Modeling and Analyzing Gamification Behavior with Badges

Author: José A. Ruipérez-Valiente
Supervisor: Pedro J. Muñoz-Merino

Abstract—This manuscript is focused in two of the most prominent techniques for the next years for education advancement, which are learning analytics and gamification. There are a lot of research works in both of these lines independently, but not many research papers combine both techniques in order to improve the educational experience. One of the most used gamification techniques is the use of badges as a reward for making specific student actions. The analysis of users' interactions and behaviors with the badge system can be used to improve the learning process. We present four high level indicators related to the behavior of students with a badge system and we particularize it for the Khan Academy platform. An extensive analysis of 291 different students interacting with the Khan Academy badge system is done processing real data from freshmen courses at Universidad Carlos III de Madrid. This analysis includes an overview of the global usage of badges, correlations between the badge indicators and other indicators related to the learning process and a Two-Step Cluster Analysis to group students by their badge preferences in order to personalize future experiences.

Index Terms— Data mining, decision support, distance learning, education, games, human factors

1 Introduction

The existence of low level data about students' events when interacting with different educational platforms and services broadens the possibilities of research. Learning analytics emerges as the science that analyzes these educational data in order to obtain useful information that can be used to improve the learning process [1]. On the other hand, gamification techniques have been used for a long time ago, but it was recently when the term was coined and become famous when some important players of the industry started to apply it. Gamification can be defined as the use of game design elements in a non-game context [2].

The K-12 Horizon Report 2014 [3] identifies both learning analytics and gamification as important developments in technology for education with a time of adoption from 2 to 3 years. The main reason to implement gamification elements in education is to improve the motivation and engagement of the student towards the learning goals. The scope and uses of learning analytics are very broad, for example educators can benefit from such things as identifying problems and analyzing student and class data in-depth whereas students can gain much self-awareness and compare activity or participation with other peers.

The application of learning analytics for gamified environments in education can bring useful information for the learning process such as students' behavior through gamification, total use of the different functionality or course differences with gamification elements. Different learning analytics research works offer indicators related to the learning process, such as number of accesses to each resource or number of correct exercises. However, indicators related to gamification in education are not provided [4] and there are not many researches that focus on learning analytics for gamification elements.

One of the most used gamification techniques in education is the provision of badges. Badges are awards that students can receive for doing specific actions on an educational platform (e.g. for watching videos during some time or for solving a set of exercises correctly). The Innovating Pedagogy report 2013 [5] presents badges to accredit learning with a high potential impact and learning analytics with a medium/high impact, and timescales for adoption for both of them from 2 to 5 years.

The knowledge of the students' behaviors and interactions with badges is very important. For example, if we can infer which students make actions with the intention to acquire badges, then we can know the students who are motivated by achievements and we can apply specific learning techniques for these students. In addition, if we know the badges that are more used and compare them with others, then we can make hypothesis about which badges might motivate more students or which of them have requirements too demanding; then we can make the proper actions to improve the learning process.

In this study, we focus on analyzing students' interactions and behavior with badges in an experience with 291 students using the Khan Academy platform at Universidad Carlos III de Madrid. The specific objectives of this research are the following:

- Analyze the general achievement of badges by students, including the comparison of badges of different types and the resources that triggered such type of badges
- Propose indicators for analyzing the students' intention for the achievement of badges, taking into account two different types of badges: topic and repetitive ones
- Propose indicators for analyzing the students' concentration on the achievement of badges (i.e. if students' actions are consecutive for the achievement of badges without interruption) and the students' efficiency for the achievement of badges
- Analyze the students' intention, concentration and efficiency for the achievement of badges, including clustering of students to group students with same type of behavior so that personalized learning can be provided to them

- Analyze the relationship among students' intention, concentration and efficiency for the achievement of badges with other typical indicators of students' performance and behavior. The objective here is to establish relationships between gamification behaviors and other relevant indicators of the course.

The remainder of the paper is organized as follows. Section 2 presents related works about gamification and learning analytics, where few of them combine both lines of research. Section 3 reviews the gamification elements which are present in the Khan Academy platform by default whereas in section 4 we propose some indicators which provide higher level information about the attitude of students towards badges (intention, concentration, efficiency). Section 5 makes an extensive analysis of the interaction of students with badges using data from three pre-graduate courses using Khan Academy at Universidad Carlos III de Madrid. Section 6 finishes with conclusions and future work.

2 Related Work

Games are widely spread among a complete whole range of different contexts, which encourages the use of them for educational purposes in order to provide a more immersive process and improve the engagement [6]. There are different initiatives that attempt to introduce educational games as a path to improve the learning process. As an example the work developed by Freire [7] presents the introduction of serious games in the MOOC platform edX, which is one of the most used worldwide. Serious games have been proven to be a successful effort in order to implement educational immersive experiences to improve engagement of students. Therefore, many works apply some of the common elements from games in non-game contexts by using gamification techniques. These techniques have been used in many different contexts; for example the introduction of game achievements in a photo sharing service [8] which concluded that these elements did not play a key role in the system or the inclusion of gamification elements or the impact of gamification elements in eco-driving [9] showing a positive correlation with the use of the proposed eco-driving tips. Gamification has been tested in different e-learning experiences similar to the case study hereby presented, for example a gamified AutoCAD tutorial [10] reported that the system improved the subjective engagement levels and a provided a higher completion ratio. Another e-learning experience using gamification [11] has been implemented as a plugin for the Blackboard e-learning platform, with the objective of increasing students' motivation towards completing exercises by introducing game elements such as trophies, achievements or leaderboards. The results [11] reported were not very encouraging as many students informed that the system was not motivating enough and it was not fun to compete with other mates for a place in the leaderboard. The reported results from gamification studies are very diverse and cannot be generalized for all experiences. Hamari, Koivisto and Sarsa [12] presented a literature review of the empirical studies on gamification analyzing 24 research works. The results indicated that gamification yielded positive effects as a general rule but these effects were strongly dependent on the contexts and the users of the experience.

Learning analytics has also emerged as a strong field of research in order to improve the learning process. However still most of the Virtual Learning Environments (VLEs) do not make use of the educational data available generated by students and more research is needed. In this direction there are some research works that have developed learning analytics tools to improve the learning analytics support that VLEs have by default. As an example ALAS-KA [13] is a plugin for the Khan Academy platform that provides a set of more than 21 new different indicators related to the learning process; all the indicators have been grouped in 5 different modules of knowledge that can be used by instructors and students. Another example is Moodog [14] which is a plugin for Moodle Course Management System (CMS) which uses the data generated by the CMS to provide new statistics which have been categorized into 4 different groups. Most of the works in the literature provide indicators which are very basic, such as number of accessed resources or score in assignments, but in this work we propose some higher level indicators that infer knowledge about more complex behaviors related to gamification. Other learning analytics jobs have also proposed higher level parameters such as forgetful user [15] or hint abuse [16] that transmit complex information. Learning analytics have also been applied to educational games in different works; the amount of events triggered by users playing games makes it harder to select which data can be useful to infer useful information. The work presented by Serrano-Laguna [17] uses the source of data from an educational game to feed a learning analytics system to infer knowledge about the effectiveness of the students. There has also been proposal [18] about providing real-time learning analytics in educational games that would adequately match the dynamics of a game environment.

The work lead by Hakulinen [19] describes and evaluates the use of achievement badges in TRAKLA2 online learning environment with 281 students. They found that some badges can be used to affect the behavior of students and that differences in students' behavior exist between each type of badge. There are other researchers [20], [21] who have work towards modeling the behavior towards badges which findings seem to indicate that indeed badges affect the normal behavior of students. However to the best of our knowledge no work has proposed information about the behavior of students towards a badge system as we do in this proposal.

