Application of K-means clustering algorithm in the analysis of college students' online entertainment consumption

Huanqin Li*

1Zhengzhou Normal University, Zhengzhou, Henan, 450044, China
*Corresponding author’s e-mail: Sz2008peak@163.com

Abstract: It is benefit for college students to explore the characteristics of online entertainment consumption and give them appropriate guidance. The Online entertainment consumption of some undergraduate students is investigated in this paper, and the K-means algorithm is used in the survey to carry out clustering analysis. The results show that in terms of online entertainment consumer behaviour, the consumption amount is polarized, consumption is gender-differentiated, consumer contents are diversified, consumer has more copyrighted-awareness., etc. From the clustering results, the K-means algorithm is more effective for analyzing the characteristics of online entertainment consumption behavior.

1.Introduction
Entertainment is an important function of the Internet. Proper online entertainment has a positive effect on regulating people's life rhythm, alleviating the stress of study or work, enhancing the interest of life, and improving the quality of life. For college students, the Internet mainly plays an entertaining role, which is mainly reflected in: listening to music, playing online games, reading news, downloading pictures, and so on. However, they should correctly understand the functions of online entertainment and prevent indulging in “pan-entertaining”, so as to establish correct entertainment values and achieve healthy growth. The K-means algorithm is used to mine survey data, which provides support for guiding college students to establish correct entertainment values and promoting healthy growth.

2.Literature review and combing
In terms of online entertainment time, research in 2007 showed that younger college students spend more time on online entertainment than the elder grades.[1] A survey of college students' online time shows that many students spend 5-15 hours per week.[2]The "Report on Online Behavior of Youth in Beijing from 2017 to 2018" pointed out that the use of Internet services for young people are concentrated on leisure and entertainment, while some functions that really benefit young people have not been fully utilized. As a result, young people lack a correct view of online consumption.

In terms of motivation for online entertainment consumption, "Player Thinking: Self-Cultivation of Game Designers" explains why lonely people who meet basic needs are addicted to games, which is based on the "selfishness" and "rational person" theories of economics[3] Researchers such as Beranuy used grounded theory to conduct qualitative interviews with 9 online game players in therapy, and found that seeking pleasure, escapism and virtual dating are the main motivations for players to play...
games.[4] Dauriat used online surveys to analyze the online game motivations of 696 players and found that: sense of achievement, socialization, immersion, relaxation, and escapism are the main motivations for players to participate in MMORPG.[5] On the research of influencing factors of online entertainment consumption, the representative views are as follows: Marmorstein H Grewal believes that interactivity gives customers more autonomy and can increase their money spending pleasure. [6] Lohse G L believes that computer interface friendliness affects online consumer behaviour. [7] Colquitt J.A, Scott BA, LePine J.A believes that the personality of users play an important role in the consumption of online games. [8]

Based on the previous studies, we plan to obtain quantitative research data via questionnaire surveys and field visits. Through the analysis of data, we abstract the basic characteristics of college students' online entertainment consumption and discover the problems to analyse and gain a rational understanding of their online entertainment consumption.

3.K-means algorithm
K-means algorithm is one of the commonly used algorithms for clustering. In the Euclidean space, n data objects are divided, and K objects are selected as clustering centres from the initial centre strategy. The most recent centroids are classified. We recalculate the average value of the data in each cluster again to get a new cluster centre and then repeatedly advance the process and iteratively calculate until all clusters converge.

4.Empirical Investigation and Analysis of College Students' Online Entertainment Consumption
According to the survey results, it is found that there are not a few people who have experience in online entertainment in the college student group. 77.14% of students have experienced paying for online entertainment. This data is only the consumption caused by online entertainment itself. It is understood that many students are more inclined to play in Internet cafes when playing online games, and Internet cafes charge prices ranging from five to tens of yuan per hour. According to the questionnaire on "University Students' Online Entertainment Consumption Behavior", contemporary college students present multiple characteristics in the process of online entertainment consumption.

