An analysis of the fans of Chinese virtual idols represented by A-SOUL

Wei Fu

Cardiff University, Cardiff CF10 3AT

iridiumsfw@gmail.com

Abstract. This study aims to analyse A-SOUL fans, the top virtual idol group in China. Public data from 6,778 fans were ultimately selected, collecting data related to their gender, platform account levels, and whom they follow. It was found that fans of these virtual idols prefer five areas: Japanese anime, games, life, music and knowledge. A general direction for the operation of virtual idol projects is proposed to help the companies concerned reduce the number of failures caused by meeting the needs of fans.

Keywords: virtual idols, fan analysis, A-SOUL, centrality, network

1. Introduction

Numerous companies are currently starting to make their mark on the metaverse. The virtual idol industry has become a scorching topic recently. There has been a great deal of research detailing virtual idols' past, present and future. Nevertheless, few studies have examined the audience for virtual idols. In terms of research data, there is a lack of relevant data to quantify the fans of virtual idols. Therefore, this study focuses on analyzing A-SOUL's fan base, the top virtual idol girl group in China. This research is able to give a general picture of the audience of virtual idols. This will enable virtual idol companies to operate their projects better and reduce failures caused by not understanding their fans.

2. Literature Review

As a result of the Covid-19 outbreak in 2020, the culture of non-contact began to accelerate. Major companies such as Meta, Microsoft and others have developed 'metaverse'. And virtual idols have become an area where some companies are making a push. For example, the American company Brud created Miquela Sousa, Japan's Crypton Future Media created Hatsune Miku, and China's ByteDance created A-SOUL, which is a Virtual Girl Group. There has been a great deal of research both domestically and internationally on the important role that virtual idols play in driving the development of the metaverse, but there is a lack of research on the audience groups for virtual idols. Recently, several companies have not been able to run their projects as well as they would like due to a lack of understanding of the needs of their fans, and some have even closed down their projects. When the needs of fans are understood by companies and changes are made to target them, the growth of companies in the field of virtual idols may be positively enhanced. This paper will discuss the role of virtual idols, projects and fans' behaviour in developing virtual idols. The paper mainly draws on articles related to virtual idols from China and Japan..

2.1. Virtual Idols

In 2008, this paper shows that the popularity of the virtual idol Yuki Terai is since virtual idols do not change due to biology and that they are able to conform to the taste of the audience for a long time[1]. Virtual idols are also programmed to constantly modify their appearance to conform to aesthetic principles. Hatsune Miku was born, and Miku differs from Yuki Terai in that Miku's voice can be customised using the Vocaloid package, and Miku can perform using the MikuMikuDance software [2]. And these softwares are publicly available so that fans can create customised performances of their idols to their liking. These new technologies have led to a dramatic change in the production of such stars[2-4]. Then came the birth of Kizuna AI, part of the company Kizuna AI, one of the founders of VTuber. She is a considerable departure from traditional virtual idols in that she is able to do live and respond to her audience. Many fans believe that the company uses motion capture technology as well as facial capture technology, allowing actors to play as Kizuna AI's movements. But the company does not admit to this. After that, a lot of capital poured into the virtual idol industry, but the 3D live broadcast was not ideal due to technical and other reasons, with incoherent movements and stiff facial expressions. It wasn't until December 2020 that the virtual girl group A-SOUL under the ByteDance banner started live streaming on the BiliBili platform. Previous studies have shown that the success of the group was due to the top technology at the time, the singing and dancing prowess of the performers, more varied live content and more interactive methods than ever before [5]. This has enabled the A-SOUL girl group to gain a million followers in less than a year, and the user base is extremely sticky. Such a scenario looks like the virtual idol industry still has a lot of room for growth in the future.

2.2. Project and Fans

According to previous research, during the Hatsune Miku era, some users used the Hatsune Miku sound library and the MukiMukiDance software to create their own compositions, while others commented on them or created more when they saw them. As creative users create various types of works and continue to attract various users, the number of interactive users attracted by the content will increase, which is a virtuous circle [6]. Absorbing an audience by spreading the work through fan networks has become an essential method for subsequent virtual idol fan development. And in the long run, music, film and television, variety shows, animation, web series and other directions should all become the direction in which the idol industry should gradually develop in the future.[7] However, Kizuna AI went down a different path because of the company's lack of understanding of what audiences wanted to see. After creating three identical Kizuna AI's, fans were very unhappy with the company's actions. This incident is considered in the study as the company simply treating the Kizuna AI as a commodity, a commodity that is constituted by the labour of voice actors, etc. [8]. The A-SOUL project then caused a great deal of fan discontent due to the high price of selling handshake coupons (users with handshake coupons can use force feedback gloves to shake hands with virtual idols), and the number of comments on related posts reached 10,000, with almost all users expressing their anger.