TABLE 1
TYPES OF BADGES PER CATEGORY AND REQUISITES DESCRIPTION

Badge category	*Type of badge*	*Requisites*	*Number of different badges*
Exercise	Streak	To solve several exercises correctly in a row of the same skill.	5
	Timed Problem	To solve a certain number of correct problems within a specific amount of time	5
	Exercise Completion	To complete some specific count of exercises	4
	Recovery Problem	To get exercise problems correct after having some problems solving exercises	2
	Unfinished Exercise	These badges are awarded when the user does not acquire a proficient level but he is answering many exercises correctly	3
	Topic	These badges are awarded when achieving a proficiency level in a subset of exercises	Different for each course
	Challenge	These badges are awarded when achieving a proficiency level in a set of exercises which are related to the same topic of knowledge	
Video	Topic Time	To watch a specific amount of videos in the same topic	4
	Video Time	To spend a certain amount of time watching videos	1
Social	Feedback	To receive up votes in your questions or answers	6
	Discussion	To flag or vote the questions and answers of other peers	4
	Profile	To customize your personal profile	1
General	Points	To earn a certain amount of points	2
	Power Time	To watch an specific amount of video and completing a certain amount of exercise problems within a set amount of time	3
	Consecutive Activity	To consecutively perform an activity on the site for a set of days in a row	3

Another important question related to this work is if badges are useful in education. Results from [22] concluded that badges could have a positive effect on learner motivations but as an extrinsic motivator it could have a bad influence on learning. Finally the systems Septris and SICKO [23] implement both learning analytics and gamification techniques to investigate its viability in medical education obtaining good opinions from the users.

3 GAMIFICATION ELEMENTS IN THE KHAN ACADEMY PLATFORM

The Khan Academy platform incorporates several features for gamification and learning analytics. The Khan Academy includes an *Energy Points* system. These points can be acquired by carrying out activities related to the learning process, such as solving exercises or watching videos. In addition, a strong badge system is implemented which is described in detail now. There are some other activities that can be considered as gamification components such as the possibility of establishing learning goals which are composed of exercises and videos; these goals can be seen as an analogy of achieving certain stages in games and receiving an award when the objectives are fulfilled.

Khan Academy implements a very complete badge system where students can receive badges as an award for their interaction in the platform. The Khan Academy system grants badges at real time whenever the conditions for that badge are fulfilled. The conditions are usually related to actions that are good for the learning process. The user can access detailed information of each badge and the conditions to get it. Students receive a notification whenever they earn a badge. Khan Academy divides the different badges into six different categories, some of these badges are very easy to acquire while others are incredibly difficult and would require a lot of effort. The first five categories are called *Meteorite*, *Moon*, *Earth*, *Sun* and *Black Hole Badges*, which go respectively from easiest to the hardest badges. The conditions required to acquire these badges are very diverse, some of them are related to solving a certain amount of exercises or watching a specific time of videos. Others requirement can be associated to social interactions (such as writing or voting a comment) or earning a certain quantity of points. However, Khan Academy also has some other badges which conditions are more complex such as "*Answer a problem correctly after having some trouble with many problems*" or very difficult to achieve such as "*Post 100 answers that earn 3+ votes*".

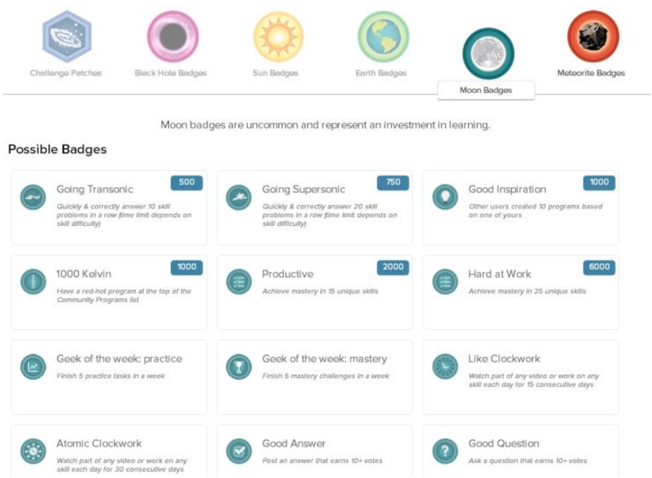

Fig. 1. Badge interface of achievements in the official Khan Academy server.

Some badges can only be earned by students once while others can be earned several times by the same student (repetitive badges). For example the badge *Nice Streak* which reads "*Correctly answer 20 problems in a row in a single skill*" can be earned many times, on the contrary the badge *Making Progress* which reads "*Achieve mastery in 7 unique skills*" can be earned only one time. In addition, the sixth badge category named as *Challenge Patches* englobes challenges where students need to finish all the exercises of a complete field of knowledge. Figure 1 shows the badge classification and some examples of the badges that can be achieved.

Table 1 shows the division of all badges in 4 categories which we have utilized for the purpose of this research. Each category contains the different types of badges that can be acquired and next a description of the requisites for each type of badge is presented. Finally, the last column has the number of different badges of the same type that can be earned of the same type of badge. For example the *Streak* type of badge have 5 different badges which are called *Nice*, *Great*, *Awesome*, *Ridiculous* and *Ludicrous Streak Badge* which are earned whenever the student correctly solves 20, 40, 60, 80 and 100 exercises in a row respectively. The quantity of topic and challenge badges is different for each course as we will explain this in-depth in section 5.1.2. Therefore, there is a base amount of common badges of 43 in all courses and a different amount of topic and challenge badges.

4 Modeling Students' Behavior Towards Badges

Some indicators which model the students' behavior towards gamification based on badges are proposed in this section. The objective is to propose some measures which can provide a deep insight about how the interaction of users with badges has been. The metrics defined are three (intentionality for topic and repetitive badges, and concentration), which are explained each one in a section. For each metric, we will first introduce the base idea about the behavior we are trying to detect and then how to apply to the particular case of our experience using the Khan Academy platform.

4.1 Intentionality for *Topic* badges

The requisites needed to receive each one of the badges are available in the achievements interface of Khan Academy. Therefore, students have a full access to this information and they can know what actions must be fulfilled to obtain each award. In this direction, it is interesting to get to know if students are intentionally maximizing the amount of badges that they are earning. This idea tries to detect whether the users' purpose when interacting with the platform is earning badges.

In order to make the implementation of this algorithm, we have focused onto the *Topic* badges which are described in table 1. The exercises required for acquiring these badges are always different and students must reach a proficient level in a specific set of exercises. In addition, the exercises required are not repeated in the requisites of the different badges. Therefore, the objective is to learn the number of badges the user has received, and compare it with the maximum number of badges he could have received, taking into account the total number of proficient exercises the user has mastered. As an example, think about an user who has achieved proficiency in 21 exercises earning only 2 badges when with that amount of exercises the user could have earned a maximum amount of 7 badges, in case the student had mastered other different exercises; thus this user does not seem to be very concerned towards earning the maximum amount of badges possible and he is not focused on earning badges.

Figure 2 shows the flow diagram that system performs in order to infer if the student is maximizing the number of badges acquired, and thus showing an intentional behavior. The system first needs to retrieve the number of skills and badges that the student has received. Next, we have designed an algorithm which obtains the maximum number of badges that the student could have received with the number of skills that he/she has mastered. The last step is to compare the number of badges that the student has received with the maximum number of badges that the student could have received with those skills. We present the measure from 0 to 100 where a value of 100 would mean that the student has earned the maximum number of badges which were possible with the skills he/she has.

