Ahmed Tlili · Maiga Chang
Editors

Data Analytics Approaches in Educational Games and Gamification Systems

Springer

Editors
Ahmed Tlili
Smart Learning Institute
Beijing Normal University
Beijing, China

Maiga Chang
School of Computing
and Information Systems
Athabasca University
Edmonton, AB, Canada

ISSN 2522-0888 ISSN 2522-0896 (electronic)
Smart Computing and Intelligence
ISBN 978-981-32-9334-2 ISBN 978-981-32-9335-9 (eBook)
https://doi.org/10.1007/978-981-32-9335-9

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Chapter 13
Data Analytics Approaches in Educational Games and Gamification Systems: Summary, Challenges, and Future Insights

Ahmed Tlili and Maiga Chang

Abstract This chapter summarizes the reported findings of this book to facilitate the adoption of data analytics in educational games and gamification systems. Specifically, this chapter presents the objectives of adopting data analytics which is finding individual differences; doing learning assessments and knowing more about the learners. It then presents the collected metrics and applied analytics techniques in order to achieve these objectives. Additionally, this chapter highlights several limitations reported by other authors during the adoption of learning analytics. These limitations should be considered by researchers and practitioners in their context to facilitate learning analytics adoption. Finally, this chapter provides future insights about the learning analytics field.

1 Objectives of Adopting Data Analytics

The inclusion of data analytics within educational games and gamification systems can make them smart by achieving several objectives, highlighted in this book, as follows:

- **Finding individual differences**: Traditional learner modeling instruments, such as questionnaires, have been reported to be lengthy and not motivating. With the help of data analytics, a system includes educational game is capable of modeling learners implicitly. Furthermore, individual differences like competences (computational thinking in particular) and motivation can also be found (see Chaps. 11 and 12 for more details). The modeling process is considered a crucial step in order to provide personalized and adaptive learning services based on individual differences that are further highlighted in this chapter).

A. Tlili (✉)
Smart Learning Institute of Beijing Normal University, Beijing, China
e-mail: ahmed.tlili23@yahoo.com

M. Chang
School of Computing and Information Systems, Athabasca University, Athabasca, Canada
e-mail: maiga.chang@gmail.com

A. Tlili and M. Chang (eds.), *Data Analytics Approaches in Educational Games and Gamification Systems*, Smart Computing and Intelligence,
https://doi.org/10.1007/978-981-32-9335-9_13

477240_1_En_13_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:22/8/2019 Pages: 255 Layout: T1-Standard

- **Do learning assessments**: In online learning environments like MOOCs, massively multiplayer educational games, where thousands of learners are taking their pace for studying and doing learning activities. It becomes very difficult for teachers to monitor learners' learning progress, assess learners' skills and knowledge levels, and take care of individual needs. Data analytics have been applied to solve this issue and help teachers. For instance, Tlili and his colleagues develop iMoodle that can identify at-risk students and provide them personalized learning support (see Chap. 6). With data analytics' help, the completion rate of a course can also be improved. Seaton and her colleagues provide learning analytics dashboard for the learners so they can easily keep track of their progress, realize their habits and weaknesses to help them overcome obstacles and achieve better learning outcome (see Chap. 7).
- **Know more about the learners**: Traditional educational games and gamification systems are black boxes where teachers cannot see or know, besides the final scores and levels cleared, how their learners do in the learning process and behave towards the learning goal. Data analytics approaches have been applied to overcome this limitation. Ifenthaler and Gibson explore the learning engagement and its relationship with learning performance in the context of game-based learning (see Chap. 3). Also, Shute, Rahimi, and Smith explore the importance of including learning supports and its impact on learning performance when using the Physics Playground game (see Chap. 4).

Based on the reported chapters in this book, Fig. 1 presents a generic framework of adopting data analytics in educational games and gamification systems to achieve the three objectives mentioned above. When learners use and interact with the developed educational game or gamified system, several metrics (traces) are created based on the interactions and collected into the database. The data analytics module(s) is (are) developed either as the built-in the game and system or accessories of the game and system. The module takes the collected metrics as inputs and does proper analysis and produce results as outputs for achieving a particular objective.

It should be noted that no chapter reports the use of cloud computing technology to store the collected metrics. Also, no chapter has parents, as stakeholders, for the application of data analytics in educational games and gamification systems. Therefore, further investigation is needed regarding these two matters.