2.3. Summary

It can be seen that previous research has provided insight into the virtual idol industry, the need for virtual idols to receive relevant protection, and the need for companies to make a shift in the positioning of virtual idols. However, there is no research to analyse how this web-like fan base is constituted and which part of the fan base it mainly attracts. Understanding what fans want plays a vital role in the subsequent development of the company and its plans. This paper will analyse the fan base's composition, understand the fans' needs and deepen the understanding of the needs of virtual idol fans.

3. Methodology

This paper uses a quantitative research method to analyse the fans of the A-SOUL virtual girl group. The quantitative approach is effective in obtaining the nature of the majority of A-SOUL fans, while the qualitative analysis may lead to some bias in the final results due to the choice of interviewees. As this is a quantitative analysis, a large amount of fan-related data is required as a sample for the study. The

data was collected using a Python crawler and an API provided by BiliBili. The reason for using Python is that the number of A-SOUL fans is so large that it would be cumbersome to collect them manually. The author wrote the Python crawler code. The crawler starts with the users who have commented on all videos posted by the five A-SOUL idols at the first level (meaning that they have commented directly on the content of the video, not on other users), as it is not possible to obtain all the fans of a user due to official restrictions. But the ones obtained this way may not be real fans, so the total number of comments per user is analysed, and when a user has posted more than ten jumps under the comments section, there is a high probability that this user is a fan of A-SOUL. Afterwards the 250 most recently followed users are read through the official API for these filtered users, and if the number of users followed is less than 250, then all the followed users are recorded. Since some users do not open their follow list, any record is made. The last users with the following information were used as the subject of this study. The data obtained this time is as of July 3, 2022. After the data was obtained, the analysis of the composition of this group of followers was started. Firstly, the basic information of the research subjects was statistically analysed. This was followed by categorising the videos sent by the study subjects' followers and labelling these study subjects' followers, for example, whether they were video creators in the food section or the dance section. But it is also necessary to remove some of the noise. For example, some users will follow their classmates, etc. Therefore only those who are followed more than 20 times will be recorded again. At this point, the subjects are also tagged, and these tags are used to give an idea of the subject's preferences through the users they follow. The tags of each subject are then statistically analysed, and a network is constructed. Each user half has two and more tags, using any two tags of a research subject as edges. When another user has this label at the same time, the weight of this edge is added by one until the whole network is constructed. After the network is constructed, the centrality of this network is analysed to understand which categories are the areas that A-SOUL should work on in the future.
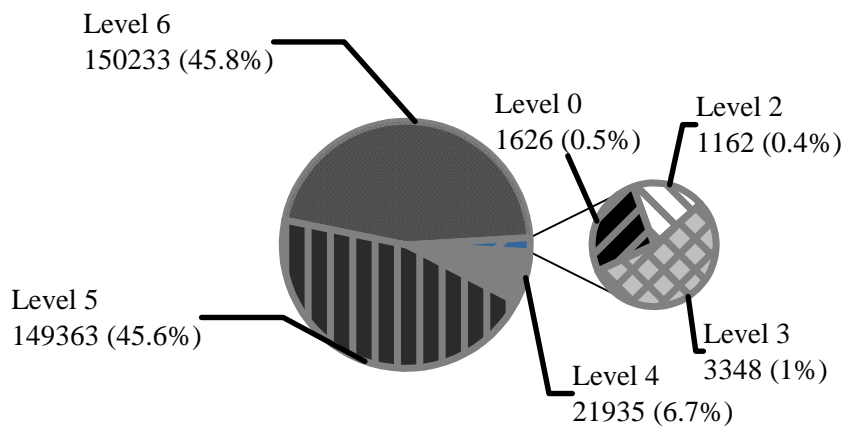
4. Result

Figure 1 shows a count of the gender of the users collected and it can be briefly noticed that 52.5% of the users kept their gender information confidential. The ratio of men to women who disclosed their gender was roughly 3:1.
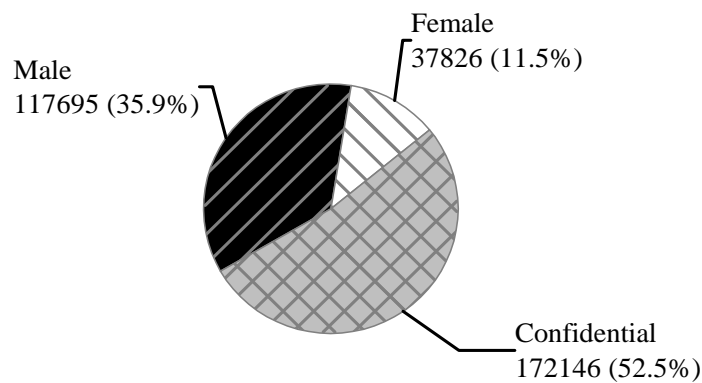