4.2 Intentionality for repetitive badges

The intentionality measures whether students are earning badges on purpose, or they are receiving these badges just as part of their interaction with educational activities. In other words, we want to infer if a student is intentionally completing the badge conditions in order to acquire badges. This knowledge would allow us to determine which students are motivated about earning badges and to adapt some decisions to this knowledge. However this is a very hard task taking into account the variability badge system which students are using in Khan Academy. There are many different types of requisites and it is very hard to detect if students have received a badge because they were trying to earn it or just due to their interaction with exercises and videos. For example, how would it be possible to know if a student has commented a video in order to obtain a social badge or just because he wanted to make a question?

Our research suggests that it is not possible to infer this information for all types of badges. Therefore for this implementation we have only analyzed those badges where we think we have strong basis in order to infer intentionality. We have selected those badges which can be received repetitively when solving exercises. There are two types that fulfil this requisite which are *Timed Problem* and *Streak* badges which can be consulted in table 1. Therefore for this implementation we take into account a total number of 10 different badges.

Figure 4 shows the flow chart describing how to infer intentionality for this special case. Students should solve exercises until the system informs them that they have achieved a proficient level in that skill. Therefore, when a student keeps solving exercises after achieving proficiency in the skill and they earn a badge, we make the hypothesis that they keep solving exercises not to learn (they have already mastered the concept) but to earn more badges. At the end of the process we can calculate the percentage of repetitive badges that were intentionally earned. It will be presented from 0 to 100 where a value of 100 would mean that all the repetitive badges were earned intentionally

4.3 Concentration

Students can devote all their actions into fulfilling the requirements of one badge or carry out different actions which are non-related to the requirements before actually receiving the badge. This information can also allow us to detect if students are concentrating all their interaction into achieving a badge, therefore they are intentionally earning them. However the wide variety of requisites in the badge system makes very difficult to propose a model that could be applicable for all of them. In addition, some requisites do not apply well into this measure, for example it does not make sense to say that a student is concentrated into voting a video.

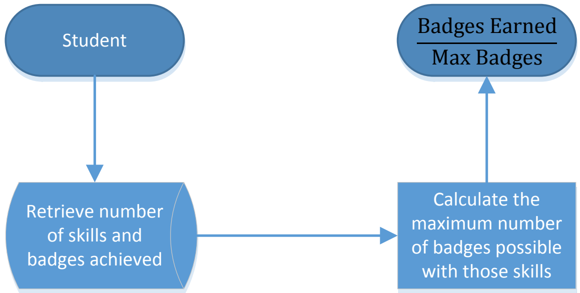

Fig. 2. Flow diagram to calculate the intentionality for topic badges

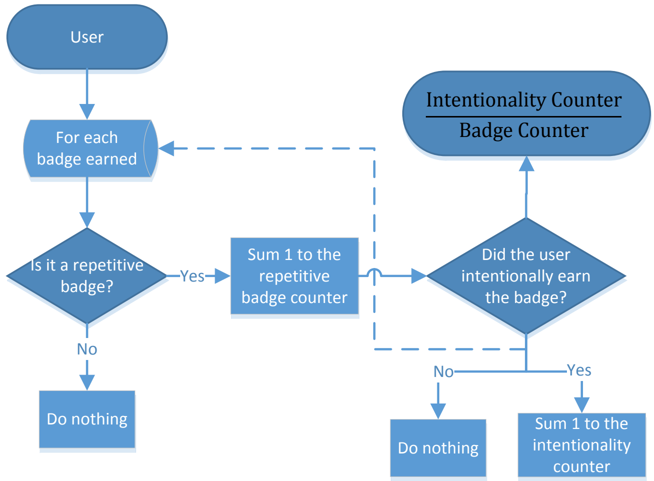

Fig. 3. Flow diagram to detect intentionality earning badges.

Therefore we have designed an algorithm to detect students' concentration of only *Topic* badges. It is reasonable to argue that since *Topic* badges have as requisites a fixed set of exercises, it is easier to track if students have done the required exercises or others before earning a *Topic* badge. Following this criterion we can infer the proportion of exercises that a student attempted that belong to the requisites before actually earning the badge

Figure 4 shows the flow diagram to infer the concentration of students when earning topic badges. We will retrieve all the badges earned by the user and analyze the previous actions for only those which are *Topic* badges. For each *Topic* badge that the student has received, the algorithm retrieve all the attempted exercises between the last *Topic* badge (in case is the first one, it retrieves all the exercises attempted until the current time) and the current badge. Next, the algorithm analyzes which of the exercises belong to the requisites of the current badge. Then we can calculate the concentration of the student when acquiring the current badge. Finally repeating this process for all *Topic* badges we can calculate a global level of concentration of the student.

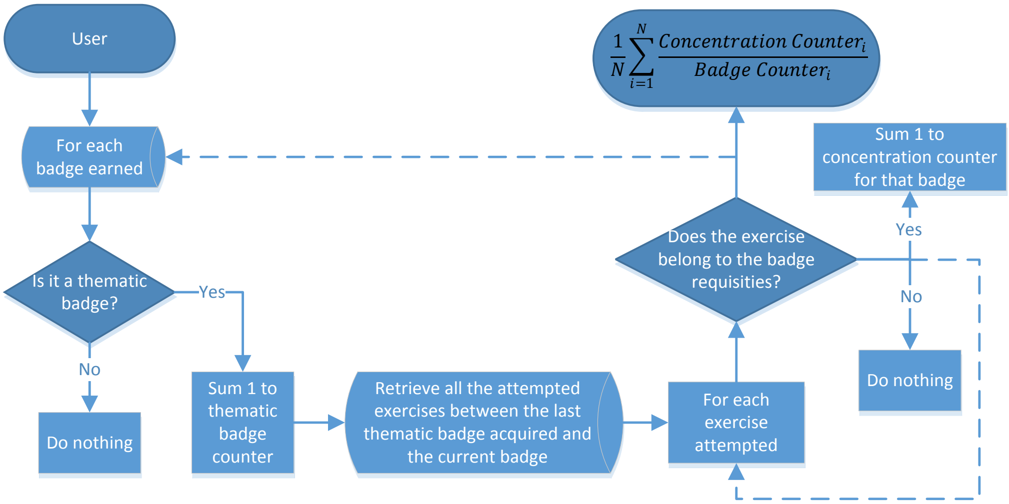

Fig. 4. Diagram to infer concentration earning topic badges.

This level of concentration gives an idea about if the student is focusing in fulfilling the requisites to receive badges in a continuous and concentrated way or just attempting exercises and they earn the badges by chance or they like to earn badges but likes to do in a disperse way.

4.4 Efficiency in Badges

Students devote a certain amount of time to the interaction with the platform. In addition, this time can be divided into different activities such as watching videos or solving exercises. The time invested by students can be used to obtain measures which give insight about the number of badges they earn per unit of time. This can allow a comparison between the different students to see which of them are more efficient towards achieving badges. We can also use the badge division by categories that we described in table 1 in order to provide information divided by the different categories of badges. In this direction we propose the next indicators related to the number of badges acquired by a student per unit of time, in other words the efficiency of the student earning badges (EB).

$$EB_{total} = \frac{total\ number\ of\ badges}{total\ time\ in\ the\ platform}$$

$$EB_{exercises} = \frac{exercise\ badges}{time\ spent\ solving\ exercises}$$

$$EB_{video} = \frac{video\ badges}{time\ spent\ watching\ videos}$$

This metrics can provide more information regarding the interest of students towards earning badges and also make a differentiation between the different badge categories that we have included. This will allow us to compare the tendencies of different students towards badges and also between different courses.

5 Badge Analysis of a Case Study

This section analyzes the use of badges in a real experience using the Khan Academy platform. Subsection 5.1 describes the experience and how the badge system was adapted. Subsection 5.2 analyzes in-depth the interaction of students with the badge system while subsection 5.3 presents the cluster analysis of students depending on the badge indicators.

