Relevance of non-activity representation in traveling user behavior profiling for adaptive gamification

María Dalponte Ayastuy*
mdalponte@unq.edu.ar
LIFIA, Facultad de Informatica, Universidad Nacional de La Plata
La Plata, Buenos Aires, Argentina
Depto CyT, Universidad Nacional de Quilmes
Bernal, Buenos Aires, Argentina

Diego Torres*
diego.torres@lifia.info.unlp.edu.ar
LIFIA, CICPBA-Facultad de Informática, Universidad Nacional de La Plata
La Plata, Buenos Aires, Argentina
Depto CyT, Universidad Nacional de Quilmes
Bernal, Buenos Aires, Argentina

ABSTRACT

Collaborative location collecting systems (CLCS) are collaborative systems where users collects location-based data. When these systems are gamified and aim to adapt the game elements to each user, it may require a user traveling behavior profile. This work presents two approaches of traveling user behavior profiling: a raw series built up with categorical data that describes the user's activity in a period, and a timed series that is an enhanced version of the first that includes a representation of the non-activity time frames. The profiling of user traveling behavior can be used in adaptive gamification strategies. The approach is evaluated over a behavioral atoms dataset based on a year of Foursquare check-ins. The results showed that both approaches reflex different aspects of traveling user behavior, and also both could be used in a complementary manner.

CCS CONCEPTS

• **Human-centered computing** → *Reputation systems*; • **Computing methodologies** → **Cluster analysis**.

KEYWORDS

User profiling, Adaptive gamification, Collaborative location collecting systems, Dynamic time warping clustering

ACM Reference Format:
María Dalponte Ayastuy and Diego Torres. 2021. Relevance of non-activity representation in traveling user behavior profiling for adaptive gamification. In *XXI International Conference on Human Computer Interaction (Interacción '21), September 22–24, 2021, Málaga, Spain.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3471391.3471431

1 INTRODUCTION

Adaptive gamification [7] is a sub-area in the gamification [6] research field that study the adaptation of gamification strategies. Indeed, it is not desirable that a gamification strategy be generalized to any user, and it should be tailored according to either user's motivations, needs, values, or personalities[5]. As adaptive gamification is a promising research area, it could be applied to several contexts. This article contextualized the adaptive gamification in collaborative location collecting systems.

Collaborative location collecting systems (CLCS) are collaborative systems where users collects data with their location using a mobile application. There is a wide family of CLCS. One of the most popular was Foursquare, where people collect information about public places such as restaurants or pubs. Also, citizen science projects use CLCS to collect scientific data. For example, E-bird[18] project to spot in a map birds visualizations, or iNaturalist[21] project to collect biodiversity information[21], a citizen science project where people collect information about biodiversity with a mobile application.

Most of these projects motivate users to collect as much data as possible, such as collecting information from an area with little information or adding new areas into the visited places. Several cases use gamification strategies to motivate users, such as badges and challenges in Foursquare or contest in iNaturalist. Unfortunately, none of these declared an adaptive gamification strategy.

Challenges are one of the most used game elements in gamified collaborative systems[3]. There are rigid challenges that are tight to the general rules of the gamification, and the emergence challenge offers flexibility to the player to solve them[8, 10, 22]. Most of the use of this game element is not adapted to the user[3].

Vahlo and Karhulahti[20] enumerate an exhaustive list of types of challenges. In particular, this article pays attention to two of them: endurance and rhythm challenges. The first involves the endurance faculties of the players, and the second requires sustaining a temporally and rhythm in the activities. The adaptation of both types of challenges needs a user behavior categorization based on their traveled distances and the time among data collection check-ins.

One of the characteristics of CLCS users is their traveling behavior. In general, CLCS logs every activity. At least, CLCS record the timestamp of the collection, the user's ID, the item collected and, the location. The users' activity is uneven, with some recording activities regularly and isolated participation. Periodicity, the intensity of activity, and rhythm are elements to be taken into account.

This work focuses on analyzing and categorizing user traveling behaviors (UTBs) as time series. A UTB represents the activities a user performs in a CLCS over time, for example, day by day in a year. Each activity day is categorized as an atom. Then, a year

*Both authors contributed equally to this research.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
Interacción '21, September 22–24, 2021, Málaga, Spain
© 2021 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-7597-9/21/09…$15.00
https://doi.org/10.1145/3471391.3471431

of activities is represented as a stream of atoms. The UTB looks like the ADN sequence of the traveling behavior of a user. All these sequences can be grouped to detect user traveling profiles: the input of tailored challenge gamification.

