Towards Factor-oriented Understanding of Video Game Genres using Exploratory Factor Analysis on Steam Game Tags

Xiaozhou Li
Faculty of Information Technology and Communication Sciences, Tampere University, Tampere, Finland
xiaozhou.li@tuni.fi

2020 IEEE International Conference on Progress in Informatics and Computing (PIC) | 978-1-7281-7086-2/20/$31.00 ©2020 IEEE | DOI: 10.1109/PIC50277.2020.9350753

***Abstract*—*Context.* Genre classification, as a tool of information managing, facilitates greatly the users' needs in identifying and retrieving information items of interest within mutually exclusive divisions of collection. Different from that of other products, current video game genre classifications suffer from providing such exclusive divisions. In the domain of game studies, game genre has long been an essential topic, when consensus has not yet been reached. On the other hand, the user-generated tags have been widely adopted enabling end users to annotate and interact with freely. It provides a unique set of crowd-sourced metadata that facilitates the description and understanding of the target products. Such feature is also adopted for video game distribution platforms. *Objective.* Thus, the goal of this work is to investigate the factor-oriented understanding of video game genre classification based on the analysis of large volume of game tag data from Steam platform. It shall contribute largely to the research on game genres within the game studies domain. *Method.* For such purpose, exploratory factor analysis is used herein towards the exploration of latent grouping of game tags seeking the underlying patterns. *Results.* As a result, 29 factors are extracted based on the over 20,000 video game data with 77 gameplay-based tags. It shows that the majority of video games on Steam can be rather seen as a combination of several factors to clear-cut classification.**

***Keywords*—Video game, Game genre, Exploratory factor analysis, Game tag, Steam**

I. Introduction

Genre classification is a normative tool of information managing facilitating the identification, locating and retrieving items of interest given the collections being exclusively divided. When classified properly, genres are seen as the guidance to the audience in information retrieval based on the common understanding of their characteristics [1]. For other media products, e.g., music, literature, movies, etc. their unique observable and objective characteristics, e.g., rhyme of music [2], discourse of literature [3], plot of movies [4] are often used for the genre classification. Towards proper genre classification, mutual exclusivity and joint exhaustivity are seen as the best practice [5]. However, it is barely possible to achieve such clear-cut classification [6]. Thus, flexible approaches in genre classification may be more useful than the strict and rigid taxonomies, according to the "family resemblance" notion by Wittgenstein [7]. Wittgenstein uses game as an example and describe such different similarities and relationships as a complicated network of similarities overlapping and "criss-crossing" [8]. Similar to other traditional media products, video games genre classification is a common practice in the game industry and has been long studied in the academia. One of the earliest studies on game genre is the taxonomy given by Chris Crawford that classify computer games into skill and action games and strategy games [9]. Such classification is based on the perceptual and cognitive skills of the players required for winning the games. Comparatively, another "overly-detailed" 42 categories of games are given by Mark J.P. Wolf based on the gameplay and interactivity [10]. Towards the similar aim, many studies have proposed multiple set of game genres in order to achieve mutual exclusivity and joint exhaustivity [11]–[14]. On the other hand, other studies focus on the structural perspectives and provide meta-categories of video games [15], [16]. Even until the recent, game studies in general investigate game genre from either an inductive or a deductive perspective, which, however, have not reached a consensual solution [17]. Meanwhile, such strict and rigid taxonomies on video game genre classification are increasingly challenged by the rapid evolution of video games together with the blending of multiple genre elements, where a more flexible approach could be useful [18]. Furthermore, concurring with Wittgenstein's statement and taking into account the overlapping similarity amongst contemporary video games, the study of Li and Zhang provides a network-oriented understanding of game genre classification [19].

The rapid-developing data mining and machine learning techniques provide valuable toolkit on solving the classification problems. For example, many studies have used the computational approaches to provide solutions towards music and movie genre classification [20]–[23]. Accordingly, the metadata of video games and players becomes increasingly important facilitating various studies towards understanding games and player behaviors [24]–[26]. However, limited studies have attempted such approaches on game genre classification. One example of using data mining techniques for game genre classification is Faisal and Peltoniemi's study [27]. They adopt Latent Dirichlet Allocation (LDA) topic modeling technique on game descriptions from online game databases and obtain 31 topics/genres. Suatap and Patanukhom propose the use of Convolutional Neural Networks (CNN) to classify games in pre-defined genres based on their icons and screenshots [28]. However, it does not contribute to the understanding of game genre but focuses on the validation of the method. On the other hand, many video game related data are used in academic research, such as, community data [29], profile and gameplay behavior data [25], review data [30], playtime data [24], and so on. The user generated tags, which allow end users to annotate and interact with information objects freely, provide a unique set of crowd-sourced metadata that facilitates the description and understanding of such information objects with decent agreement level to formal taxonomies [31], [32]. Specifically towards video games, Steam platform provides a unique user tagging feature enabling players to denote some aspects of the "aboutness" of the games [33]. Thereby, such game tag data shows great potential towards an alternative understanding of game genres from a collective intelligence perspective via inter-correlation of such tags provided.