5.1 Description of the Experience

The experience is framed in the contexts of zero courses that freshmen students have to take before starting their first year of an engineering degree at Universidad Carlos III de Madrid. These courses have an online period of time in August where the students review the concepts by themselves. Next, students take the face-to-face lessons which take place in September. This methodology in which students should prepare the lessons at home before receiving the actual lecturing class sessions where they will be able to solve their questions is known as “flipping the classroom”. This methodology was applied for the first time as a pilot initiative in August 2012 for a physics course and repeated again in August 2013 and 2014 for physics, mathematics and chemistry courses due to the success of the pilot experience. The data analysis that is presented here belongs to the physics, mathematics and chemistry courses of August 2013. The courses are composed by a series of exercises and videos which have been developed by the instructors of each course. A total number of 30 exercises and 25 videos in mathematics, 30 exercises and 30 videos in physics, and 49 exercises and 22 videos in chemistry. Chemistry is the course with more exercises and less videos, while mathematics and physics are more similar. This is important as many of the badges are related to correctly solve exercises and watching videos.

5.1.1 Participants

Students who participate in these experiences are freshmen who have enrolled to an engineering degree. Most of them have an age comprised from 17 to 19 years. The number of students is different for every course taking into account that they might have different requisites depending on the engineering degree that they have enrolled. The number of students whose data has been analyzed in this study is 167 for the physics course, 73 in chemistry and 243 in mathematics, which make a total number of 483 different data samples. Although the number of enrolled students is higher, we have taken into account those who logged into the Khan Academy platform at least once. It is also noteworthy that as some students had to enroll in more than one course depending on their engineering degree, thus the number of different students taken into account for this study is 291.

5.1.2 Customization of the Badge System

For the purpose of this experience, the badge system was customized at different levels. First, we redefined the name of the types of badges using the names of different touristic emplacements of Madrid, so that students can be more contextualized. One of the six categories was removed as these badges were impossible to obtain in our specific context. The new type names read like *Medallas Sabatini*, *Atocha*, *Neptuno*, *La Cibeles* and *Puerta de Alcala*. In addition, new images to represent the badge types were created and incorporated to the platform as can be seen in figure 5.

Fig. 5. Customized images and badge types for the experience. Customized challenge badges for the chemistry course.

Another customization was related to removing all the *Topic* and *Challenge* badges that Khan Academy has by default. New badges were added for each one of the different courses based on the new exercises provided by the team of teachers. This is necessary due to the fact that all the exercises of this experience were developed from scratch and all the exercises that Khan Academy incorporates by default were removed. In this direction, figure 5 shows the *Challenge* badges which were configured for the chemistry course. Similarly, *Topic* badges were also configured to be adapted to the existing exercises of each course and some examples can be seen in figure 6. A total amount of 7, 12 and 16 *Topic* badges and 3, 6 and 6 *Challenge* badges were added for mathematics, physics and chemistry respectively. Taking into account the common base of 43 badges explained in section 3, the total amount of badges considered for this research is 53, 61 and 65 badges for mathematics, physics and chemistry respectively.

Fig. 6. Configuration of new *Topic* badges for the chemistry course.

5.2 Badge Analysis

This section presents an analysis of the use of badges by the students of this real experience. Section 5.2.1 is focused on descriptive analysis of the different metrics defined. Section 5.2.2 presents the relationship among these gamification metrics and among other important metrics for the learning process. Finally, section 5.2.3 presents a clustering analysis depending on the student behavior regarding badges. It is noteworthy that when presenting descriptive data such as the mean value of badges per user we use all the data samples (483 data samples) but when we perform statistical analysis we will use only different students (291 data samples) as independent cases in the samples are required for some statistical tests. The data sample of different students is obtained by randomly assigning those students who have participated in more than one course to one of the courses only. The data sample of different students is composed by 164 mathematics students, 88 physics students and 39 chemistry students.

5.2.1 Evaluation of the use of badges

This section presents a general analysis of the achievement of badges by students of the three courses. As some badges are quite straightforward to acquire with the interaction with the platform, then most of the students have obtained some of them (even if they did not have the intention to get them). Nevertheless, there are others that are very hard to earn. The total number of badges delivered is 1153, 1609 and 4773 for the chemistry, physics and mathematics courses respectively. Taking into account the number of students in each course, the number of badges per student is 15.8, 9.6 and 19.64 respectively. There exists a notorious difference between physics (9.6 badges per user) and mathematics (19.64 badges per user) that almost doubles the physics number. A One-Way ANOVA was used to check whether these differences are statistically significant. The data samples are independent as each of them belongs to a different course and different users; both Kolmorov-Smirnov and Shapiro-Wilk let assume that the sample come from a distribution which is normally distributed. Finally the Levene test let assume the homogeneity of variances between groups. As the assumptions are met, the ANOVA test (F=7.9, p=0.000) demonstrates that the differences between the amount of badges per student in each course are statistically significant. These results might be associated with the fact that the progress in mathematics and chemistry courses was higher than the one in physics, and earning badges is strongly associated with the time invested by students. But there are other possible causes.

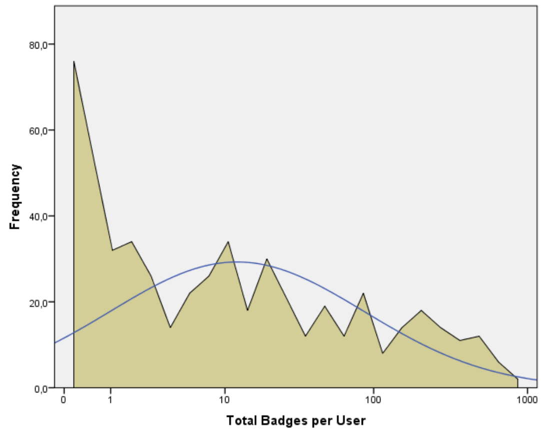

Fig. 7. Histogram of the total quantity of badges earned by each user in all courses

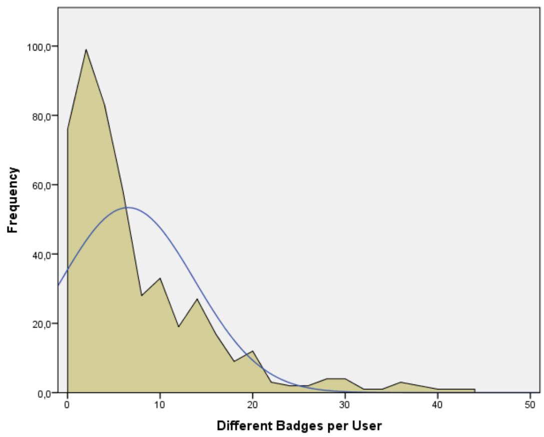

Fig. 8. Histogram of the quantity of different types of badges earned by each in all courses.

Figure 7 shows a histogram chart of the total amount of badges earned by each user taking into account all the courses. The y-axis represents the frequency of each case and the x-axis the amount of badges earned by the user. The y-axis is in a linear scale whilst the x-axis is logarithmic in order to improve visual details of the representation. Additionally the blue line represents the normal curve that fits the histogram. This graph describes the distribution of badges earned by each user and can be used to see the evolution and the most common cases. For example, we can see an important peak in 0 and 1 badges which are those users who have interacted very little with the platform. In addition, it is interesting to see how there are many students in the interval from 100 to 1000 which has achieved a big amount of badges. We should also note out that the distribution have many peaks in different places.

Analogously figure 8 represents a histogram of the quantity of different badges earned by each user considering all courses. This time both axes have a linear scale. The vast majority of the population is concentrated in the interval from 1 to 10 different badges. There are important differences with figure 7 since the previously analysis took into account all those badges that can be earned repeatedly and this only different types of badges. This distribution has more abrupt descendant curve that the other one.