However, as several users do not participate in a day, the UTBs representation strategy could imply two alternatives. On the one hand, the UTBs could only represent the performed activity logged in the CLCS, called in this work as Raw-UTBs. On the other hand, Timed-UTBs enhance the Raw-UTB information by injecting an atom that represents the null activity in the non-active days (similarly to a place holder or the null object design pattern[12]).

The contribution of this article is to analyze the implications in the categorization of users traveling behaviors with the use of Raw-UTBs or Timed-UTBs. Specifically, this work compares the categorization of UTBs, using machine learning time series clustering, with Raw and Timed UTBs representation. The following questions conducted the research: What information does each approach provide in isolation?, What kind of information do they provide together? Are there similarities between the categorizations?

In the following, Section 2 introduces other related approaches in the literature. Then, Section 3 details both the raw and timed UTBs, and the problem definition. The time series clustering approach is introduced in Section 3.1. A detailed evaluation over a Foursquare dataset is described in Section 4. Then, several discussions of the approach used in adaptive gamification are detailed in Section 5. Finally, Section 6 describe some conclusions and further work.

2 RELATED WORK

The user profile is a central component of information systems such as adaptive systems, and it has been widely studied. Ponciano et al. [17], and Aristeidou et al. [1] worked in profiling the users' motivations and contribution patterns in citizen science projects, looking at engagement metrics. Also, several works have been done specifically with Foursquare datasets to estimate the user's behavior. The work in [15] studies the geo-temporal dynamics of user activity to detect transitions between visited places and identify sequences of activities. Also, mobile users' spatial-temporal activity preference was inferred from the user-generated digital footprints in location-based social networks [23]. Long et al. [11] focus on exploring the local geographic topics using the Latent Dirichlet Allocation (LDA) model to discover the local geographic topics from the check-ins datasets.

To estimate sequence similarity and feature representations for sequence classification and clustering is one of the main tasks of exploratory data mining and is used in many fields such as bioinformatics, pattern recognition, image analysis, or machine learning.

None of the mentioned contributions are related to the user's time series categorization considering the null activity periods, as is introduced in this article.

3 PROBLEM DEFINITION

A **behavioral atom** is a categorical value that describes the user's interaction with the CLCS within a time frame from registered traveling activity about the user. Specifically, the traveled distance, the spent time, and the number of check-ins are three possible metrics to be sampled in each given time frame. Several automatic analyses could categorize the activities based on the three former variables. Also, the CLCS could include a description of the type of traveling activity a user does. This work is based on four activity atoms: Low, Medium, High, and Max. This clusters were obtained from a dataset in [2] which is detailed in Section 4. However, the approach could be generalized to any number of activity atoms.

A **user traveling behavior series (UTB series)** is a sequence of behavioral atoms organized as a time series in chronological order to describe the user's behavior during the check-in dataset period. This UTB series, called Raw-UTBs, can have different sizes, given that users may not be using the application in every time frame sampled. In this work, two approaches to match UTB series of different sizes are analyzed. On the one hand, the Raw-UTBs could be enhanced by injecting an atom representing the null activity in a specific time. These enhanced UTBs are called Timed-UTBs. Both UTBs' representations are described below.

A **Raw-UTB** series is a sequence of behavioral atoms that describe the activities of a user in a time. Each atom describes the behavioral activity in a specific time frame that the user participates in collecting data. However, the days without participation do not appear in the sequence, and the activity is not linked to the specific day in the calendar, making it impossible to identify, for example, the weekends.

A **Timed-UTB** series is an enhanced version of a Raw-UTB where a null activity atom represents the time frames without activity. A null activity atom acts as a placeholder to have the specific time frame sequence. As the raw-UTB, timed-UTB registers the activities sequence. Furthermore, it is possible to know which activity a user performed in a specific time frame, the period of time without activity (like a silence) and, an idea of rhythm or repetition of behavioral patterns. For example, each weekend, a user has a Low, High, and Low activity atoms sequence.