Hence, in this study, we use the user-generated game tags metadata from Steam platform and apply factor analysis in order to detect the latent connections amongst them towards a factor-oriented game genre understanding. The result of this study shall answer the following research question: *What are the factors determining the genre of a particular video games?* The main contribution of this work includes: an introduction of data-driven approaches to solve a long-existing domain specific research question in game studies, a novel way of understanding game genre that differs greatly from the previous works in game studies. Compared to the work of using game tags data for network analysis [19], factor analysis shall result in a more practical genre classification framework that can be adopted easily supporting information management. Herein, using factor analysis and focusing on gameplay-related tags, we aim to extract a set of factors, each of which contains a set of inter-correlated tags, by which the factor can be interpreted in terms of gameplay.

The remainder of the article is organized as follows. Section 2 presents the related works in video game genre studies and the studies using Steam platform as data source. Section 3 describes the data used in this study and introduces factor analysis method adopted. Section 4 presents the results of factor analysis with factor interpretation and further discusses relevant issues. Section 5 concludes the article.

II. Related Work

The most common characteristic selected for video games genre classification is the gameplay. The gameplay is the all-important quality factor and the pure interactivity of the games [34]. Many game genre classification are proposed focusing on the gameplay that the designer intend to advocate in particular games. The different abstraction levels of the understanding of gameplay can lead to various genre classification outcomes, for example, the "skill-and-action" and "strategy" genres by Chris Crawford [9] or the 42 genres by Mark J.P. Wolf [10]. These early studies and genre definitions are critiqued by the later for not being able to accommodate new genres, e.g., multiplayer online battle arena (MOBA) [35]. Many other studies attempt to contribute in the middle ground and still pursuing to the best practice, i.e., mutual exclusivity and joint exhaustivity, while also reasoning their proposal with the focus on different perspectives. Therein, Apperley focus on players' interaction and relations with the the game genres and propose four commonly used game labels: Simulation, Strategy, Action and Role-playing [13]. A 4-tiered hierarchy emphasizing interactivity rather than narrative is proposed by King and Krzywinska indicating genre is only one of the tiers for classification but other factors such as, platform, mode, and Milieu shall also be considered [36] when it is also critiqued by Whalen indicating such hierarchy falls short in creating a common language [12]. Both Whalen and Clearwater indicate that game elements occur simultaneously instead of hierarchically [14]. Lee et al. use facet analysis and propose a 12-facet-and-358-foci scheme describing the game genre information when emphasizing such method shall represent the different types of information embedded in current video game genre labels in a flexible and extensible way [35]. Vargas-Iglesias focuses on the different functions within video games and presents four elemental genres and a way of configuring hybrid genres with binary function relations [17]. On the other hand, for understanding game genres in regarding the communicative purposes and actions with both purposes and elements of form [37], Bogost indicates that game genres are classified by the players through the on-screen effects and controllable dynamics they experience [38]. On the other hand, other scholars also suggest to understand game genre as a socially constructed phenomena instead of a clear-cut taxonomy. Clearwater indicates that formal and aesthetic considerations, industrial and discursive context, and social meaning and cultural practice shall all be taken into account regarding video game genres [14]. Arsenault also emphasizes that the genre of a game is tied not to a checklist of features, but to the phenomenological, pragmatic deployment of actions through the gameplay experience, as gameplay is both functional and aesthetic [39].

