obtain ground truth labels of WTF behavior. Indicators of WTF include running the same trial a large number of times, pausing the simulation a large number of times, and changing variables many times without stopping to think before running the simulation. The detectors were fit using 11 common classification algorithms and the PART algorithm achieved the best performance.

Another study (Carpenter et al, 2020) used dialogue data collected from a group chat feature integrated into Crystal Island, a collaborative game-based learning environment for science, to investigate off-task behavior. Data was labeled by researchers using a rubric that was developed to identify on-task and off-task chat messages. The features used were the number of times the student had previously contributed to the group conversation, a score representing the polarity of the message's sentiment, the number of characters in the message, and the Jaccard similarity of the message with the game's text content, and the average word embedding for the message. They used logistic regression to perform binary classification of the messages. They also used an LSTM-based sequential model to analyze the sequences of data. The performance of logistic regression and LSTM were evaluated using accuracy, precision, and F1. Based on the results, sequential techniques for modeling off-task behavior outperform static techniques.

2.2 What-If Hypothetical Implementations in Minecraft (WHIMC)

What-If Hypothetical Implementations in Minecraft (WHIMC) is a set of simulations that learners can explore in order to learn more about science, technology, engineering, and mathematics (STEM). The alternate versions of Earth present learners with opportunities to observe the planet under altered conditions. Although the worlds are fictional, they are created in consultation with scientists: They accurately depict conditions on Earth under these circumstances.

In each of these alternate Earths and exoplanets, learners explore the terrain, describe the environment, report observations about how life on Earth is affected by these circumstances, and possibly create habitats that will enable them to survive. By immersing learners in these activities, WHIMC hopes to generate interest in and excitement for STEM among participating learners.

Previous studies analyzed the dataset from the module implementations in WHIMC which includes both in-game and out-of-game data. A previous study in WHIMC examined the relationship of features extracted from location data such as area, distance traveled, and MSI to assessment outcomes (Esclamado & Rodrigo, 2022). A study also analyzed and compared American and Filipino learner traversals and in-game observations against canonical answers from experts to determine the extent to which students achieved the desired learning outcomes (Casano & Rodrigo, 2022).

3. Contribution of the Proposed Research

This study contributes to the body of knowledge in 4 ways:

First, off-task behavior in Minecraft might be more difficult to detect because of the nature of the game. Open-ended games like Minecraft provide players autonomy on how to do the tasks in the learning environment, thus there is no prescribed learning sequence and analysis of behavior becomes significantly important (Käser, & Schwartz, 2020). With the open-ended nature of Minecraft, it is challenging to detect behavior patterns such as off-task behavior. In this study, the features that would predict off-task behavior will be identified.

Second, this will contribute to the literature on how off-task behaviors affect student's learning outcomes. This behavior has been associated with poor learning outcomes (Sabourin & Lester, 2014).

Third, since this study will also examine how off-task behavior relates to learning outcomes, this may provide an opportunity for a real-time intervention to guide the learners to better learning outcomes (Baker & Clarke-Midura, 2013).

Lastly, this will also contribute to the modeling of users in the Philippines, who tend to be underrepresented, since most studies are data from western learners.

4. Proposed Research Methodology

4.1 Data Collection

The data, consisting of in-game data and assessment data of WHIMC module implementations of partner schools in the Philippines, will be analyzed in this study. Partner teachers from the schools developed learning modules to explore the WHIMC worlds. The students then explore the WHIMC worlds based on the learning modules. After the exploration, the learners answer the knowledge assessments.

4.2 Off-task Behavior Detector

Log files will be converted to text replays, easy-to-read versions of the log files and are effective for providing ground truth labels for behaviors (Wixon et al., 2012). Human coders will examine the text replays and determine whether the actions or events are off-task behaviors. Using text replays, ground truth labels of off-task behavior for each action or event can be obtained.

Indicators of off-task behaviors in WHIMC are identified which include visiting other worlds when the task is not yet completed, not visiting the assigned worlds, visiting assigned worlds that need exploring for a very short time and/or only covering a very small area, not making observations, and not accessing the science tools especially if this is part of the learning task. Based on these indicators, features of off-task behaviors will be extracted from the log files.

RQ1: What features predict off-task behavior?

Detectors will be built using a Latent Response Model (LRM), which predicts whether an action is an off-task behavior and assess how frequently each of the students is off-task (Baker, 2007). Fast Correlation-Based Filtering and Forward Selection will be used to select the features and the model.