2 Collectable Metrics and Traces

It has been seen that the more data is collected within educational games and gamification systems, the more possibilities we will have to enhance the learning process. Kinshuk et al. [1] highlight that to provide smart learning every bit of information that each learner comes into contact with should be collected. For example, to predict at-risk students, Tlili and his colleagues in Chap. 6 use the following five metrics, namely: (1) Number of acquired badges which highlights the number of conducted

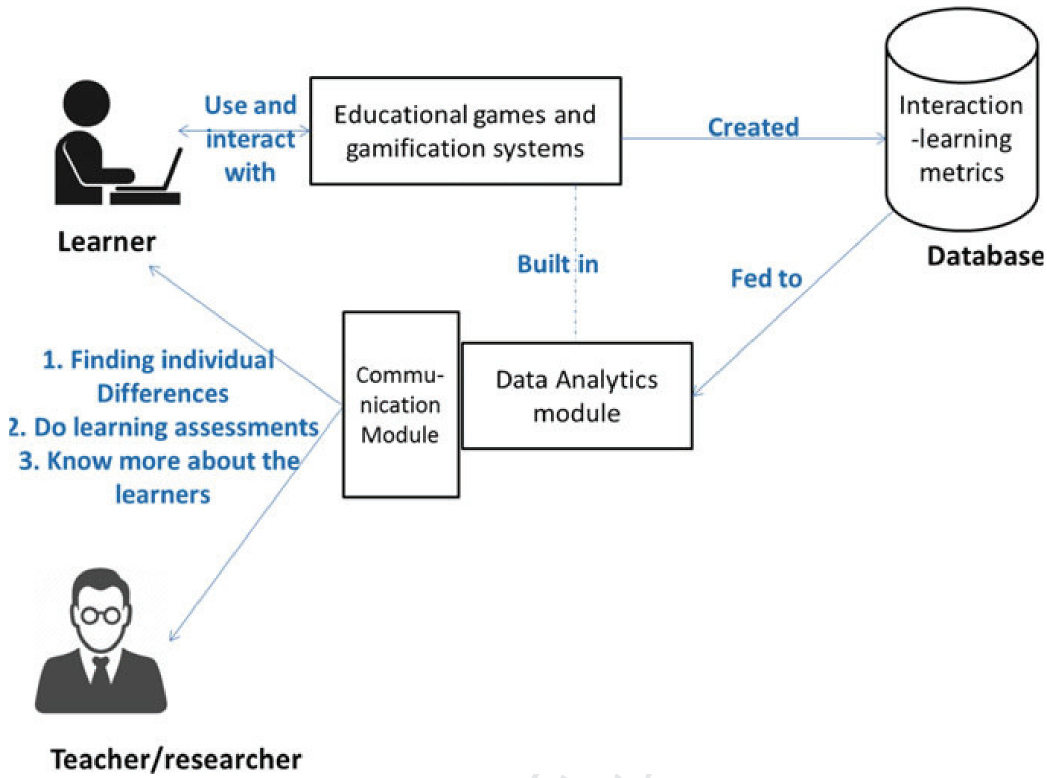

Fig. 1 Generic process of adopting data analytics in educational games and gamification systems

learning activities, since every time a student finishes a learning activity, he/she gets a badge; (2) Activities grades which refer to the value assigned by teachers to assignments and quizzes requested and delivered by students; (3) Student's rank on the leaderboard which is based on the acquired number of points (4) Course progress which can be seen in the progress bar; and, (5) Forum and chat interactions which refers to students' participation in online discussions, such as the number of posts read, posts created and replies.

To identify the motivation of students in an educational game, Flores, Silverio, Feria, and Cariaga use in Chap. 12 the following five metrics, namely: (1) difficulty versus accuracy: compared to assess students' behavior; the student's choice of difficulty based on their result in the previous problem (correct, wrong or skip); (2) number of non-easy problems chosen: the total number of selected medium, hard and expert difficulty problems; (3) number of non-skipped problems: measured to give students a reasonable score for this metric as skipping is generally considered a negative factor; (4) accuracy versus time: were also compared to identify students who only guess the answers; and, (5) perks versus accuracy: were compared to examine students' engagement or mastery in solving a problem.