Despite the studies in video game genre classification that provides various classification outcomes, the issues of such classification being rigid and falling short at contributing to the original purpose of classification persist. Considering the primary purpose of genre classification being to help users find similar items, the previous classifications perform ineffectively when games with different characteristics located in same genre lacking concrete identification criteria [18]. Such lack of concrete definition then results in the heavily overloaded genre labels that representing multi-dimensional information [35]. In the discussion of his study, Apperley also emphasizes that it is crucially important to video game genres that the notion of each individual game belonging to several genres at once being taken into account [13]. Towards such end, metadata are used to describe video games and interactive media. Lee et al. introduce a schema of 16 elements to formally describe video games based on a user-centered design approach [40]. Despite the importance of the works towards preserving and retrieving video game legacy as well as the usefulness of metadata reflecting the characteristics of video games [41], it falls short on providing approaches towards game genre classification based on the analysis of such metadata. Li and Zhang provide an alternative network-oriented understanding of game genre classification using the user generated game tag metadata [19]. However, it falls short in providing practical guidelines on how to apply the genre network into the information retrieval practice of video games.

Steam, as one of the most popular digital game distribution platforms, has drawn attention from the academia. Variety of video game data can be obtained freely and anonymously via the official API provided by Steam as well as from other web crawling tools. Targeting the studies on multiple aspects of the Steam community, video games and the players, scholars have published multiple valuable work applying data-driven approaches on Steam data. For example, Lim and Harrell examine the players' social identity, as well as the connection between players' behaviors on maintaining their profiles and their social network [25]. Regarding player behaviors, Sifa et al. analyze the players' different playtime frequency distribution and investigate their engagement and cross-game behavior on Steam [24], [42]. Slivar et al. analyze the the impact of game types and video adaptation strategies on the quality of experience via a case study on Steam platform [26]. Lin et al. focus on the video game development and maintenance practices and analyze the urgent update strategy of popular games on Steam [43]. Windleharth et al., conduct a conceptual analysis on the user-generated game tags on Steam and propose a categorization of them according to the Video Game Metadata Schema (VGMS) category [33], [44]. Furthermore, other perspectives regarding player communities [45], game reviews [46], game recommendation mechanisms [47], and etc. have also been studied using the data obtained from Steam. Specifically using game tag data from Steam, Li and Zhang use social network analysis and provide a network-oriented understanding of game genre classification.

III. METHOD

A. Data

To obtain the game tag data from Steam, we use Scrapy [1], an open-source Python-written web crawling framework. Scrapy was first released in 2008 with its latest version 2.4.0 being compatible with Python 3.6 and later versions. In order to crawl structured review data from the store page of a particular game, a Spider is defined with Scrapy and ran through the crawler engine. The crawling process starts with a request on the URL of the game store page and calls the default callback method looping through the elements with CSS selectors and yielding a dictionary with the requested information. Scrapy is able to crawl the content loaded with Javascript via users' scrolling that cannot be obtained using BeautifulSoup[2], which is crucial for crawling the data.

I choose Steam platform as the source of data crawling as it contains the largest video game collection compared with other game platforms on PC, e.g., Origin, GOG, Uplay, and so on. On the other hand, console-exclusive games are of limited amount compared to the volume of games on Steam, when the majority of games on consoles are also published on Steam. On the other hand, despite the mobility of mobile devices and unique ways of interaction for mobile players [48], mobile platform contains limited unique game genres that unique to the devices [49]. Thus, the data obtained from Steam platform shall be seen statistically representative for PC and console based video games.

In this study, 23,034 random video games on Steam platform are crawled. The release years of the crawled games range from 1985 to 2020. For majority of the selected games, their top 20 user generated tags are crawled. Some particular games contain less than 20 tags (for example, *WWE 2K19*[3]) when only the given number of tags are crawled. On the other hand, 11 games were not given any tags when crawled which are eliminated from the dataset. According to the user voting mechanism of game tags, the tag with more user voting for a particular game shall appears ahead of the other tags of this game. As the detailed voting numbers are inaccessible, a weighted value (20 to 1) is assigned to each tag of a game by its rank, where the first tag is weighed as 20 with the last as 1. Similarly, for the games with less than 20 tags, the nth tag is weighed as $20 - n + 1$. Furthermore, due to the variety of themes in the game tags, only gameplay-related tags are selected as gameplay are the essential factor determining the playability of games. Using the VGMS [44] and the categorization of Windleharth et al. [33], 93 innitial gameplay tags are selected out of the original 374. The data items that contain no gameplay tags are deleted from the dataset making 22,749 games remain.