Figure 1 series shows the two versions of a UTB series for a given user. At the top is shown the Raw-UTB series, and at the bottom the Timed-UTB series. The y-axis details the atom: 0 for the null activity atom, from 1 to 4 for Low, Medium, High, and Max level atoms. Figure 1 shows that the Raw-UTB sequence appears with tight lines and much shorter than the second sequence. The graph line of the Timed-UTB shows the series stretched in the period of time and allows us to understand a long period of silence in the user activity.

Although both UTB strategies represent the traveling user behavior, each of them has particular characteristics. In order to describe user behavior categories to offer a tailored gamification strategy, this article seeks to analyze what aspects the different representations can contribute. As either raw and timed-UTB series provides different aspects, the first goal is to understand which characteristics provide each of them in isolation. Then, the second goal is to analyze whether these characteristics can complement each other.

3.1 Approach

From a general perspective, the approach of this article consists of generating from a user activity (check-in) dataset, which includes for each user the atom and the time frame the activity occurs the two sets of UTB series. Figure 2 shows an example of a check-in

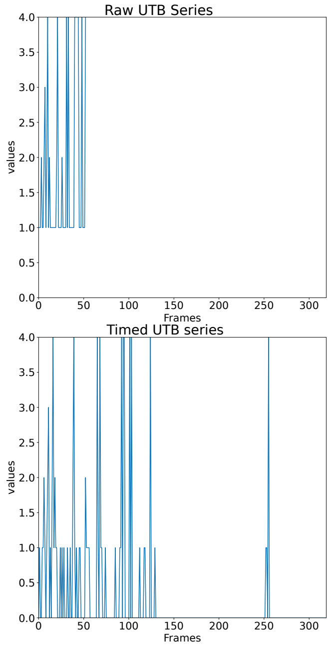

Figure 1: Time series comparison

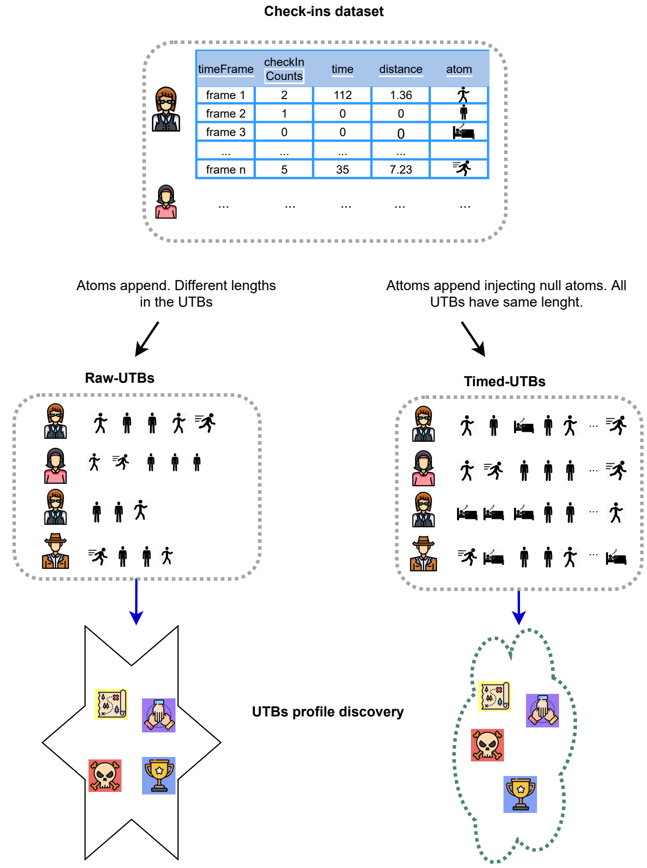

Figure 2: Data transformation into UTBs. It shows the two approaches to be compared. The final sets will be compared

dataset and the transformation into a UTB series set through both mentioned approaches. Then, clustering analysis is performed in each UTB series dataset isolated. The results, illustrated in Figure 2 as a star and a cloud, include the detected UTB profiles. Finally, the results will be compared.

In order to detect traveling user profiles, an unsupervised cluster analysis is presented. The approach proposes analyzing each UTB sequence as a time series and applying a cluster analysis using dynamic time warping (DTW)[4, 14]. DTW is used to compare two temporal series that do not precisely fit. The Dynamic Time Warp Barycenter Averaging (DBA)[16] is proposed to measure the average among the UTB series in a cluster.

