Exploring Social Dynamics in Online Media Sharing

Martin Halvey
School of Computer Science & Informatics, University College Dublin, Belfield, Dublin 4, Ireland.
+353-1-7165355

Martin.Halvey@ucd.ie

Mark T. Keane
School of Computer Science & Informatics, University College Dublin, Belfield, Dublin 4, Ireland.
+353-1-7162470

Mark.Keane@ucd.ie

ABSTRACT

It is now feasible to view media at home as easily as text-based pages were viewed when the World Wide Web (WWW) first emerged. This development has supported media sharing and search services providing hosting, indexing and access to large, online media repositories. Many of these sharing services also have a social aspect to them. This paper provides an initial analysis of the social interactions on a video sharing and search service (www.youtube.com). Results show that many users do not form social networks in the online community and a very small number do not appear to contribute to the wider community. However, it does seem those people who *do* use the available tools have much a greater tendency to form social connections.

Categories and Subject Descriptors

H.1.2 User/Machine Systems

General Terms

Measurement, Performance, Human Factors.

Keywords

Media, social dynamics, search, video.

1. INTRODUCTION

Websites that allow users to share media online have become a phenomenon of growing social, economic and cultural importance, routinely attracting millions of users and hits. Prime examples of this phenomenon can be found in Flickr's support for the sharing of photographs and YouTube's sharing of video files. Apart from sharing, some other websites explicitly support social networking (e.g., MySpace and Bebo) allowing media sharing and the sharing of tastes and interests. In this paper, we examine social interaction between users on YouTube. This social interaction is examined by looking at the tools people use to share online media and index video clips (see Ducheneaut et al [2] for a related study in online gaming). Our focus is to determine whether aspects of this behavior that can be used to improve collaborative-based search on future systems (e.g., [4]). Such an approach could significantly enhance the metadata attached by creators of the media, thus enhancing the searching of such repositories.

2. ANALYSING USERS USING SOCIAL FEATURES

In June 2006, we carried out a crawl of the YouTube video sharing and search engine. Over 100,000 pages were downloaded involving videos added by 57,588 YouTube users; coming from 204 countries with a mean age of 25. In September 2006, the online behavior of these users was analyzed focusing on their use

of website-supported methods for social interaction (e.g., linking to other users as friends and in groups, and interacting with other users through comments and subscriptions). About half of these users were male, a quarter being female and the remainder not indicating their sex.

	Mean	*Max*	*Number Equal to Zero per Video*
Views	966.32	361347	324
Uploads	11.03	2261	4030
Friend	6.80	7455	31454
Favourite	15.91	2412	30635
Groups	0.38	228	50564
Subscribers	10.87	24357	29845
Subscription	1.06	469	42966
Comments	2.20	983	43696

Table 1: Analysis of Users of the YouTube Service

Table 1 illustrates some of the key figures from the analyses of interaction carried out. They suggest the following:

- Users tend to visit the site to view videos rather than to add their own media; on average they view 966 videos though they only upload 11 videos on the site (though note that the variance is huge compare the Max with the Min/Equal-to-Zero).

- Users are more likely to save videos to their profiles rather than upload videos, each user saved an average of approximately 15 videos to their profile

- There is a widespread failure amongst users to exploit the community facilities available on the website (i.e., most users are not members of groups, do not have any subscriptions, do not list friends or favorite videos, and have not posted any comments on videos).

These findings have several implications for any attempt to leverage this data for personalization or recommendation purposes. Specifically, these usage patterns suggest that usage data might not be that useful for generating predictions for communities of users. However, the fact that some of these averages are reasonably high indicates that the users who use these facilities use them quite often, and that the minority of users who use these facilities might benefit from such facilities. In fact since an initial (and less detailed) crawl of these users profiles was carried out by us the average number of uploads has

increased by approximately 9, the average number of friends increased by approximately 3 and the average number of views increased by 500. Of those 3 categories the only one to show a large decrease in the number equal to zero is the number of uploads. This indicates that the small number of users who use these facilities continue to do so at a regular rate.
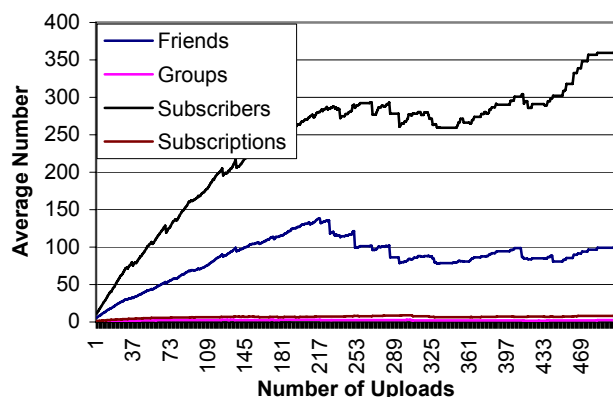

Figure 1: Average number of social features with respect to number of uploaded videos on YouTube website.