B. Exploratory Factor Analysis

To uncover the underlying structures of the Steam user profiles, an exploratory factor analysis (EFA, [50]) is conducted. Factor analysis in general aims to detect the latent variables sharing a common variance and being unable to be observed [51]. Such variables are obtained from a larger set of measurable and observable ones. These detected factors are not directly measured but are essentially hypothetical constructs that are used to represent variables [52]. Specially, EFA aims to discover not only the number of factors, but also what measurable variables together influence which individual factors [53]. It enables the reduction of complexity of the data, and explains the observations with a smaller set of latent factors and discover the relations between variables. Therefore, regarding the case of game tag data herein, EFA shall discover the underlying factors characterizing game genres and their differences with the influencing tag-variables for each factor.

Herein, parallel analysis (PA) [54] is applied in order to find the number of factors. In PA, the Monte Carlo simulation technique is employed to simulate random samples consisting of uncorrelated variables that parallel the number of samples and variables in the observed data. From each such simulation, eigenvalues of the correlation matrix of the simulated

[1]https://scrapy.org/
[2]https://www.crummy.com/software/BeautifulSoup/
[3]https://store.steampowered.com/app/817130/WWE_2K19/

data are extracted, and the eigenvalues are averaged across several simulations [54]. The eigenvalues extracted from the correlation matrix of the observed data, ordered by magnitude, are then compared to the average simulated eigenvalues, also ordered by magnitude. The decision criteria is that the factors with observed eigenvalues higher than the corresponding simulated eigenvalues are considered significant. To simplify interpretation of the factor analysis result, the *varimax* rotation technique [55] is employed to maximize the variance of the each factor loading.

IV. Results

A. Factor Analysis

To verify the sampling adequacy, both Bartlett's Test of Sphericity [56] and Kaiser-Meyer-Olkin (KMO) Test [57] are applied. In Bartlett 's Test of Sphericity, the p-value is 0.0 (Approx. Chi-Square = 250978.05, df = 4278), indicating the test being statistically significant and the observed correlation matrix being not an identity matrix. And the KMO test measures the suitability of data for factor analysis with its value estimating the proportion of variance among all the observed variable. Herein, the KMO score is 0.661, which shows the adequacy of the sampling and the usefulness of applying factor analysis.

When applying PA to the dataset of 22,749 games with 93 gameplay tags, 36 factors are obtained (Observed Eigenvalue $>$ 1.0). Therein, 14 tags of the 93 had correlation coefficients ranging from –.5444 to .1932. These tags include *'Action-Adventure'*, *'Arena Shooter'*, *'Bowling'*, *'Casual'*, *'God Game'*, *'Hidden Object'*, *'Lemmings'*, *'MOBA'*, *'On-Rails Shooter'*, *'Physics'*, *'Pinball'*, *'Sniper'*, *'Space Sim'*, and *'Word Game'*. It implies that these tags are either very specific on gameplay and hard to be applied together with other tags (e.g., *Hidden Object* and *Pinball*) or so general that can be applied to a wide range of tags (e.g., *Casual* and *Action-Adventure*). These tags are deleted with 79 tags retained while 21,874 games are retained as the ones with only tags of these tags are also deleted. By applying the same process again on the newly obtained dataset, two other tags, *'Puzzle'* and *'Chess'*, are also identified having low correlation ranging from –.5965 to .1375. They are also deleted with also the games with only these two tags. Thereafter, 21,298 games are retained with 77 gameplay tags. The sampling remains adequate, as p-value of the Bartlett 's Test of Sphericity is 0.0 (Approx. Chi-Square = 217420.84, df = 2926) when KMO value remains 0.661. By applying PA to the refined dataset, 29 factors are obtained with the corresponding factor loadings shown in Fig. 1.

B. Factor Interpretation

Based on the result of EFA, I interpret each of the 29 factors into a game genre factor based on the correlated game tags.

The results of such factor interpretation show that in general a particular video game, except for the ones that fit only in the 16 gameplay tags that are deleted, can always be seen as belonging to one single genre/factor or a combination of several factors out of the 29 with various degrees. For example, Grand Theft Auto V[4] shall be labeled as a *Shooter* game. However, on the other hand, the *Racing* factor obviously plays a critical part in terms of its gameplay. Another example shall be Divinity: Original Sin 2[5], as it belongs to the *RPG* genre and contains factor of *Turn-based Tactics* as well. Thus, using such factor-oriented genre classification, each video game can be described as a list of factors (e.g., GTA V: ['Shooter', 'RPG', 'Racing']). Provided the voting value of each game tag is accessible, the importance of the factors can then be quantified.