The approach is detailed below through a case study using the Foursquare dataset for New York between April 2012 and February 2013.

4 EVALUATION

An evaluation was carried out in order to evaluate the introduced approach. The evaluation will answer the following questions:

(1) What information does each UTB series representation strategy provide in isolation?
(2) Are there similarities between the categorizations?

4.1 Method

In order to answer the former questions, the evaluation is organized in several steps. Firstly, the elbow method curve is analyzed to determine the number of clusters in each UTB dataset. Secondly, the DTW clustering algorithm is run over both the Raw-UTB dataset and Timed-UTB dataset. Thirdly, how the clusters are formed and distributed are plotted and analyzed. This step will answer question 1. Fourthly, a visual comparison between the clusters plots and a similarity analysis among both approaches clusters through Jaccard distance[9] is made to answer question 2.

All the analysis was developed in Kaggle environment[1], using Python language with Pandas and tslearn as main libraries.

4.2 Data

The behavioral atoms dataset [2] has 93,862 items from 1083 users through 318 days between April 2012 and February 2013. Each record is a spatial-temporal activity aggregation for each user each day of the sampled period, classified through a k-means clustering process. Note that the activity aggregation comprises users' check-ins count, the spent time, and the traveled distance. The clustering of these records is shown in Figure 3, showing a precise segmentation of the clusters by the time dimension. Also, Figure 4 shows the cluster's distribution, highlighting the cluster with the lowest intensity.

4.3 Results and analysis

[1]https://www.kaggle.com

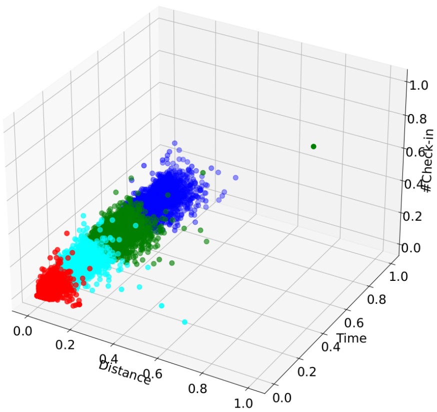

Figure 3: Atom clusters

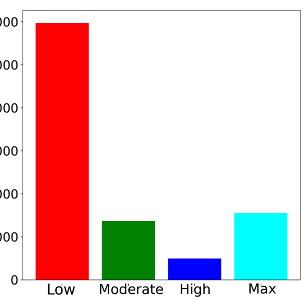

Figure 4: Atom clusters distribution

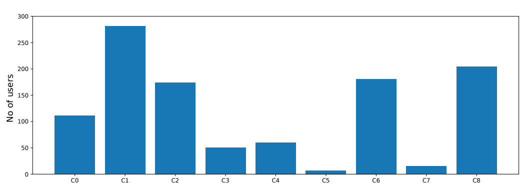

Figure 5: Raw UTBs clusters distribution

4.3.1 Individual strategies analysis. In order to develop the Raw-UTB approach, the UTB series was built from the atoms dataset. After this step, each of the 1083 users is associated with a series of variable sizes. These series were the input for a time series k-means algorithms -from tslean library- using the DTW distance. This user clustering process gave rise to the distribution shown in Figure 5. In addition, the average series with Dynamic Time Warping Barycenter Averaging (DBA is an averaging method that is consistent with Dynamic Time Warping) are described in Figure 6. Notice that this average notion is formed in some quantity by values that are not in the atoms alphabet.

The most significant cluster is *C1*, with 281 users, followed by clusters C8, C6, and C2, having 204, 181, and 174 items, respectively. Looking at the DTW barycenter averages (DBA) of these clusters,