To assess computational thinking skills, Montaño, Mondragón, Tobar-Muñoz, and Orozco use in Chap. 5 the following seven metrics, namely, (1) abstraction and pattern recognition: focus on not having unused code, the use of functions in the code, and the use of clones of blocks of code (a specific functionality of the Scratch environment);

(2) flow control: assessment of the correct use of every control instruction (such as if and for statements), and also the adequate use in nesting those statements; (3) input control: assessment of the adequate use of statements designed to capture user input into the code, the naming of variables, and the use of non-user-defined variables; (4) data analysis: assessment of the treatment and transformation of the data through the use of data transformation blocks or statements, and also their adequate nesting if necessary; (5) parallelism and threading: assessment of the adequate use of threading and multi-tasking enabling blocks; (6) problem-solving: assessment of the student's ability to decompose a problem into multiple smaller ones in order to address them more easily; and, (7) algorithmic thinking: assessment of the student's ability to develop sequences of tasks, that would be translated into blocks of code, in order to solve a problem.

While Ghergulescu and Muntean [2] mentioned that little is known about the collected traces and used metrics in game-based learning environments, it has been seen that different types of metrics could be collected by asking three questions as Fig. 2 shows:

(1) *What types of metrics should be collected?* Two types of metrics (traces) can be collected, namely, (1) generic metrics which can be found in most educational games and gamification systems, such as the number of signing into an educational game or gamification system, and the time spent on the game or system; and, (2) specific metrics which are defined based on the designed learning envi-

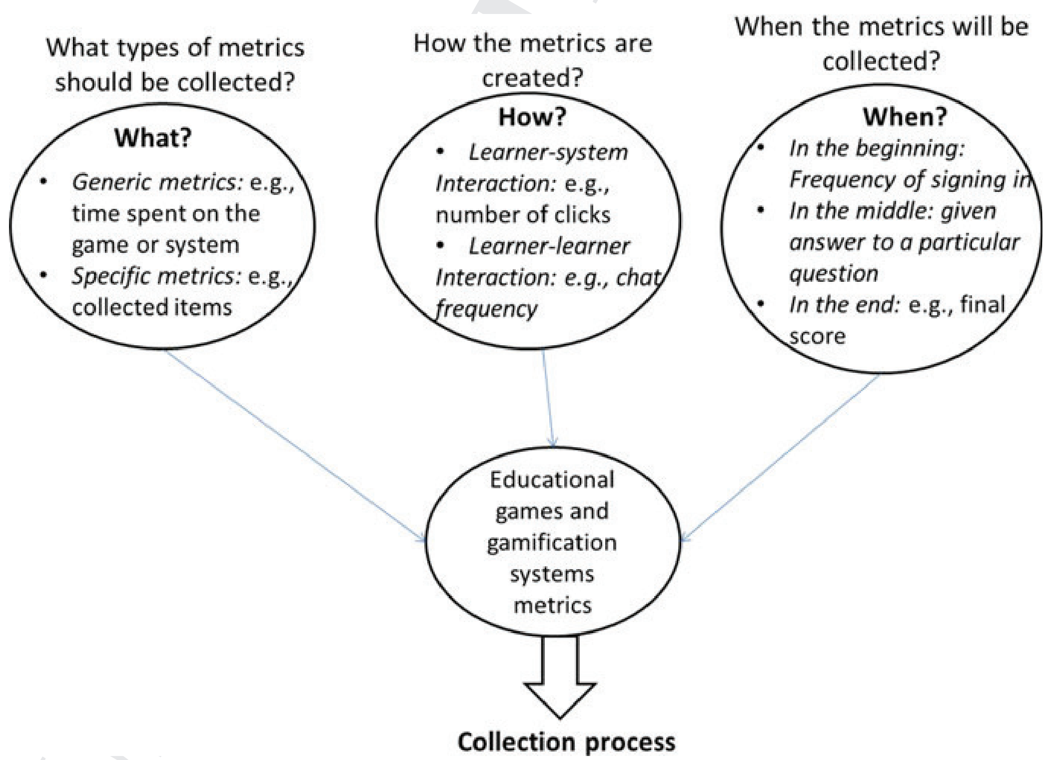

Fig. 2 Types of collected metrics in educational games and gamification systems

ronment, such as number of collected items (badges, points, coins, etc.) and the selected game path.

(2) *How the metrics are created?* One kind of metrics (traces) is created when the learners interact with the educational game or the gamification system, while the others can be created when the learners are interacting with others within the game or the system. For instance, chat frequency is created when learners start to chat together within the game or the system to solve a particular learning activity.

(3) *When the metrics will be collected?* Metrics (traces) can be collected at the beginning, in the middle or at the end of the game-play or the usage of the gamification system. For instance, the final score is collected at the end while the number of times to sign in the game or the system is collected at the beginning.