The results of EFA show that, based on the analysis of large volume of video games with the according user generated game tags, it is reasonable to understand video game genre classification from a factor-oriented perspective, instead of adopting clear-cut classification. Arguably, it is still possible to classify each individual video game into a single genre based on the primary factor of it, provided the voting values of each game tags are accessible with the importance of each factor quantified. However, considering the goal of genre classification being facilitating customers searching, browsing, locating and retrieving media items, based on the understanding on their characteristics [1], the factor-oriented understanding of game genre shall perform more accurately targeting such customer needs. Nonetheless, the similarity between any two games can be calculated by the Jaccard similarity [58] or KH similarity [59] of their factors, which shall ease the customers' effort in searching for related games.

C. Further Discussion

On the other hand, due to the fact that each game contains very limited number of gameplay related tags, the majority of the dataset are zeros. Therefore, though the sampling is adequate for factor analysis, it shall be emphasized the results can show differences and shall be interpreted differently compared to those obtained from applying EFA on survey data. One of the significant difference is the interpretation of the deleted items (i.e., game tags). In the studies dealing with survey data, such as [60], the variables are deleted due to the lack of correlation. Comparatively, for this study, the deleted tags are the ones being too generally connecting to a number of tags or too specific with very limited ones. It does not necessarily mean they are not part of the genre classification. For example, *Hidden Object* is a well-known gameplay style and a unique game genre. However, it is so independent from other types of gameplay that it fails to connect to any other gameplay tags. On the contrary, *'Casual'* describes games with less complicated controls and complexity in gameplay [61]. It is hard to fit casual games into a particular genre based on gameplay, as it can be understood from various perspectives such as gameplay, attitudes, and etc. [62], [63]. Thus, this tag can be connected to any other gameplay tags. In addition, another difference is that some of the 29 factors contain low

[4] https://store.steampowered.com/app/271590/
[5] https://store.steampowered.com/app/435150/

Fig. 1. Final Factor Loadings

internal consistency based on the calculation of Cronbach's coefficient alpha [64]. For example, the alpha values for Factor 3 and 27 are respectively .287 and .168. The reason for such phenomena is likely due to the low volume of cases supporting the correlation or broad connection of particular tags within the factor to the ones in other factors. Regarding the example, only 49 games selected contain the tag 'Soccer' while 62 the tag 'Football'. Meanwhile, over 10,000 games contain the tag 'Adventure' connecting to nearly all the gameplay tags. To be emphasized, such phenomenon does not necessarily indicate the failure of applying EFA or the lack of trustworthiness of the obtained factors, but shows the latent tendency of tag connections. After all, pursuing for high alpha value in EFA is arguably a common misconception [65]. Specially for the game tag data, it is reasonable to claim that meaningful interpretation can be drawn herein.

V. Conclusion

This study proposes the use of EFA on large volume of game tag data towards a factor-oriented understanding of video game genre classification. Based on the results of EFA, 29 factors are obtained indicating for the vast majority of the video games, each of which can be described with a list of selected factors in terms of its gameplay. The contribution of this work is to enrich the understanding of video game genre in the game studies domain with a novel factor-oriented genre classification and to introduce a data-driven approach of tackling research questions in game studies domain. The result of this study provides a easy-to-adopt middle ground between clear-cut genre classification and network-oriented understanding. It shall facilitate not only the information management practice of video game industry but also the theoretical foundation of game studies. The future works of this study shall include the investigation of methods in automatically identifying genre factors for any particular game, as well as the study on the design of recommending mechanisms based on such factor analysis results.

References

[1] L. M. Chan and A. Salaba, *Cataloging and classification: an introduction*. Rowman & Littlefield, 2015.

[2] R. Mayer, R. Neumayer, and A. Rauber, "Rhyme and style features for musical genre classification by song lyrics." in *Ismir*, 2008, pp. 337–342.

[3] G. Steen, "Genres of discourse and the definition of literature," *Discourse Processes*, vol. 28, no. 2, pp. 109–120, 1999.

[4] P. Cook, *The cinema book*. British Film Institute, 2007.