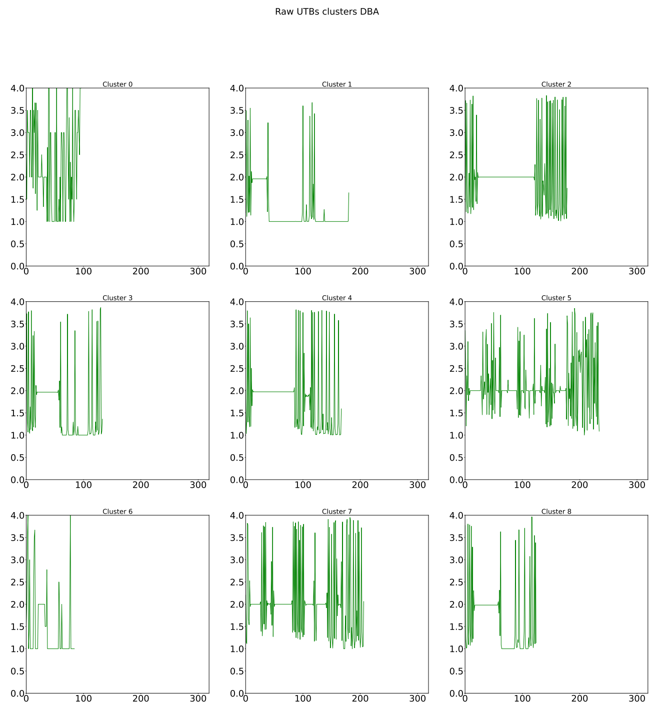

Figure 6: Raw UTBs clusters with DTW Barycenter Average

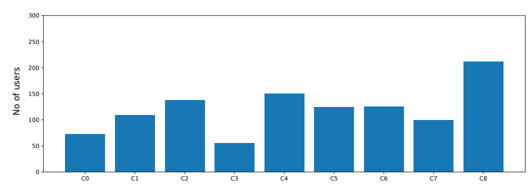

Figure 7: Timed UTBs clusters distribution

most users (77,5 percent) have less than 200 days with registered activity.

Cluster C1 describes an activity of approximately 180 days, starting with a precise sub-period with ups and downs, two extended periods in the low state, and one shorter period in the moderate state.

The users in cluster C8 had an activity of 120 days, starting with intensity peaks, followed by a plateau of about 50 days in the moderate state. Then it falls to the low state, finishing with about 30 days in maximum-low rhythm.

The cluster C6 groups users with the lower general participation, meaning that their UTB series is the shortest, about 90 days, and with a predominant low state. They have some points with higher intensity, but they are low in number. Nevertheless, this cluster reaches the maximum activity level, being the only representative cluster (Cluster 0 does reach this level, but it has very few users) with this feature. It can be identified by *explosive users* because they have some motivation to participate intensely in a short period of time.

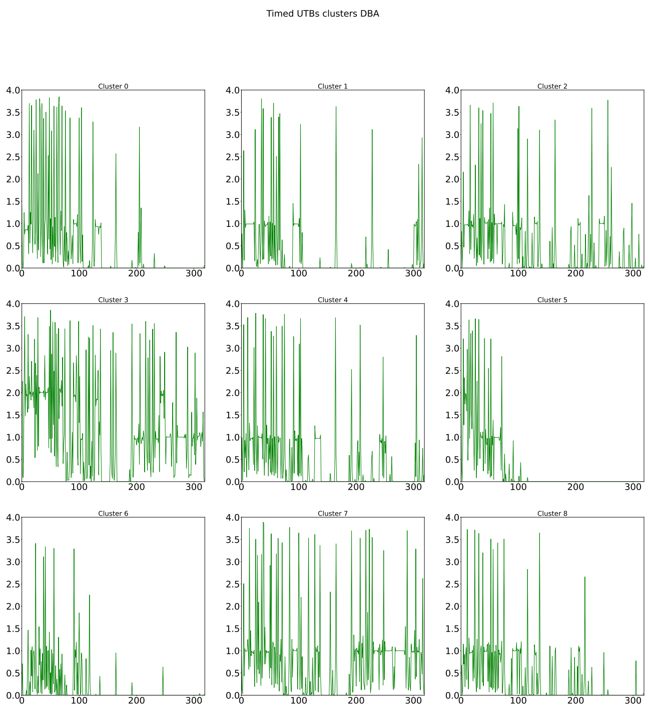

Figure 8: Timed UTBs clusters with DTW Barycenter Average

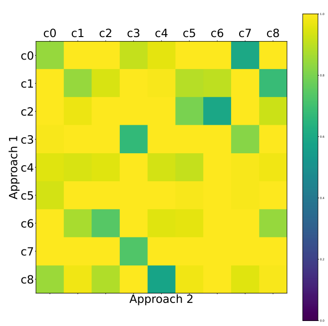

Figure 9: Jaccard distances between clusters of approach 1 and 2

In cluster C2 the number of days with activity is approximately 180 days. A significantly big plateau of 100 days in the moderate state is clearly shown, preceded by an up and downs sequence of 10 days and finishing with another ups and downs sequence of approximately 70 days.