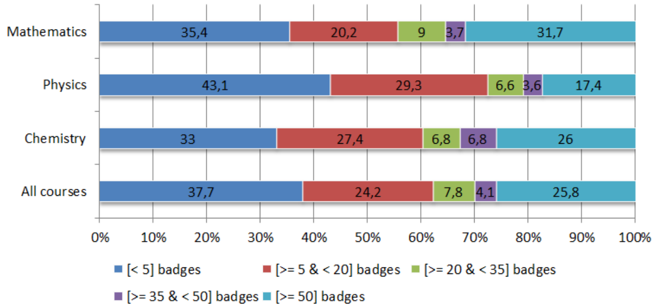

Fig. 9. Percentage cumulative bar chart of the distribution of all earned badges by each student

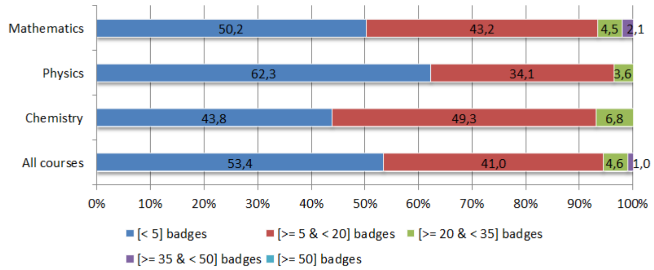

Fig. 10. Percentage cumulative bar chart of the distribution of different types of badges earned by each student.

We can make deeper analysis by establishing intervals and comparing courses. Figure 9 shows a percentage cumulative bar chart of the distribution of the amount of badges earned by each user in five intervals which are [<5], [>= 5 & < 20], [>= 20 & < 35], [>= 35 & < 50] and [>= 50] badges earned by the user. We have divided the data for each course so we can compare class tendencies. The comparison between chemistry, physics and mathematics shows that the distributions of badges earned by each user are very alike in the three courses. The most populated intervals are the extreme ones, being users who earn few badges ([<5] and [>= 5 & < 20]) and users who earn a lot of badges ([>= 50]). Here we can identify that the most common cases is to either earn a lot of badges or just few of them which could be related to those users who have been motivated by earning badges and those who were not motivated. Analogously, figure 10 shows a percentage cumulative bar chart but this time it represents the distribution of different badges earned by each user. The intervals are the same and we can see important differences with the distribution of figure 9 and 10. A percentage of 53.4% of all the students means that more than half of the students have received less than 5 different badges and a 41.0% of the students received between 5 and 20 badges. This leaves a total percentage of only 5.6% of students who earned more than 20 different badges, which seems to be a very low number. These huge differences are probably due to the fact that some badges can be received repeatedly and some users with an elevated *Intentionality for repetitive badges* indicator earn many times the same badge (even when they have already mastered the skill) but not too many different badges. In addition, there are not huge differences between the distributions of the three courses between figures 9 and 10.

It is also interesting to look into which badges are more commonly awarded in these experiences. Table 2 shows the position and count of the ten most awarded badges for each of the courses. The count is expressed in badges per user in order to normalize the value with the quantity of students of each course, so it can be comparable. *Nice* and *Great Play List Time* belong to *Topic Time* type of badge, *Act One Scene One* to *Video Time* type, *Ten Thousandaire* to the *Points* type and *Getting Started* and *Making Progress* to the *Exercise Completion* type. These badges have different requirements as can be consulted in table 1. An interesting detail of these results is that the badge type order for both physics and chemistry is exactly the same except for the positions 9th and 10th which are exchanged. Anyhow we can see that the ranking of badges of the three courses are very similar, most of these badges have easy requirements or can be awarded many times, thus it makes sense that students have received these badges the most. None of the *Topic* or *Challenge* badges are in the top positions due to it is harder to obtain and can be received only once. In addition, there are not social badges in top positions which can be related to the low social participation of students using the Khan Academy platform in this experience, as we have reported in previous work [24]. We should take into account that students' main communication was using the forums of Moodle instead of the Khan Academy platform.

Some badges are triggered when solving exercises or for watching videos, thus it is interesting to analyze which exercises and videos trigger the biggest amount of badges. The causes can be very diverse, for example an easy exercise can be used to obtain many correct exercises in a row or a difficult topic might trigger more video badges because students need to watch the video more than once. Table 3 shows for each one of the courses the two exercises and videos which triggered more badges. There are some important differences between those exercises and videos which triggered the biggest amount of badges and those which triggered the least; to put it into an example for the math course, the exercise which triggered more badges is the *Biquadratic equation* exercise with a 301 badge count whilst *Basic operations with complex numbers* triggered only 31 badges. These differences are also applicable for the others courses and in videos as well. We have analyzed the possible reasons for these different results such as the position of the video and exercise in the course order (at the beginning or end of the course), video length or possible complexity of the topic. Despite we have not found conclusive causes for these results, we can hypothesize that the complexity of the topic (easier exercises are correctly solved more often and complex videos are re-viewed) and the course order (the contents that are placed at the beginning of the course tend to be more accessed) play an important role.

TABLE 2
POSITION AND COUNT OF THE TEN MOST AWARDED BADGES

Badge type	*Badge ranking position: Badge count*		
	Math	*Physics*	*Chemistry*
Nice Play List Time	2nd: 3.3	1st: 3.2	1st: 3
Nice Streak	3rd: 3.1	2nd: 1.8	2nd: 3
Nice Timed Problem	1st: 3.3	3rd: 1.3	3rd: 1.9
Great Streak	4th: 2	4th: 0.92	4th: 1.5
Act One Scene One	6th: 0.8	5th: 0.5	5th: 0.8
Ten Thousandaire	8th: 0.53	6th: 83	6th: 0.6
Great Timed Problem	5th: 1.2	7th: 0.4	7th: 0.6
Getting Started	10th: 0.47	8th: 0.3	8th: 0.5
Great Play List Time	7th: 0.7	10th: 0.2	9th: 0.1
Making Progress	12th: 0.3	9th: 0.2	10th: 0.3

TABLE 3
EXERCISES AND VIDEOS WHICH TRIGGERED MORE BADGES

	Exercise	*Video*
Math	Biquadratic equation	Functions: domains
	Derivatives I	Derivatives II
Physics	Force diagram	Circular motion
	Scalar and vector magnitudes	Newton's Laws
Chemistry	Atom and subatomic particles	Electronic levels
	Quantum numbers	Le Chatelier's Principle

5.2.2 Relationship between badge indicators and others

This sub-section presents the correlation of the badge metrics that have been presented in section 4 with other metrics related to the students behavior during the learning process. Many of these metrics have been presented by the authors of this work in previous research [24]. It also presents the correlation between the badge behavior metrics with themselves. The objective is to present the existing and underlying relationships between the badge indicators and others, in order to learn more about the relationships of gamification behavior towards badges of the students in a Khan Academy course and make reasonable hypothesis. Table 4 describes the Bivariate Pearson correlation (N=291, two-tailed) of the badge indicators with several indicators. We apply a Bonferroni correction in order to reduce the probabilities of finding false significant correlations. The asterisk marks those correlations which are significant at the 99% level after applying the Bonferroni correction. We have divided this analysis into four groups which are separated by lines. The indicators from the first three groups have been previously described in our work [25] and the objective is to join together indicators which provide information about the same topic. The last group is the correlation between the four metrics about badge behavior.