TABLE I
GAME GENRE FACTORS

Factor	Genre Factor	Correlated Gameplay Tags
1	Arcade Shooter	['Arcade', 'Bullet Hell', "Shoot 'Em Up", 'Shooter']
2	Fighting	['2D Fighter', "Beat 'em up", 'Fighting']
3	Soccer	['Football', 'Soccer', 'Sports']
4	Sandbox	['Base Building', 'Building', 'Crafting', 'Sandbox', 'Survival']
5	Turn-based Tactics	['Tactical', 'Tactical RPG', 'Turn-Based Combat', 'Turn-Based Strategy', 'Turn-Based Tactics']
6	Resource Management	['Base Building', 'Building', 'City Builder', 'Management', 'Resource Management', 'Strategy']
7	Music	['Music', 'Rhythm']
8	Shooter	['Action', 'FPS', 'Shooter', 'Third-Person Shooter']
9	Rogue	['Dungeon Crawler', 'Rogue-like', 'Rogue-lite']
10	Character Action	['Character Action Game', 'Hack and Slash', 'Spectacle fighter']
11	Strategy	['4X', 'Grand Strategy', 'Strategy', 'Turn-Based Strategy', 'Wargame']
12	Classic	['Classic', 'Cult Classic']
13	Board/Card	['Board Game', 'Card Game', 'Trading Card Game']
14	Gal Game	['Dating Sim', 'Otome', 'Visual Novel']
15	Platformer	['Metroidvania', 'Platformer', 'Puzzle-Platformer', 'Side Scroller']
16	Survival	['Crafting', 'Survival', 'Survival Horror']
17	MMO	['Massively Multiplayer', 'MMORPG']
18	Racing	['Driving', 'Offroad', 'Racing', 'Sports']
19	RPG	['Dungeon Crawler', 'JRPG', 'Party-Based RPG', 'RPG', 'Turn-Based Combat']
20	Strategy RPG	['Strategy RPG', 'Tactical RPG', 'Turn-Based Tactics']
21	ARPG	['Action RPG', 'Hack and Slash']
22	Top-down Shooter	['Top-Down Shooter', 'Twin Stick Shooter']
23	Interactive Fiction	['Choose Your Own Adventure', 'Interactive Fiction']
24	Tower Defense	['RTS', 'Strategy', 'Tower Defense']
25	RTS	['Real Time Tactics', 'RTS', 'Tactical']
26	Simulation	['Management', 'Simulation', 'Strategy']
27	Exploration	['Adventure', 'Exploration', 'Walking Simulator']
28	Parkour	['3D Platformer', 'Parkour', 'Platformer', 'Puzzle-Platformer']
29	Education	['Education', 'Programming']

[5] C. Beghtol, "The concept of genre and its characteristics," *Bulletin of the American Society for Information Science and Technology*, vol. 27, no. 2, pp. 17–17, 2000.

[6] K. P. Jones, "The environment of classification: the concept of mutual exclusivity," *Journal of the American Society for Information Science*, vol. 24, no. 2, pp. 157–163, 1973.

[7] L. Wittgenstein, "Philosophical investigations. philosophische untersuchungen." 1953.

[8] A. Manser, "Games and family resemblances," *Philosophy*, vol. 42, no. 161, p. 210–225, 1967.

[9] C. Crawford, "The art of computer game design," 1984.

[10] M. J. Wolf, "Genre and the video game," *The medium of the video game*, pp. 113–134, 2001.

[11] E. Aarseth, S. M. Smedstad, and L. Sunnanå, "A multidimensional typology of games." in *DiGRA Conference*, 2003.

[12] Z. Whalen, "Game/genre: A critique of generic formulas in video games in the context of "the real"," *Works and Days*, vol. 22, no. 43/44, pp. 289–303, 2004.

[13] T. H. Apperley, "Genre and game studies: Toward a critical approach to video game genres," *Simulation & Gaming*, vol. 37, no. 1, pp. 6–23, 2006.

[14] D. Clearwater, "What defines video game genre? thinking about genre study after the great divide," *Loading...*, vol. 5, no. 8, 2011.

[15] C. Elverdam and E. Aarseth, "Game classification and game design: Construction through critical analysis," *Games and Culture*, vol. 2, no. 1, pp. 3–22, 2007.

[16] A. Järvinen, *Games without frontiers: Theories and methods for game studies and design*. Tampere University Press, 2008.

[17] J. J. Vargas-Iglesias, "Making sense of genre: The logic of video game genre organization," *Games and Culture*, p. 1555412017751803, 2018.

[18] R. I. Clarke, J. H. Lee, and N. Clark, "Why video game genres fail: A classificatory analysis," *Games and Culture*, vol. 12, no. 5, pp. 445–465, 2017.

[19] X. Li and B. Zhang, "A preliminary network analysis on steam game tags: another way of understanding game genres," in *Proceedings of the 23rd International Conference on Academic Mindtrek*, 2020, pp. 65–73.

[20] C. Xu, N. C. Maddage, X. Shao, F. Cao, and Q. Tian, "Musical genre classification using support vector machines," in *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03).*, vol. 5. IEEE, 2003, pp. V–429.