4.1 Significant differences in online entertainment consumer spending among college students
During the investigation, it was found that the college student group has obvious differences in online entertainment consumption expenditure. This is specifically reflected in the questionnaires, “Your expenses on online entertainment in the previous month” and “The largest amount of online entertainment content spent in 2018”. Among the 530 responding objectives, 33.33% cost less than 50 yuan, but more than 25% of the respondents said that they spent more than 200 yuan on online entertainment last month. The monthly living expenses of college students between 500-1,000 yuan accounts for a large proportion of the population. Online entertainment costs more than 200 yuan per month means that about one fifth of the monthly living expenses are spent, which is shown as in Fig.1.

![Fig1: Last month's spending on online entertainment.](image-url)

College students who have experienced online entertainment consumption indicate that online games and online music consumption account for a relatively high proportion in online entertainment
consumption. Among them, online games cost the most, accounting for nearly forty percents. The next is the cost of online music, which accounts for nearly 30%, and the last is the cost of online live streaming, which has the lowest proportion for college students. The above data also shows that online entertainment consumption may be inseparably related to the values of the college students themself, which is shown as in Fig.2.

4.2 Diversification of Online Entertainment Consumption Content
According to the results of the questionnaire survey, among the students who have experienced online entertainment consumption, 62.5% of them have paid the corresponding fees for online music; 77.46% of them have paid the corresponding fees for online games; 39.32% of students paid corresponding fees for online video watching; 34.04% of students paid corresponding fees for online literature; and 29.64% of them paid corresponding fees for webcast. It can be seen that among the various expenses of online entertainment, college students account for the largest proportion of expenses related to online games, and the expenses for online live broadcasts are relatively small, which is shown as in Fig.3.

4.3 Gender differentiation of online entertainment consumption
Online entertainment consumption also shows a certain difference in terms of gender, which is manifested in the following aspects.

First, sorted by terms of online entertainment consumption. Generally speaking, there is a small difference between male and female college students in the consumption amount of less than 50yuan and 100-200yuan. Among the spending amount of less than 50yuan, female students accounted for 26.87%, which is about 2 percentage points higher than that of male students. Among the 100-200yuan range, boys accounted for 9.59%, while girls accounted for only 4.48%, and male college students were about 5 percentage higher than female students. Among the consumption of 200yuan or more, boys accounted for 21.92%, while girls accounted for 13.43%, and the difference between boys and girls was about 9 percentage points. This shows that in terms of the amount of online entertainment consumption, there is a gender difference to a certain extent, which is shown
as in Fig.4.

![Fig4.Gender differences in online entertainment spending](image)

The second is the consumption time of online entertainment behavior. Male students and female students show a small difference in online entertainment time consumption of 0-2 hours, 4-6 hours, and more than 6 hours, while there’s a big difference in online entertainment time consumption of 2-4 hours. The duration of online entertainment for 2-4 hours was relatively high, reaching 56.61%, while that of female students was only 31.34%, about 27 percents lower than that of male students, which is somewhat consistent with the content preferences of male and female students during online entertainment. Of all the survey groups of male college students, 75.36% prefer multiplayer online competitive type games in online games.

4.4 Strong Consciousness of Copyright in Online Entertainment Consumption
According to the survey results, college students have relatively independent and correct values and value judgments in terms of online entertainment consumption behavior, and have a stronger sense of copyright and protection of copyright. In terms of copyrighted film and television and music works, college students generally reflect high copyright awareness and support copyright behavior. Taking online music as an example, the proportion of students who consume for listening to copy-righted music is higher, reaching 62.22%. At the same time, when asked if they would like to pay the corresponding fees for copy-righted music, 48.15% of the students said they were willing to and very willing to, while only 12.59% of the students said they were unwilling to and very unwilling to pay for copy-righted music; instead they are more likely to obtain those free pirated music or games.