To extract useful information from the collected metrics (discussed above), different analytics techniques are applied, as discussed in the next section.

3 Analytics Techniques

Based on the chapters included in this book, three analytics techniques are usually adopted in educational games and gamification systems:

- *Data Visualization*: It uses visualization such as pie charts and histograms to represent data. This can help to communicate information clearly and efficiently to stakeholders (e.g., teachers, students, etc.). For instance, the authors in Chaps. 6 and 7 all adopt data visualization techniques to create dashboards for both teachers and students.
- *Data Mining*: It aims to discover hidden information and meaningful patterns from massive data. In this context, several algorithms are adopted. For example, Tlili and his colleagues in Chap. 6 adopt association rules mining and Apriori Algorithm to predict at-risk students.
- *Sequential Analysis*: It allows exploring, summarizing, and statistically test cross-dependencies between behaviors that occur in interactive sequences. For example, Moon and Liu in Chap. 2 conduct a systematic literature review on 102 articles that work on sequential data analytics (SDA) in game-based learning.

Several studies also reported the abovementioned analytics techniques are commonly adopted in games [3–5]. Several challenges, on the other hand, are reported by the authors, in their chapters, which might hinder the adoption of data analytics in educational games and gamification systems. These challenges are discussed in the next section.

4 Challenges

Based on the chapters included in this book, several challenges are reported by the authors. These challenges should be considered by researchers and practitioners in their context to enhance the adoption of data analytics in educational games and gamification systems.

Moon and Liu in Chap. 2 highlight two limitations while adopting data analytics approaches, specifically sequential analysis, in educational games and gamification systems, namely: (1) the need for high computational power in order to collect and analyze big data; and, (2) sequential analysis is often performed as post hoc analysis. Therefore, it is challenging to ensure the validity of the results without cross-validating with the participants. In addition, the participants may not even recall some certain behaviors because the data is captured at a fine granularity. Another issue with post hoc analysis is if the scope of the study is biased, data collection will be biased which in turn leads to an invalidated biased results.

Ifenthaler and Gibson in Chap. 3 and Montaño, Mondragón, Tobar-Muñoz, and Orozco in Chap. 5 report that one of the challenges is collecting large enough data so the applied data analytics approach within their gamified systems can be more accurate. Tlili and his colleagues in Chap. 6 highlight the challenge of protecting learners' privacy while applying educational games and gamification systems. They also discuss the importance of having a predefined time of keeping the learners' stored data.

5 Conclusion

Game-based learning environments and learning analytics are gaining increasing attention from researchers and educators since they both can enhance learning outcomes. Therefore, this book covered a hot topic which is the application of data analytics approaches and research on human behavior analysis in game-based learning environments, namely educational games and gamification systems, to provide smart learning. Specifically, this book discussed the purposes, advantages, and limitations of applying theses analytics approaches in these environments. Additionally, this book helped readers, through various presented smart game-based learning environments, integrate learning analytics in their educational games and gamification systems to, for instance, assess and model students (e.g., their computational thinking) or enhance the learning process for better outcomes. Finally, this book presented general guidelines, from different perspectives, namely, collected metrics, applied algorithms and the encountered challenges during the application of data analytics approaches, which facilitate incorporating learning analytics in educational games and gamification systems.

Future directions for readers to consider and focus could be: (1) investigating the use of data analytics in educational games and gamification systems for health education in particular; and, (2) investigating how Internet of Things (IoT), which is a new technology that is gaining increasing attention from researchers and practitioners, could help the application of data analytics in educational games and gamification systems.

References

1. Kinshuk., Chen, N. S., Cheng, I. L., & Chew, S. W. (2016). Evolution is not enough: Revolutionizing current learning environments to smart learning environments. *International Journal of Artificial Intelligence in Education, 26*(2), 561–581.
2. Ghergulescu, I., & Muntean, C. H. (2012). Measurement and analysis of learner's motivation in game-based e-learning. In *Assessment in game-based learning* (pp. 355–378). New York: Springer.
3. Chaudy, Y., Connolly, T., & Hainey, T. (2014). Learning analytics in serious games: A review of the literature. In *European Conference in the Applications of Enabling Technologies (ECAET)*, Glasgow, (2014).
4. Siemens, G. (2010). *What are learning analytics*? Retrieved August 12, 2016, from http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics/.
5. Kim, S., Song, K., Lockee, B., & Burton, J. (2018). What is gamification in learning and education? In *Gamification in learning and education* (pp. 25–38). Cham: Springer.