[21] C. N. Silla, A. L. Koerich, and C. A. Kaestner, "A machine learning approach to automatic music genre classification," *Journal of the Brazilian Computer Society*, vol. 14, no. 3, pp. 7–18, 2008.

[22] H. Zhou, T. Hermans, A. V. Karandikar, and J. M. Rehg, "Movie genre classification via scene categorization," in *Proceedings of the 18th ACM international conference on Multimedia*. ACM, 2010, pp. 747–750.

[23] G. S. Simões, J. Wehrmann, R. C. Barros, and D. D. Ruiz, "Movie genre classification with convolutional neural networks," in *2016 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2016, pp. 259–266.

[24] R. Sifa, A. Drachen, and C. Bauckhage, "Large-scale cross-game player behavior analysis on steam," *Borderlands*, vol. 2, pp. 46–378, 2015.

[25] C.-U. Lim and D. F. Harrell, "Developing social identity models of players from game telemetry data." in *AIIDE*, 2014.

[26] I. Slivar, M. Suznjevic, and L. Skorin-Kapov, "The impact of video encoding parameters and game type on qoe for cloud gaming: A case study using the steam platform," in *2015 Seventh International Workshop on Quality of Multimedia Experience (QoMEX)*. IEEE, 2015, pp. 1–6.

[27] A. Faisal and M. Peltoniemi, "Establishing video game genres using data-driven modeling and product databases," *Games and Culture*, vol. 13, no. 1, pp. 20–43, 2018.

[28] C. Suatap and K. Patanukhom, "Game genre classification from icon and screenshot images using convolutional neural networks," in *Proceedings of the 2019 2nd Artificial Intelligence and Cloud Computing Conference*, 2019, pp. 51–58.

[29] X. Li, C. Lu, J. Peltonen, and Z. Zhang, "A statistical analysis of steam user profiles towards personalized gamification," in *3rd International GamiFIN Conference, GamiFIN 2019*. CEUR-WS, 2019.

[30] D. Lin, C.-P. Bezemer, Y. Zou, and A. E. Hassan, "An empirical study of game reviews on the steam platform," *Empirical Software Engineering*, vol. 24, no. 1, pp. 170–207, 2019.

[31] S. A. Golder and B. A. Huberman, "Usage patterns of collaborative

tagging systems," *Journal of information science*, vol. 32, no. 2, pp. 198–208, 2006.

[32] C. Veres, "The language of folksonomies: What tags reveal about user classification," in *International Conference on Application of Natural Language to Information Systems*. Springer, 2006, pp. 58–69.

[33] T. W. Windleharth, J. Jett, M. Schmalz, and J. H. Lee, "Full steam ahead: A conceptual analysis of user-supplied tags on steam," *Cataloging & Classification Quarterly*, vol. 54, no. 7, pp. 418–441, 2016.

[34] J. Juul, *Half-real: Video games between real rules and fictional worlds*. MIT press, 2011.

[35] J. H. Lee, N. Karlova, R. I. Clarke, K. Thornton, and A. Perti, "Facet analysis of video game genres," *iConference 2014 Proceedings*, 2014.

[36] G. King and T. Krzywinska, "Computer games/cinema/interfaces." in *CGDC Conf.*, 2002.

[37] J. Swales, *Genre analysis: English in academic and research settings*. Cambridge University Press, 1990.

[38] I. Bogost, *Persuasive games*. Cambridge, MA: MIT Press, 2007, vol. 5.

[39] D. Arsenault, "Video game genre, evolution and innovation," *Eludamos. Journal for Computer Game Culture*, vol. 3, no. 2, pp. 149–176, 2009.

[40] J. H. Lee, J. T. Tennis, R. I. Clarke, and M. Carpenter, "Developing a video game metadata schema for the seattle interactive media museum," *International journal on digital libraries*, vol. 13, no. 2, pp. 105–117, 2013.

[41] J. H. Lee, R. I. Clarke, and A. Perti, "Empirical evaluation of metadata for video games and interactive media," *Journal of the Association for Information Science and Technology*, vol. 66, no. 12, pp. 2609–2625, 2015.

[42] R. Sifa, C. Bauckhage, and A. Drachen, "The playtime principle: Large-scale cross-games interest modeling." in *CIG*, 2014, pp. 1–8.