Figure 1 shows the figures for different social interactions with respect to the number of files that users upload, for all users. Figure 1 shows that users who are forming social groups, favour being friends as a facility to form these groups, rather than explicitly being member of a group. It also shows that users have more subscriptions on average than any other social interaction. As expected the number of subscribers that a user has increases with the number of videos that a user has uploaded. When the same graph is plotted with respect to the number of views for all users, the number of subscribers increases with respect to views and then falls off after a certain point (figure not shown due to space requirements). The shape of the graphs for friends, groups and subscriptions are similar with respect to both views and uploads. In both scenarios it can be seen that the number of friends increases to a point and then begins to tail off slowly. Again in both scenarios the number of subscriptions that a user has and the number of groups that they are a member of are negligible.

3. DISTRIBUTION OF VIEWS

In this section, the distribution of views for users of the website are investigated. Figure 2 shows the distribution of views of videos files for the users of YouTube. The distribution of uploads and the distribution of favorites was also investigated, but for reasons of space is not shown here. We found that the distribution of views does not reflect the Zipf-type distributions typically observed in desktop and mobile browsing contexts [3]. Instead a slight curvature in the distribution can be seen. This distribution is similar to one found in search logs [1] and indicates that users find videos through searching rather than browsing.

In web scenarios, the Zipf distribution is typically a consequence of the fact that there are very few pages with large numbers of hits, and many pages have very few hits. However, in this case, users with high numbers of "views" increase to a certain point and then drop off in a more Zipf-like way. This distribution suggests

that users in this scenario actually visit more pages within the site, or within a session than in the desktop or mobile browsing paradigm. However, when the distribution of uploads and the number of favourites was investigated a Zipf-like distribution was indeed found.
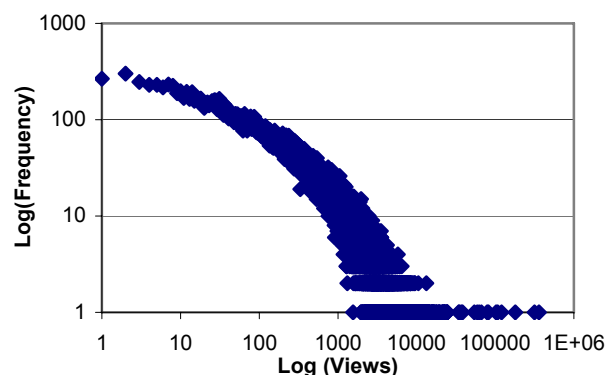

Figure 2: Frequency distribution of views per user for video files

4. CONCLUSIONS AND FUTURE WORK

For this analysis a database of information about users that use a video sharing and search system, YouTube, was created. This database also contained information about the users that have added those video files to the search engine. It has been shown that in general, a large number of users do not use the facilities for social interaction available to them in media sharing services. However, those users that do appear to use those services quite frequently, creating social networks within those media sharing services. Hopefully such social networks can be exploited in order to aid the user experience within these services, e.g. media search within the service. With the growth of media online and as it becomes more accessible to users it is important that better search methods are discovered, and differences between search paradigms and users are identified. To that end, this work will lead the way and help improve video search. This information can help make content more accessible to users of the Internet.

5. REFERENCES

[1] Beitzel S, Jensen E, Chowbury A, Grossman D and Freider O. Hourly Analysis of a very large categorized web query log. In Proceedings of the 27th International SIGIR Conference, Sheffield, United Kingdom, (2004).

[2] Ducheneaut, N., Yee, N., Nickell, E., and Moore, R.J. "Alone Together?" Exploring the Social Dynamics in Massively Multiplayer Online Games. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems . (2006).

[3] Halvey, M., Keane, M.T. and Smyth, B. Mobile web surfing is the same as web surfing. Communications of the ACM 49(3) : 76-81 (2006)

[4] Smyth, B., Balfe, E., Boydell, O., Bradley, K., Briggs, P., Coyle, M. and Freyne, J. A Live-User Evaluation of Collaborative Web Search. In Proceedings of IJCAI 2005 Conference. (2005).