TABLE 4
BIVARIATE PEARSON CORRELATION OF BADGE INDICATORS WITH OTHERS INDICATORS

Bivariate Pearson Correlation two-tailed (N=291)	*Intentionality for topic badges*	*Intentionality for repetitive badges*	*Concentration*	*EB_{total}*
Exercises Accessed	0.456*	0.464*	0.361*	0.438*
Videos Accessed	0.305	0.322*	0.228*	0.225*
Exercise Abandonment	-0.456*	-0.327	-0.399*	-0.352*
Video Abandonment	-0.177*	-0.125*	-0.168	-0.49
Total Time	0.510*	0.409*	0.372*	0.338*
Optional Activities	0.489*	0.358*	0.345*	0.446*
Follow recommendations	0.169*	0.041	0.202*	-0.017
Forgetful user	0.010	-0.053	0.010	-0.104
Unreflective user	0.032	0.024	0.027	0.019
Video avoidance	0.004	-0.047	-0.012	-0.009
Hint avoidance	-0.027	-0.025	-0.063	0.009
Hint abuse	-0.06	-0.031	-0.015	-0.087
Proficient exercises	0.737*	0.511*	0.563*	0.625*
Completed videos	0.326*	0.302*	0.251*	0.279*
Intentionality for topic badges	1	0.445*	0.859*	0.552*
Intentionality for repetitive badges	0.445*	1	0.417*	0.573*
Concentration	0.859*	0.417*	1	0.459*
EB_{total}	0.552*	0.573*	0.459*	1

The first group of indicators composed by *Exercises* and *Videos Accessed*, *Exercise* and *Video Abandonment*, *Total Time* and *Optional Activities* are related to the *Total Use* of the *Platform.* As we can see many of these indicators have been found statistically significant with badge metrics, which makes sense and we can determine that the more use of the platform students do the more badges they earn. Especially, we can see that all the indicators from *Use of the Platform* have been found statistically significant with the indicator of *Efficiency in Badges*. In addition, both *Exercise* and *Video Abandonment* have been found negatively and significantly correlated with several badge metrics, which proves to be true that the platform engagement plays an important role in the motivation towards earning badges.

The indicators which were less correlated to badge metrics are *Videos Accessed* and *Video Abandonment*, which are coherent results taking into account that most of badge metrics does not take into account video activity. Finally *Exercises Accessed*, *Exercise Abandonment*, *Total Time* and *Optional Activities* are strongly correlated with most of the badge metrics.

The second group of indicators is related to the *Exercise Solving Habits* of students. These are behavioral indicators about how students act when solving exercises. These indicators are *Follow Recommendations*, *Forgetful User*, *Unreflective User*, *Video Avoidance*, *Hint Avoidance* and *Hint Abuse*. As we can see all these behavioral indicators are not correlated with badge metrics with the exception of the *Follow Recommendations* indicator. Two statistically significant but low correlations where found between *Follow Recommendations* with *Intentionality for topic* *badges* and *Concentration*, it is not strange to think that an user who follows recommendations might also behave with certain order when solving exercises and thus, earning badges.

The third group of indicators is related to the *Correct Progress in the Platform*. The correlation with *Completed Videos* is significant but low, as we have said before there are less video badges than those who are earned by solving exercises. Still the correlation exists and this is due to the fact that those users who completed a lot of videos probably show an engagement with the platform, hence they probably have earned many badges also. The correlation with *Proficient Exercises* is the highest of all indicators which is probably related to the big amount of exercise badges that can be earned repeatedly. Two correlations are especially significant, first with *Intentionality for topic badges* (0.737, p=0.000) which makes sense because for those users who have skilled proficiency in many exercises would have earned also many *Topic* badges. Secondly with *Efficiency in Badges* (0.625, p=0.000) which also makes sense due to the more proficient exercises, more badges the student will earned hence *Efficiency in Badges* will also be higher.

The last group separated by a double line presents the correlations between the badge metrics themselves. All the correlations have resulted to be statistically significant due to probably the fact that when one student shows motivation towards earning badges it will be reflected in several of these indicators. We should point out the correlation between *Intentionality for topic badges* and *Concentration* (0.859, p=0.000) which is the highest of all the correlation analysis. This correlation is probably very high as the students who are concentrated earning topic badges are also probably maximizing and earning a many topic badges. Anyhow we can present hypothesis and possible reasons but not assure the causes for the results provided in this section.

5.2.3 Clustering students by their badge indicators

A common problem is to be able to cluster data samples into different groups in regards of their similarity. Most of the times a decision should be done without having any prior knowledge of the number of data groups or how the data is distributed. With the objective of solving these problems, unsupervised learning algorithms also known as clustering are often used. In the context of this experience, we would like to know which students have been motivated by the badge activity and group them according to their preferences. This information can be used to divide all the students into different groups in order to customize each of the later classes with the preferences of each group. For example, a group of students who have not participated much using Khan Academy and they did not have been motivated by badges might perform better with a more traditional learning approach whilst those who really enjoy the gamification elements, will perform better if these elements are introduced in future lessons.

In this section, we present the application of a clustering algorithm in order to cluster the students depending on their badge activity. We have decided to use for this work a Two-Step Cluster Analysis procedure since we do not know how many groups we should find. This algorithm gives its best performance when all variables are independent and have a normal distribution, however it is very robust and it performs well when these conditions are not met. Both Kolmogorov-Smirnov and Shapiro-Wilk test confirm that the four badge metrics can be considered as normally distributed and the variables are independent, thus the conditions are met. Anyhow, you can always determine if the solution is satisfactory for your needs after applying the algorithm.

We apply the Two-Step Cluster algorithm leaving the number of groups to be determined by the execution of the algorithm automatically, with a Log-likelihood measure for distance and a Schwarz's Bayesian Criterion for clustering. Finally, we choose the badge indicators *Intentionality for topic and repetitive badges*, *Concentration*, and *Efficiency in Badges* as continuous variables for the clustering algorithm. We will also use the *Proficient Exercises* and *Total Time* indicators to support the evaluation of results from the algorithm output.

The Two-Step algorithm has selected three clusters providing a good cluster quality in terms of cohesion and separation (0.75 in a quality ranging from -1 to 1). The smallest cluster has 70 students (24.1%) whilst the largest has 149 students (51.2%) providing a ratio of sizes of 2.13; the middle-sized cluster has 72 students (24.7%). The predictors' importance for the four continuous variables has been 1.0, 0.94, 0.79 and 0.31 for *Intentionality for repetitive badges*, *Concentration*, *Intentionality for topic badges* and *Efficiency in Badges* respectively. We can see results about the distribution of the clusters which have been obtained in figure 11. Each column represents one of the clusters; the first row describes the size of the cluster with the number of students and the percentage. The second row is composed by the four badge indicators which have been the inputs to the Two-Step algorithm, this row shows the relative distribution of each of the badge indicators for the population of students belonging to the cluster. In addition, in the upper right side of the badge indicator distribution, the mean value for the indicator in that cluster is represented. Finally, the last row contains the evaluation fields *Proficient Exercises* and *Total Time* which we are going to use to support the evaluation of the cluster. The evaluation fields are detailed in the same data format explained for the inputs. The background color of both inputs and evaluation fields' distribution codifies the importance of that variable as predictor; the legend can be consulted in the upper part of the chart.

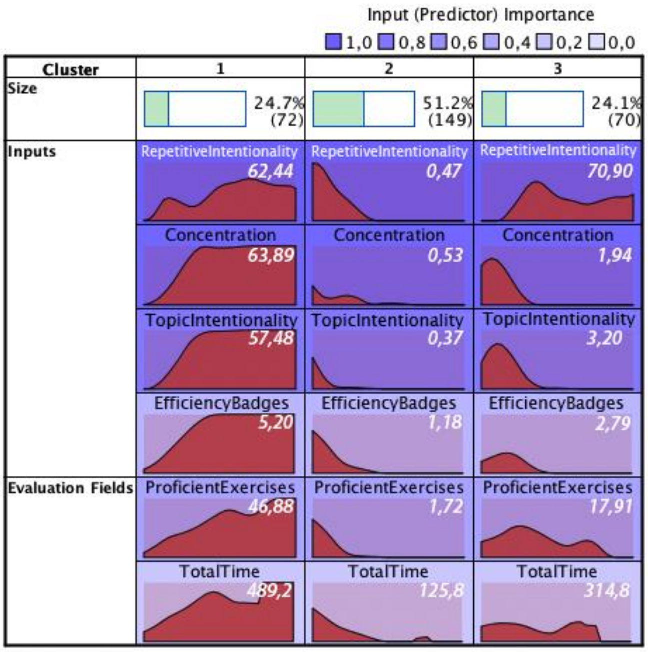

Fig. 11. Overview of the results of a Two-Step Cluster Analysis using badge indicators as inputs.