Figure 1. The pie chart depicts the number and proportion of gender in the study population

A 6-level division has been implemented on the BiliBili platform, with unregistered users being level 0. Users are upgraded to level 2 directly after registration and each subsequent upgrade takes a significant amount of time, with an average viewing user taking approximately 4-6 years to upgrade to level 6. Moreover, from Figure. 2, the number of Level 5 as well as Level 6 users accounts for 91.4% of the overall. And Level 4 users also account for 6.7%.

Level 6
150233 (45.8%)

Level 0
1626 (0.5%)

Level 2
1162 (0.4%)

Level 5
149363 (45.6%)

Level 3
3348 (1%)

Level 4
21935 (6.7%)

Figure 2. The pie chart depicts the distribution of the corresponding accounts of the study participants on the BiliBili platform in terms of rank

The network in Figure 3 consists of labels for each of the subjects studied. Each node in the network represents one type of label. Table 1 explains in detail the labels corresponding to each node. Node 8 (GuoChuang) means works created in China, while node 17 (Douga) refers to Japanese anime, for example. Since there are 21 categories of tags in total when a user's object of interest generates a plural number of tags, these tags correspond to two nodes of the network two by two, and the number of correspondences is the weight of the edges of the two nodes. The graph gives a more visual representation of the importance of the nodes. For example, nodes three (corresponding to the TV series tag) and four (corresponding to the documentary tag) have a small weight for each edge. It is not possible to obtain accurate information by just looking at a network, the centrality of the network nodes is more than enough to more accurately reflect the label of the object of study.
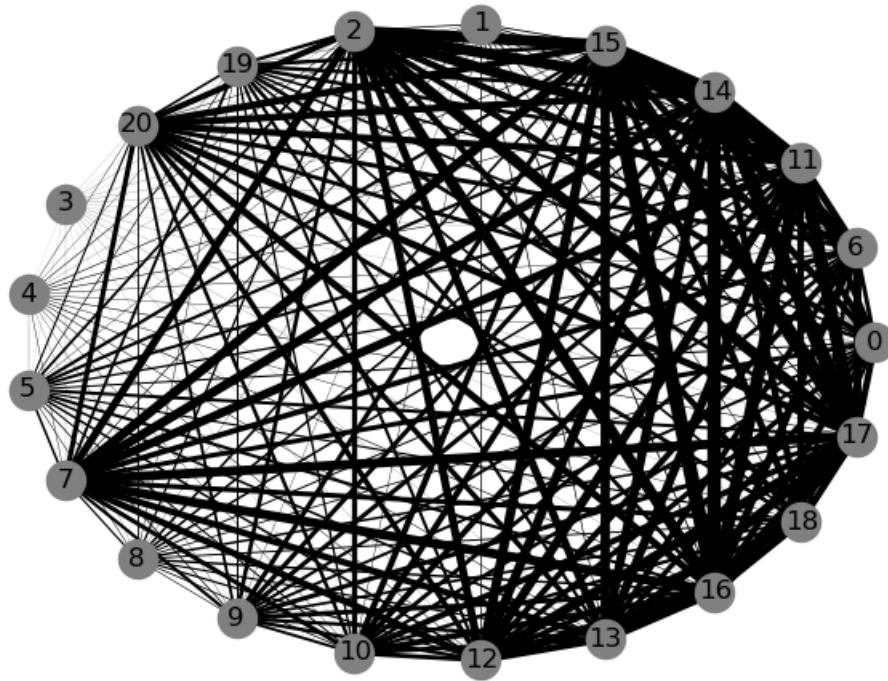
Figure 3. A network formed by the tags of the followers of the research subject

Table 1. Tags corresponding to nodes (The odd numbered columns refer to the node number corresponding to the tag, the even numbered columns are the specific name of the tag)

Tag Number	0	1	2	3	4	5	6
Tag Name	Sport	Movie	Food	TV	Documentary	Anime	Entertainment
Tag Number	7	8	9	10	11	12	13
Tag Name	Kichiku	Guochuang	Fashion	Dance	Knowledge	Cinephile	Technology
Tag Number	14	15	16	17	18	19	20
Tag Name	Music	Life	Game	Douga	Information	Car	Animal

After finding the degree centrality of the network in table 2, the betweenness centrality, closeness centrality, eigenvector centrality and Katz centrality.