Regarding the Timed-UTB series approach, the temporary frames are fixed depending on the total period of the dataset (or period of analysis). Considering this, after the behavioral atoms sequence is built, those frames that do not have associated activity are filled in with zero values, forming a timed-UTB series for each user. These series were also clusterized with time series k-means, giving place to the distribution shown in Figure 7, and the DBA plot lines shown

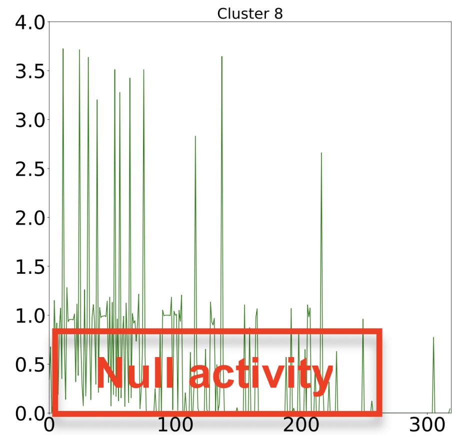

Figure 10: Timed-UTB Cluster 8 detail

in Figure 8. It can be seen that in this case, the clusters' sizes are more similar.

The Timed-UTB approach clearly shows user loyalty. It is possible to see if the activity is maintained during the sampled period or there is no activity. If the lines last until the end of the year, it means that there was activity.

The biggest cluster is cluster 8, with 211 users. It has little activity and the activity peaks decrease over time. In almost the whole sampled period, null activity is shown as a string of ups and downs, probably affected by the injection of null atoms. This idea is depicted in Figure 10.

Cluster 4 has 150 users and has more significant activity than cluster 8, but the first sub-period concentrates most of the activity and falls in the second sub-period.

Cluster 2 has 138 users and generally moderate activity but sustained over time. However, a more intense activity can be noted at the beginning of the sampled period. These users can be identified as *loyal bored users*. The users of clusters 5 and 6 have a common feature: they have a short period of activity during the first 100 days, and then it decreases almost to zero. They show no user loyalty towards the application. In cluster 5, the activity varies among levels moderate and high. It can be explained by users who used the application in a season (*spring break users*) and then abandoned the use of the application. Cluster 6 represents users that participate sporadically in the application, not active users or *lurking users*.

4.3.2 Similarity analysis. To answer the 2nd questions, Jaccard distance is used. The Jaccard distance (Definition 4.1) measures dissimilarity between sample sets with values between 0 (totally dissimilar) and 1 (equal sets).

Definition 4.1. $d_J(A,B) = 1 - J(A,B) = \frac{|A \cup B| - |A \cap B|}{|A \cup B|}$

In this work, the Jaccard distance is used to measure the dissimilarity among clusters in both approaches. Specifically, the similarity is in terms of the users that belong to the clusters. Figure 9 shows a heat matrix with the Jaccard distance values: yellow squares mean a high level of dissimilarity, and dark green squares show some similarity. In general, it can be seen a notable difference among clusters from one strategy to the other. All the Jaccard distance values are significantly low to relate the pair of clusters precisely. Particularly, notice that the most alike clusters are: cluster number 8 (with 204 users) in the Raw-UTB strategy (R8) and cluster 4 (with 150 users) in the Timed-UTB (T4) strategy, with distance: 0.58; followed by cluster number 2 (R2 with 174 users) in Raw-UTB strategy and cluster 6 (T6 with 125 items) in Timed-UTB strategy, with distance: 0.59. Nevertheless, their similarity does not allow further analysis.

On the one hand, R8 and T4 are the most numerous clusters in each strategy, and thus this aspect can make them similar. Hypothetically, it can be thought that the activity plateau of cluster R8 would impact the persistence of values close to low in T4. However, this requires a more detailed analysis.

R2 and T6 are clusters with fewer users, although they are still representative in each strategy. The simultaneous analysis of the DBA graphs indicates that many of the R2 activities have been carried out during the first 100 days of the sampling (Spring season in New York).