The information provided about the distribution, the inputs and evaluation fields in each cluster can be used to learn what type of students composed each group:

- Cluster 1: The first cluster is composed by the 24.7% of students. We can rapidly perceive that the students who belong to this cluster are those who have put the greatest effort in the platform in terms of number of proficient exercises and total time. The mean value of proficient exercises is 46.88%, and the total time is 489.2 min per user on average, which are both high values. In addition, they have high values in all badge metrics when compared to the rest of the clusters. The average user of this cluster has made an important effort in time, and has made a considerable advancement in exercises showing also interest in the badge system.
- Cluster 2: The second cluster is composed by the 51.2% of the students and it is quite the contrary of the first one. These students have made a low effort using the platform, we can see that on average they have invested 125.8 min per user obtaining only 1.72% of proficient exercises. In addition, they have not shown interest in any of the badge indicators.
- Cluster 3: The third cluster is less concrete than the others two being composed by the 24.1% of the population. We can see that the students in this cluster have invested a decent amount of time with 314.8 minutes per student but their progress has not been so good with only 17.91% of proficient exercises. The badge metrics are more interesting; *Concentration* and *Intentionality for topic badges* show very low values (1.94 and 3.20) but *Intentionality for repetitive badges* is even higher than in the first cluster. Finally the *Efficiency in Badges* indicator shows a moderate value. Therefore, this cluster concentrates users who are very low on *Concentration* and *Intentionality for topic badges* and therefore they are not doing an organized effort towards achieving badges, but the very high *Intentionality for repetitive badges* value demonstrates that they are very eager to earn those repetitive badges, thus they are interested in the badge system. These students have invested a moderate amount of time but they have not achieved a great progress, additionally they have shown low *Concentration* and *Intentionality for topic badges* but high *Intentionality for repetitive badges*, we could say that they have behaved a bit anarchically doing what they enjoy without paying much attention to their progress or suggestions.

This information can be used to divide all the students who took these courses into different classes where the teaching method can be adapted to their preferences. To this purpose we could create e.g. four different classes with the results that have been obtained. The first two classes can be composed by half of the students in cluster 2 in each class. These students have not shown a high motivation neither using Khan Academy platform and its learning concept by exercises and videos nor by earning badges. Therefore, these two classes could perform better if a more traditional learning approach is applied. The third class can be composed by the students of cluster 1, which have both found interest in the Khan Academy format with online videos and exercises and their badge indicators are also high. Therefore, this class will probably perform well using innovative methodologies like 'flipping the classroom' and if part of the learning process involves VLEs with gamification elements. Finally, the last class can be composed by students from cluster 3. These students performed a moderate progress, but they have shown a very important motivation towards earning repetitive badges. Therefore instructors teaching this class could add some gamification elements in order to motivate students in their learning process.

Additionally, we also performed a cross tabulation between the clusters that have been obtained and the gender and course of students. This categorical variable analysis reveals if there exists any association between these variables. The Pearson Chi-Square Test reports that there are not significant associations between the cluster of the student and his/her gender or the course they were taking. Therefore, we assume that the gender and the course did not have an important influence on how students were elected for each cluster.

Finally, we wanted to include a clustering visualization of the results. Since we have used 4 different variables for the clustering process we cannot use common 2 dimensional visualizations. We can either apply a multidimensional scaling or use a high dimensional data visualization in order to represent the results. We have chosen to make use of a high dimensional data visualization that has been widely used. This method is denominated as parallel coordinates [26]; it uses consecutive parallel axes in which an *n*-dimensional point will be represented as a polyline with vertices on the parallel axes. This type of visualization can be used to detect 2D patterns and it is often use for clustering purposes. Despite this is a broadly used visualization, there are that not many free tools or libraries that implement it. In this case we have used GGobi[1] visualization system.

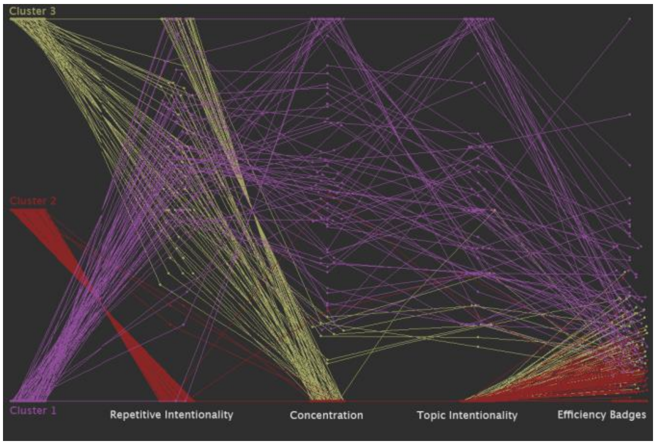

Fig. 12. Parallel coordinates visualization of the clustering variables where the polyline color represents its cluster.

Figure 12 represents the multidimensional parallel coordinates visualization for the clustering variables. Each one of the polylines starts from the cluster variable, which was assigned after the Two-Step Cluster Analysis, and its color represents the cluster to which it belongs. Cluster 1 is represented by the purple color, cluster 2 by red and cluster 3 by yellow. The parallel axes are in a descendent order by the importance of the variable they represent. The representation shows the users that represent clusters 2 and 3 are very alike as there are almost no outliers. Cluster 1 is composed of a more diverse source of students but all of them have shown an interest in the gamification indicators. As we can see the data representation is very similar as the discussion we offered previously, and the students belonging to each cluster are very similar, thus we think that the results and groups from the clustering analysis are good.

6. Conclusions

The work conducted in this paper has been focused on analyzing the interaction of students with the badge system of the Khan Academy platform in three course experiences. The analysis provides the distribution of the number of badges and different types of badges each user earned. We have also provided which badges are most earned by the students. In addition, we found that some exercises and videos triggered much more badges than others and this might be due to the complexity of the topic and the position of the course. These results can be used to put special emphasis on some topics and resources in future courses.

We have defined four high level indicators about the students' interactions with the gamification system. These indicators try to measure the intentionality of students to earn the badges (both topic and repetitive ones), the concentration or disparity to achieve this objective, and the efficiency to achieve it. Although these indicators have been defined for the Khan Academy platform, they can be easily generalized for other platforms. The intentionality for obtaining topic badges should consider a set of badges for which each badge can be achieved by obtaining the proficiency in a set of predefined exercises. Each badge cannot repeat an exercise as requisite if is part of another badge's requisites. The intentionality for achieving repetitive badges should consider badges that can be obtained several times for students. The concentration should consider topic badges that cannot have in common any requisites as exercise, while the efficiency is quite general without any restrictions. If the requirements for this four metrics are not fulfilled in other e-learning platforms, they should be redefined although ideas from them can be taken.

We show the clustering of students by means of these badge indicators for the presented case study. We discovered three different groups of students with similar attitudes towards gamification. These can be used by instructors to put them together in practical sessions or to customize their learning activities to improve their performance.