5.Problems and Causes in Online Entertainment Consumption of College Students
From the description of the current status of college students’ online entertainment consumption behavior, it is not difficult to see that during the process of online entertainment consumption, college students present the characteristics of polarization of online entertainment consumption, gender differentiation, content diversification, and copyright awareness of consumption. One is to spend time. It is learned from the survey results that many students in the process of online entertainment behavior are largely due to their boredom of passing time, and even more than 20% of students think that online entertainment itself has big influence to their leisure time, which also reflects from the side that the current campus life is vain for some students.

The the second is using consumption to show off. In the consumer society, people are not only seeking the use value of goods, but they are constantly pursuing the "symbol value" of goods. Commodities externalize their own values into the symbol of goods. People can achieve this through the consumption of symbolic values. The purpose of personality show-off is to use this logic to distinguish personal choices from others. This is not just a code of conduct to follow in real life But also a big change with the rapid development of network today, it is also constantly adapting to the changes in the era of online consumption.

6. Conclusion
With the development of the Internet, online entertainment is showing a trend of diversified development, and online entertainment consumption is more like showing-off in a conspicuous and symbolic way. As a main force for the use of the Internet, college students, their online entertainment
and corresponding consumer behavior are showing new development trends than before. The following conclusions were drawn from this field survey and interview.

In terms of online entertainment consumption behavior, it presents the characteristics of polarization of consumption amount, differentiation of consumption gender, diversification of consumption content, and copyright consciousness during consumption. Specifically, in terms of consumption amount, some college students have spent thousands of RMB on annual expenditures, while some students have never spent it; in terms of gender, the level of online entertainment consumption of university boys is higher than that of girls, especially in money and time consuming. In terms of grades, the number of freshmen and sophomores who are engaged in online entertainment consumer behavior is larger than that of juniors and seniors, which is closely related to the pressure on preparation and entrance examination for juniors and seniors. From the perspective of living expenses, the higher the monthly living expenses are, the higher the level of online entertainment consumption is.

Acknowledgments
First and foremost, I would like to show my deepest gratitude to my supervisor, Dr. Jiangfeng Li, who has provided me with valuable guidance in every stage of the writing of this thesis. Without his enlightening instruction, impressive kindness and patience, I could not have completed my thesis. His keen and vigorous academic observation enlightens me not only in this thesis but also in my future study. I shall extend my thanks to Mr. Li for all her kindness and help.

References
[1] Liang Linmei, Li Fengqing. (2012) Information Technology and University Experience: A Review of the Status and Characteristics of North American Universities. Journal of Distance Education, 2: 77-83.
[2] Liang Linmei, Zhao Jianmin. (2013) Investigation and Research on Freshman's Network Application Behavior—— A Case Study of Nanjing Colleges and Universities. Electrification Education Research, 1: 41-46.
[3] Zhang Hongxia, Xie Yi. (2008) A Model of the Influence of Motivation Process on the Behavioral Intention of Young People's Online Games, Journal of Psychology, 1275-1286.
[4] Marta Beranuy, Xavier Carbonell, Mark D. Griffiths. (2013) A Qualitative Analysis of Online Gaming Addicts in Treatment. Int J Ment Health Addiction, 11: 149-161.
[5] Francesca Zanetta Dauriat, Ariane Zermatten, Jo Bil- lieux, et al. (2011) Motivations to Play Specifically Predict Excessive Involvement in Massively Multiplayer On-line Role-Playing Games: Evidence from an Online Survey. European Addiction Research, 17: 185-189.
[6] Marmorstein H, Grewal D, Fishe RPH. (1992) The value of time spent in price-comparison shopping: survey and experimental evidence. Journal of Consumer Research, 52-61.
[7] Lohse GL, Spiller P. (1995) Internet retail store design: how the user interface influences traffic and sales. Journal of Computer Mediated Communication, 2: 112-116.
[8] Colquitt J A, Scott B A. (2007) Trust, trustworthiness, and trust propensity: a meta-analytic test of their unique relationships with risking taking and job performance. Journal of applied psychology, 92: 909.