Table 2. Degree centrality of each node(The odd-numbered rows are the nodes, and the even-numbered rows are the values of the corresponding node's degree centrality)

Tag Number	0	1	2	3	4	5	6
Degree Centrality	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Tag Number	7	8	9	10	11	12	13
Degree Centrality	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Tag Number	14	15	16	17	18	19	20
Degree Centrality	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Looking at Table 3, one can see that only node 3 has a value of 1, while all other nodes have a value of 0. In Table 4, all nodes have a value of 1. Later in Table 5, as well as in Table 6, the data in the table are the Eigenvector centrality values and the Katz centrality values for each node

Table 3. Betweenness centrality of each node(The odd-numbered rows are the nodes and the even-numbered rows are the values of the corresponding node's degree centrality)

Tag Number	0	1	2	3	4	5	6
Betweenness Centrality	0.0	0.0	0.0	1.0	0.0	0.0	0.0
Tag Number	7	8	9	10	11	12	13
Betweenness Centrality	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Tag Number	14	15	16	17	18	19	20
Betweenness Centrality	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table 4. Closeness centrality of each node(The odd-numbered rows are the nodes and the even-numbered rows are the values of the corresponding node's degree centrality)

Tag Number	0	1	2	3	4	5	6
Closeness Centrality	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Tag Number	7	8	9	10	11	12	13
Closeness Centrality	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Tag Number	14	15	16	17	18	19	20
Closeness Centrality	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Table 5. Eigenvector centrality of each node(The odd-numbered rows are the nodes and the even-numbered rows are the values of the corresponding node's degree centrality)

Tag Number	0	1	2
Eigenvector Centrality	0.1823223908099383	0.038234624108519744	0.2801425788373377
Tag Number	3	4	5
Eigenvector Centrality	0.007253242642868229	0.026118686880744296	0.08727671084307383
Tag Number	6	7	8
Eigenvector Centrality	0.20842211220365883	0.2387936019827163	0.06578944897506225
Tag Number	9	10	11
Eigenvector Centrality	0.12984636254660417	0.166840500017688	0.3104183325667527
Tag Number	12	13	14
Eigenvector Centrality	0.24419901103930092	0.24102759390047598	0.3244903695874532
Tag Number	15	16	17
Eigenvector Centrality	0.33103641204989837	0.32625758978669517	0.3316060861446514
Tag Number	18	19	20
Eigenvector Centrality	0.18751687002122436	0.10928916874234243	0.19104741996619137

Table 6. Katz centrality of each node(The odd-numbered rows are the nodes and the even-numbered rows are the values of the corresponding node's degree centrality)

Tag Number	0	1	2
Katz Centrality	0.016535398602871103	0.34353060934894586	-0.17678534707555538
Tag Number	3	4	5
Katz Centrality	0.40912489061982543	0.37092060602449956	0.2256374849399786
Tag Number	6	7	8
Katz Centrality	-0.037499424834739406	-0.10795682854860415	0.2777596133858561
Tag Number	9	10	11
Katz Centrality	0.13146028591125966	0.049988792144603965	-0.23005624288635956
Tag Number	12	13	14
Katz Centrality	-0.10887741249761775	-0.10101407826434064	-0.256582258712989
Tag Number	15	16	17
Katz Centrality	-0.2683555667006963	-0.2589100009559461	-0.2696059844857529
Tag Number	18	19	20
Katz Centrality	0.00910648188210662	0.17866036694952228	-0.005772091338838331

5. Discussion

According to Figure 1, there are far more males than females in terms of gender. Firstly this is probably because the idol group is a female group and the main audience is male. Secondly, in China, more men are into a virtual culture. Furthermore, the technology that A-SOUL demonstrated is something that has not been seen before. These are the points that attracted a large number of men and became fans. The more important point is that more than half of the users have set their gender as private. This is partly a reflection of the fact that users value their privacy. Assuming anything is done on the part of the project that violates the privacy of the users, this will have severe repercussions for the whole project.