The correlational analysis shows that the different gamification indicators are related although their definitions are different. It provides the level of relation of a student who has intention to obtain topic badges has also of obtaining repetitive badges, the concentration or the efficiency towards them. We also found that the badge indicators are strongly correlated with those indicators related to the use of the platform and also to those about the progress in the platform, which is quite logical as making activities in the platform leads to achieve badges if they are done according to the badge requirements. We also found that there were not high correlations between badge indicators and behavioral habits when solving exercises such as hint abuse.

Acknowledgment

Work partially funded by the EEE project, "Plan Nacional de I+D+I TIN2011-28308-C03-01" and the "Emadrid: Investigación y desarrollo de tecnologías para el e-learning en la Comunidad de Madrid" project (S2009/TIC-1650)".

References

[1] Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. Educause Review. Retrieved from http://www.elmhurst.edu/~richs/EC/OnlineMaterials/SPS102/Teaching and Learning/Penetrating the Fog.pdf

[2] Deterding, S., Dixon, D., Khaled, R., & Nacke, L. E. (2011). Gamification: Toward a Definition. In ACM CHI Conference on Human Factors in Computing Systems (pp. 12–15).

[3] NMC Horizon Report 2014 K-12. (2014). New Media Consortium and EDUCAUSE Learning Initiative.

[4] Dyckhoff, a. L., Lukarov, V., Muslim, A., Chatti, M. a., & Schroeder, U.

[1] http://www.ggobi.org

(2013). Supporting action research with learning analytics. In Proceedings of the Third International Conference on Learning Analytics and Knowledge - LAK '13 (p. 220). New York, New York, USA: ACM Press. doi:10.1145/2460296.2460340

[5] Sharples, M., Mcandrew, P., Weller, M., Ferguson, R., Fitzgerald, E., & Hirst, T. (2013). Innovating Pedagogy 2013. The Open University.

[6] Arnab, S., Berta, R., Earp, J., Freitas, S. De, Popescu, M., Romero, M., Usart, M. (2012). Framing the Adoption of Serious Games in Formal Education. Electronic Journal of E-Learning, 10(2), 159–171.

[7] Freire, M., Blanco, Á., & Fernández-Manjón, B. (2014). Serious Games as edX MOOC Activities. In Global Engineering Education Conference (EDUCON) (pp. 867–871). Istanbul, Turkey: IEEE.

[8] Montola, M., Nummenmaa, T., Lucero, A., Boberg, M., & Korhonen, H. (2009). Applying Game Achievement Systems to Enhance User Experience in a Photo Sharing Service. In Proceedings of the 13th International MindTrek Conference: Everyday Life in the Ubiquitous Era (Vol. 4, pp. 94–97). New York, NY, USA: ACM. doi:10.1145/1621841.1621859

[9] Magaña, V., & Organero, M. (2014). The Impact of Using Gamification on the Eco-driving Learning. In C. Ramos, P. Novais, C. E. Nihan, & J. M. Corchado Rodríguez (Eds.), Ambient Intelligence - Software and Applications SE - 5 (Vol. 291, pp. 45–52). Springer International Publishing. doi:10.1007/978-3-319-07596-9_5

[10] Li, W., Grossman, T., & Fitzmaurice, G. (2012). GamiCAD : A Gamified Tutorial System For First Time AutoCAD Users. In 25th annual ACM symposium on User interface software and technology (pp. 103–112). New York, NY, USA: ACM. doi:10.1145/2380116.2380131

[11] Domínguez, A., Saenz-de-Navarrete, J., de-Marcos, L., Fernández-Sanz, L., Pagés, C., & Martínez-Herráiz, J.-J. (2013). Gamifying learning experiences: Practical implications and outcomes. Computers & Education, 63, 380–392. doi:10.1016/j.compedu.2012.12.020

[12] Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does Gamification Work? – A Literature Review of Empirical Studies on Gamification. In 2014 47th Hawaii International Conference on System Sciences (pp. 3025–3034). Ieee. doi:10.1109/HICSS.2014.377

[13] Ruipérez-Valiente, J. A., Muñoz-Merino, P. J., Leony, D., & Delgado Kloos, C. (2014). ALAS-KA: A learning analytics extension for better understanding the learning process in the Khan Academy platform. Computers in Human Behavior. doi:10.1016/j.chb.2014.07.002

[14] Zhang, H., Almeroth, K., Knight, A., Bulger, M., & Mayer, R. (2007). Moodog: Tracking students' online learning activities. In World Conference on Educational Multimedia, Hypermedia and Telecommunications, (pp. 4415–4422). Chesapeake, VA: AACE. doi:10.1.1.82.951

[15] Muñoz-Merino, P. J., Valiente, J. A. R., & Kloos, C. D. (2013). Inferring higher level learning information from low level data for the Khan Academy platform. In Proceedings of the Third International Conference on Learning Analytics and Knowledge - LAK '13 (pp. 112–116). New York, New York, USA: ACM Press. doi:10.1145/2460296.2460318

[16] Aleven, V., Mclaren, B. M., Roll, O., & Koedinger, K. (2004). Toward Tutoring Help Seeking ; Applying Cognitive Modeling to Meta-Cognitive Skills (pp. 227–239). doi:10.1007/b100137

[17] Serrano-Laguna, Á., Torrente, J., Moreno-Ger, P., & Fernández-Manjón, B. (2012). Tracing a Little for Big Improvements: Application of Learning Analytics and Videogames for Student Assessment. Procedia Computer Science, 15, 203–209. doi:10.1016/j.procs.2012.10.072

[18] Minović, M., & Milovanović, M. (2013). Real-time learning analytics in educational games. In First International Conference on Technological Ecosystem for Enhancing Multiculturality (pp. 245–251). New York, NY, USA: ACM. doi:10.1145/2536536.2536574

[19] Hakulinen, L., Auvinen, T., & Korhonen, A. (2013). Empirical Study on the Effect of Achievement Badges in TRAKLA2 Online Learning Environment. In Learning and Teaching in Computing and Engineering (pp. 47–54). Washington, DC, USA: IEEE. doi:10.1109/LaTiCE.2013.34

[20] Anderson, A., Huttenlocher, D., & Kleinberg, J. (2013). Steering User Behavior with Badges. In *22nd international conference on World Wide Web* (pp. 95–106).

[21] Grant, S., & Betts, B. (2013). Encouraging User Behaviour with Achievements : An Empirical Study. In *10th Working Conference on Mining Software Repositories* (pp. 65–68).

[22] Abramovich, S., Schunn, C., & Higashi, R. M. (2013). Are badges useful in education?: it depends upon the type of badge and expertise of learner. *Educational Technology Research and Development*, *61*(2), 217–232. doi:10.1007/s11423-013-9289-2

[23] Tsui, J., & Edtech, S. (2014). Septris and SICKO: Implementing and Using Learning Analytics and Gamification in Medical Education. Educause Review, (March), 1–7.

[24] Ruipérez-Valiente, J.A. Muñoz-Merino, P.J. Delgado Kloos, C. Niemann, K. Scheffel, M. Do Optional Activities Matter in Virtual Learning Environments? Proceedings of the 9th European Conference on Technology Enhanced Learning, EC-TEL 2014, Graz, Austria, September, 2014.

[25] Muñoz-Merino, P. J., Ruipérez Valiente, J. A., & Kloos, C. D. (2013). Inferring higher level learning information from low level data for the Khan Academy platform. In Proceedings of the Third International Conference on Learning Analytics and Knowledge - LAK '13 (pp. 112–116). New York, New York, USA: ACM Press. doi:10.1145/2460296.2460318

[26] Inselberg, A. (1990). Parallel coordinates: a tool for visualizing multidimensional geometry. In *Proceedings of the First IEEE Conference on Visualization* (pp. 361 – 378). San Francisco, CA: IEEE. doi:10.1109/VISUAL.1990.146402