5 DISCUSSION

This work analyzed two profiling strategies to build user traveling profiles as input to tailored gamification. It has been shown that clusters in these strategies are dissimilar, so it can be concluded that the users could belong to one profile or another according to the temporal strategy. Therefore these strategies can be seen as different dimensions in user traveling behavior profiles. This idea is in line with the approach of the Hexad player type framework[19].

The Raw-UTB describes in a more qualitative way activity profiles measured in intensity. In other words, groups of users are characterized by sharing max, high, low, and intermediate levels of participation. This strategy could be interpreted to measure stamina or power user profiles. On the other hand, the Timed-UTB approach complements the profiling of users with a 'calendar' viewpoint, allowing to associate the different loyalty intensity levels within a moment of the day, a moment of the week, or the month.

This article introduces the time series analysis as a first step in the search for trends in activity patterns. When a user's behavior matches a pattern, game challenges related to that trend can be proposed.

The introduced approach provides the ability to use user activity patterns as input to adaptive gamification in CLCS. The following steps should include the reaction when a user changes its travel behavior from one pattern to another. In this sense, the Timed-UTB approach could provide the information for both the community and user analysis. Additionally, the travel behavior patterns could be combined with other approaches of adaptive gamification such as Monterrat et al. [13].

6 CONCLUSIONS AND FURTHER WORK

This article analyzes the impact of representing the non-activity in the user traveling behavior categorization in the context of CLCS. Remarkably, the article focused on analyzing travel characteristics to adapt travel challenges in a gamification strategy.

The article introduces Raw-UTBs as time series without no-activity representation and Timed-UTBs, which includes a null atom representing the absence of activity on a specific day. Then, the study of the strategies is analyzed through Dynamic Time Warping clustering analysis.

An evaluation compares the resulting clusters of both strategies by analyzing the plots and distribution and the dissimilarity of clusters with Jaccard Distance.

The results showed that both approaches reflex different aspects of traveling user behavior, and also both could be used in a complementary manner. Users could belong to one profile or another according to the temporal strategy. Indeed, results open the door to new research lines in traveling user types.

Besides the new research lines, a substantial evaluation with different datasets must be performed. Finally, developing a concrete adaptive gamification strategy applying the introduced approach.

REFERENCES

[1] Maria Aristeidou, Eileen Scanlon, and Mike Sharples. 2017. Profiles of engagement in online communities of citizen science participation. *Computers in Human Behavior* 74 (2017), 246–256. https://doi.org/10.1016/j.chb.2017.04.044

[2] Maria Dalponte Ayastuy and Diego Torres. 2021. *Behavioral atoms for NY Foursquare users in 2012.* https://doi.org/10.5281/zenodo.4728128

[3] María Dalponte Ayastuy, Diego Torres, and Alejandro Fernández. 2021. Adaptive gamification in Collaborative systems, a systematic mapping study. *Computer Science Review* 39 (2021), 100333. https://doi.org/10.1016/j.cosrev.2020.100333

[4] Donald J Berndt and James Clifford. 1994. Using dynamic time warping to find patterns in time series.. In *KDD workshop*, Vol. 10. Seattle, WA, USA:, 359–370.

[5] Martin Böckle, Jasminko Novak, and Markus Bick. 2017. Towards adaptive gamification: a synthesis of current developments. In *Proceedings of the 25th European Conference on Information Systems (ECIS)*. Guimarães, Portugal. https://aisel.aisnet.org/ecis2017_rp/11

[6] Sebastian Deterding, Dan Dixon, Rilla Khaled, and Lennart Nacke. 2011. From Game Design Elements to Gamefulness: Defining "Gamification". In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments (MindTrek '11)*. ACM, New York, NY, USA, 9–15. https://doi.org/10.1145/2181037.2181040 event-place: Tampere, Finland.

[7] Stefan Göbel and Viktor Wendel. 2016. *Personalization and Adaptation.* Springer International Publishing, Cham, 161–210. https://doi.org/10.1007/978-3-319-40612-1_7

[8] Sara Iversen. 2012. In the double grip of the game: Challenge and Fallout 3. *Game Studies* 12 (2012). http://www.gamestudies.org/1202/articles/in_the_double_grip_of_the_game

[9] Paul Jaccard. 1912. THE DISTRIBUTION OF THE FLORA IN THE ALPINE ZONE.1. *New Phytologist* 11, 2 (1912), 37–50. https://doi.org/10.1111/j.1469-8137.1912.tb05611.x arXiv:https://nph.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1469-8137.1912.tb05611.x

[10] Jesper Juul. 2011. *Half-real: Video games between real rules and fictional worlds.* MIT press.