According to table 2, it can be seen that the majority of A-SOUL's fan audience is indigenous. This means that A-SOUL's previous line of operation matched the interests of the original group of active BiliBili users. And as BiliBili's main audience is Chinese youth, another way of interpreting this is that A-SOUL's fans are as growable as its idols. But the small number of low-level users is a side indication that A-SOUL is too unattractive to newcomers to the BiliBili platform, something that may have implications in terms of future growth. The project should change or increase the activity behaviour of the virtual idols without affecting the original fans in preparation for attracting a more diverse fan base in the future.

Finally, observe the five centrality-related tables. Degree centrality reflects whether a point is directly connected to other points. The fact that all values of degree centrality are 1 indicates that the A-SOUL fan base is very well-rounded. In betweenness centrality, node 3 is the only non-zero number. Although the shortest path does not have practical significance in this network, the significance of this shortest path is that the links between this point and the other points have little weight due to the presence of weights. This means that the label node 3 is of little interest to A-SOUL fans. Therefore, the project should minimise the element of drama represented in node 3. If there are too many elements in this section, it may cause users to lose interest. Afterwards the CLOSENESS centrality is not very different from the degree centrality, which is due to the fact that every node in the diagram is connected to every other node. The more important eigenvector centrality, eigenvector centrality is a better response to the influence of nodes on the network. When a node has a high value of eigenvector centrality, this indicates that the node is very important. Although this network is a complete graph, the influence of the weights makes this value very informative. The corresponding node importance ranking can be obtained by reordering the contents of the table as follows:

17, 15, 16, 14, 11, 2, 12, 13, 7, 6, 20, 18, 0, 10, 9, 19, 5, 8, 1, 4, 3

This level of ranking reflects a large extent the preferences of the fans. For example, the top five areas are Japanese anime, games, lifestyle, music and knowledge. Projects can focus on these areas. For example, increase the number of songs and dances during the virtual idol performances, and talk more about life and knowledge during the miscellaneous talks. If the project is structured around these areas, it will capture the hearts and minds of the users and make A-SOUL better. In the subsequent Katz centrality, the ranking of importance remains almost unchanged, with only 7 and 13 changing in order, which again illustrates the importance of the first five tags.

However, the judgement of some of the data obtained this time is not rigorous enough, for example, there is no better filtering of fans as well as video publishers. Although most of the noise was removed, it also had an impact on the original data. Moreover, the population of this study is limited to A-SOUL fans, which are the top group of virtual idols in China but are not representative of other virtual idols, such as those from other platforms. Future research could use k-means algorithms to better filter fans and video publishers and obtain more data from more platforms.

6. Conclusions

The aim of this study is to determine the future direction of operation for major virtual idol projects. Based on a quantitative analysis of the fan preferences of A-SOUL, one of China's top virtual girl groups, it can be concluded that future projects should have an in-depth understanding of the five areas of anime, games, life, music and knowledge and apply them to virtual idols. The results also show that users have little to no interest in drama. To better understand the implications of these results, future research could address the noise of the fans and the corresponding tags. For example, when dealing with whether or not they are up (similar to YouTubers), a k-means algorithm could be used to analyse three dimensions: number of followers, number of videos posted, and number of comments. And this study provides a degree of analysis as well as recommendations for the virtual idol industry to understand the possible market, which is a more comprehensive analysis of fan preferences in the Chinese virtual idol industry.

References

[1] Black D 2008 The Virtual Ideal: Virtual Idols, Cute Technology and Unclean Biology Continuum vol 22 pp 37-50

[2] Le L K 2014 Examining the rise of Hatsune Miku: The First International Virtual Idol. (The UCI Undergraduate Research Journal) vol 13 pp 1-12

[3] Black D 2012 The Virtual Idol: Producing and Consuming Digital Femininity Idols and Celebrity in Japanese Media Culture(London: Palgrave Macmillan UK) pp 209-288

[4] Zaborowski R 2016 Hatsune Miku and Japanese virtual idols (The Oxford Handbook of Music and Virtuality)

[5] Wang Y Q 2022 A Brief Analysis of the Development of Chinese Virtual Idol Industry Empowered by 5G+Motion Capture Technology——Taking the Virtual Idol Group A-SOUL as an Example vol 2278 (Bristol: IOP Publishing)p 012011

[6] Kobayashi H and Taguchi T 2018 Virtual Idol Hatsune Miku: Case Study of New Production/Consumption Phenomena generated by Network Effects in Japan's Online Environment vol 3

[7] Zhang L Y 2018 The Influences of Idol Effect on the Purchasing Decisions of Their Fans vol 68 (Xian: International Symposium on Social Science and Management Innovation) pp 164-171

[8] Zhou X 2020 Virtual Youtuber Kizuna AI: co-creating human-non-human interaction and celebrity-audience relationship