[43] D. Lin, C.-P. Bezemer, and A. E. Hassan, "Studying the urgent updates of popular games on the steam platform," *Empirical Software Engineering*, vol. 22, no. 4, pp. 2095–2126, 2017.

[44] J. H. Lee, A. Perti, R. I. Clarke, T. W. Windleharth, and M. Schmalz. (2017) Uw/simm video game metadata schema version 4.0. [Online]. Available: https://cpb-us-e1.wpmucdn.com/sites.uw.edu/dist/2/3760/files/2019/09/ VGMSVersion4.0_20180824.pdf

[45] R. Becker, Y. Chernihov, Y. Shavitt, and N. Zilberman, "An analysis of the steam community network evolution," in *Electrical & Electronics Engineers in Israel (IEEEI), 2012 IEEE 27th Convention of*. IEEE, 2012, pp. 1–5.

[46] D. Lin, C.-P. Bezemer, Y. Zou, and A. E. Hassan, "An empirical study of game reviews on the steam platform," *Empirical Software Engineering*, pp. 1–38, 2018.

[47] H.-C. Yang and Z.-R. Huang, "Mining personality traits from social messages for game recommender systems," *Knowledge-Based Systems*, 2018.

[48] X. Li and Z. Zhang, "A user-app interaction reference model for mobility requirements analysis," in *ICSEA 2015, The Tenth International Conference on Software Engineering Advances*, 2015, pp. 170–177. [Online]. Available: https://academic.microsoft.com/paper/2604476080

[49] F. Mäyrä, "Mobile games," *The International Encyclopedia of Digital Communication and Society*, pp. 1–6, 2015.

[50] J. F. Hair, W. C. Black, B. J. Babin, R. E. Anderson, R. L. Tatham *et al.*, "Multivariate data analysis (vol. 6)," 2006.

[51] D. J. Bartholomew, M. Knott, and I. Moustaki, *Latent variable models and factor analysis: A unified approach*. John Wiley & Sons, 2011, vol. 904.

[52] R. B. Cattell, "The meaning and strategic use of factor analysis," in *Handbook of multivariate experimental psychology*. Springer, 1988, pp. 131–203.

[53] J. DeCoster, "Overview of factor analysis," 1998.

[54] J. L. Horn, "A rationale and test for the number of factors in factor analysis," *Psychometrika*, vol. 30, no. 2, pp. 179–185, 1965.

[55] H. F. Kaiser, "The varimax criterion for analytic rotation in factor analysis," *Psychometrika*, vol. 23, no. 3, pp. 187–200, 1958.

[56] G. W. Snedecor and W. G. Cochran, "Statistical methods, 8thedn," *Ames: Iowa State Univ. Press Iowa*, vol. 54, pp. 71–82, 1989.

[57] H. F. Kaiser, "An index of factorial simplicity," *Psychometrika*, vol. 39, no. 1, pp. 31–36, 1974.

[58] P. Jaccard, "Étude comparative de la distribution florale dans une portion des alpes et des jura," *Bull Soc Vaudoise Sci Nat*, vol. 37, pp. 547–579, 1901.

[59] B. V. Kumar and L. Hassebrook, "Performance measures for correlation filters," *Applied optics*, vol. 29, no. 20, pp. 2997–3006, 1990.

[60] C. S.-Y. Park, S. L. Yoon, S.-N. Yun, and E. Park, "Korean patient-perceived satisfaction scale of community-based case management services (korean-psccm): Development and psychometric evaluation," *Journal of Community Health Nursing*, vol. 34, no. 1, pp. 32–45, 2017.

[61] M. Wallace and B. Robbins, "Casual games white paper," *IGDA Casual Games SIG, http://www. igda. org/casual/IGDA_CasualGames_Whitepaper_2006. pdf (accessed April 9, 2008)*, 2006.

[62] J. Kuittinen, A. Kultima, J. Niemelä, and J. Paavilainen, "Casual games discussion," in *Proceedings of the 2007 conference on Future Play*. ACM, 2007, pp. 105–112.

[63] A. Kultima, "Casual game design values," in *Proceedings of the 13th international MindTrek conference: Everyday life in the ubiquitous era*. ACM, 2009, pp. 58–65.

[64] L. J. Cronbach, "Coefficient alpha and the internal structure of tests," *psychometrika*, vol. 16, no. 3, pp. 297–334, 1951.

[65] E. Cho and S. Kim, "Cronbach's coefficient alpha: Well known but poorly understood," *Organizational Research Methods*, vol. 18, no. 2, pp. 207–230, 2015.