[11] Xuelian Long, Lei Jin, and James Joshi. 2012. Exploring Trajectory-Driven Local Geographic Topics in Foursquare. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing* (Pittsburgh, Pennsylvania) *(UbiComp '12)*. Association for Computing Machinery, New York, NY, USA, 927–934. https://doi.org/10.1145/2370216.2370423

[12] James E. McDonough. 2017. *Null Object Pattern.* Apress, Berkeley, CA, 265–269. https://doi.org/10.1007/978-1-4842-2838-8_21

[13] Baptiste Monterrat, Michel Desmarais, Élise Lavoué, and Sébastien George. 2015. A Player Model for Adaptive Gamification in Learning Environments. In *Artificial Intelligence in Education*, Cristina Conati, Neil Heffernan, Antonija Mitrovic, and M. Felisa Verdejo (Eds.). Springer International Publishing, Cham, 297–306.

[14] Meinard Müller. 2007. Dynamic time warping. *Information retrieval for music and motion* (2007), 69–84.

[15] Anastasios Noulas, Salvatore Scellato, Cecilia Mascolo, and Massimiliano Pontil. 2011. An Empirical Study of Geographic User Activity Patterns in Foursquare. *Proceedings of the International AAAI Conference on Web and Social Media* 5, 1 (Jul. 2011). https://ojs.aaai.org/index.php/ICWSM/article/view/14175

[16] François Petitjean, Alain Ketterlin, and Pierre Gançarski. 2011. A global averaging method for dynamic time warping, with applications to clustering. *Pattern Recognition* 44, 3 (2011), 678–693. https://doi.org/10.1016/j.patcog.2010.09.013

[17] Lesandro Ponciano and Thiago Emmanuel Pereira. 2019. Characterising Volunteers' Task Execution Patterns across Projects on Multi-Project Citizen Science Platforms. In *Proceedings of the 18th Brazilian Symposium on Human Factors in Computing Systems* (Vitória, Espírito Santo, Brazil) *(IHC '19)*. Association for Computing Machinery, New York, NY, USA, Article 16, 11 pages.

https://doi.org/10.1145/3357155.3358441

[18] Brian L Sullivan, Jocelyn L Aycrigg, Jessie H Barry, Rick E Bonney, Nicholas Bruns, Caren B Cooper, Theo Damoulas, André A Dhondt, Tom Dietterich, Andrew Farnsworth, et al. 2014. The eBird enterprise: an integrated approach to development and application of citizen science. *Biological Conservation* 169 (2014), 31–40.

[19] Gustavo F. Tondello, Rina R. Wehbe, Lisa Diamond, Marc Busch, Andrzej Marczewski, and Lennart E. Nacke. 2016. The Gamification User Types Hexad Scale. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play* (Austin, Texas, USA) *(CHI PLAY '16)*. Association for Computing Machinery, New York, NY, USA, 229–243. https://doi.org/10.1145/2967934.2968082

[20] Jukka Vahlo and Veli-Matti Karhulahti. 2020. Challenge types in gaming validation of video game challenge inventory (CHA). *International Journal of Human-Computer Studies* 143 (2020), 102473. https://doi.org/10.1016/j.ijhcs.2020.102473

[21] Grant Van Horn, Oisin Mac Aodha, Yang Song, Yin Cui, Chen Sun, Alex Shepard, Hartwig Adam, Pietro Perona, and Serge Belongie. 2018. The INaturalist Species Classification and Detection Dataset. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

[22] Mark JP Wolf and Bernard Perron. 2014. *The Routledge companion to video game studies*. Routledge.

[23] D. Yang, D. Zhang, V. W. Zheng, and Z. Yu. 2015. Modeling User Activity Preference by Leveraging User Spatial Temporal Characteristics in LBSNs. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 45, 1 (2015), 129–142. https://doi.org/10.1109/TSMC.2014.2327053