Detectors will be assessed using metrics such as A', correlation, and kappa.

RQ2: How do off-task behaviors relate to the student's learning outcome?
Correlation analysis will be conducted between the number of incidences of off-task behaviors and the outcome assessment of the students to determine the effect of this behavior on learning outcomes.

RQ3: How will the learning outcomes of the students change after the intervention?
Another WHIMC module implementation will be conducted to apply intervention to keep the students from off-task behavior. There will be two learning modules that will be developed. In the first learning module, the learners will explore WHIMC based on the learning module and there will be no intervention yet. In the second learning module, there will be intervention to keep the students from off-task behavior. At the end of each module, there will be a knowledge assessment. The changes in the performance of the students in Modules 1 and 2 will be examined.

Acknowledgements

The author thanks H Chad Lane and Jeff Ginger for their enthusiastic collaboration, Dominique Marie Antoinette B. Manahan, Jonathan D. Casano, and Ma. Rosario Madjos for their support, the Ateneo Laboratory for the Learning Sciences, the Ateneo de Manila University, the Department of Science and Technology (DOST) as the funding agency, and the Department of Science and Technology's Philippine Council for Industry, Energy, and Emerging Technology Research and Development (DOST-PCIEERD) as the monitoring agency for the grant entitled, "Nurturing Interest in STEM among Filipino learners using Minecraft."

References

Baker, R. S. J. (2007). Modeling and understanding students' off-task behavior in intelligent tutoring systems. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.

Baker, R. S., & Clarke-Midura, J. (2013, June). Predicting successful inquiry learning in a virtual performance assessment for science. In *International conference on user modeling, adaptation, and personalization* (pp. 203-214). Springer, Berlin, Heidelberg.

Botelho, A. F., Baker, R. S., & Heffernan, N. T. (2019, July). Machine-learned or expert-engineered features? Exploring feature engineering methods in detectors of student behavior and affect. In *The twelfth international conference on educational data mining*.

Carpenter, D., Emerson, A., Mott, B. W., Saleh, A., Glazewski, K. D., Hmelo-Silver, C. E., & Lester, J. C. (2020, July). Detecting off-task behavior from student dialogue in game-based collaborative learning. In *International Conference on Artificial Intelligence in Education* (pp. 55-66). Springer, Cham.

Casano, J. D., & Rodrigo, M. M. T. (2022). A Comparative Assessment of US and PH Learner Traversals and In-Game Observations within Minecraft. In *International Conference on Artificial Intelligence in Education* (pp. 267-270). Springer, Cham.

Esclamado, M. A., & Rodrigo, M. M. T. (2022). Are All Who Wander Lost? An Exploratory Analysis of Learner Traversals of Minecraft Worlds. In *International Conference on Artificial Intelligence in Education* (pp. 263-266). Springer, Cham.

Kang, J., Liu, M., & Qu, W. (2017). Using gameplay data to examine learning behavior patterns in a serious game. *Computers in Human Behavior*, *72*, 757-770.

Pardos, Z. A., Baker, R. S. J. D., San Pedro, M., Gowda, S. M., & Gowda, S. M. (2014). Affective states and state tests: Investigating how affect and engagement during the school year predict end-of-year learning outcomes. *Journal of Learning Analytics*, *1*(1), 107–128. https://doi.org/10.18608/jla.2014.11.6

Sabourin, J. L., & Lester, J. C. (2014). Affect and engagement in game-based learning environments. *IEEE Transactions on Affective Computing*, *5*(1), 45–56. https://doi.org/10.1109/t-affc.2013.27

Sao Pedro, M., Baker, R. S., Montalvo, O., Nakama, A., & Gobert, J. D. (2010, June). Using text replay tagging to produce detectors of systematic experimentation behavior patterns. In *Educational Data Mining 2010*.

Wixon, M., d Baker, R. S., Gobert, J. D., Ocumpaugh, J., & Bachmann, M. (2012, July). WTF? detecting students who are conducting inquiry without thinking fastidiously. In *International Conference on User Modeling, Adaptation, and Personalization* (pp. 286-296). Springer, Berlin, Heidelberg.

Zhonggen, Y. (2019). A meta-analysis of use of serious games in education over a decade. *International Journal of Computer Games Technology*, 2019, 1–8. https://doi.org/10.1155/2019/4797032

WHIMC. (n.d.) What-If Hypothetical Implementations using Minecraft. Available online: https://whimcproject.web.illinois.edu/