A Glance at an Overlooked Part of the World Wide Web

Ionut Trestian
Northwestern University
ionut@northwestern.edu

Chunjing Xiao
Northwestern University
xiao@northwestern.edu

Aleksandar Kuzmanovic
Northwestern University
akuzma@northwestern.edu

ABSTRACT

Although according to surveys related to internet activity it is considered one of the most popular aspects, few studies are actually concerned with internet pornography. This paper is aimed at rectifying that overlook. In particular, we study user activity related to internet pornography by looking at two main behaviors: (*i*) watching pornography, and (*ii*) providing feedback on pornography items in the form of ratings and comments.

By using appropriate datasets that we collect, we make contributions related to the study of both behaviors pointed out above. With regards to viewing, we observe that views are highly dependent on pornography category and video size. By studying the feedback system of pornography video websites, we observe differences in the way users rate items across websites popular in different parts of the world. Finally, we employ sentiment analysis to study the comments that users leave on pornography websites and we find surprising similarities across the analyzed websites. Our results pave the way to understanding more about human behavior related to internet pornography and can impact, among others, fields such as content personalization, video content delivery, recommender systems.

Categories and Subject Descriptors

J.4 [**Computer Applications**]: Social and Behavioral Sciences

General Terms

Measurement, Human Factors

Keywords

Internet Behavior, Sentiment Analysis, Search Behavior

1. INTRODUCTION

Sexuality, although often times a tabu subject, is an important part of human behavior. As an integral part of modern sexuality, viewing, and rating pornography content on the internet are definitely an important element of internet activity. Statistics related to internet pornography have pointed out that about 42.7% of internet users visit pornographic websites. Pornography related queries to search engines account for 25% of all search engine queries. About

12% of all websites are pornographic in nature and worldwide revenue is in the tens of billions of dollars [3].

Motivated by the high ease with which users can find pornographic material on the internet, social sciences researchers have started investigating the effects that internet pornography has on individuals. The research work done only in the internet sexuality area is tremendous with hundreds of research papers written on the subject. A survey of recent research can be found in [6].

Given the obvious importance it has on human behavior and, judging by the statistics above, also on the internet, one would expect more attention from the internet research community. However the few studies that were carried out so far have taken a bird's eye view approach. A few studies looked at the infrastructure that drives such websites and tried to understand the inter workings of the industry as in [11]. A few other studies provided high level statistics of pornography consumption such as in [3], or looked at how pornography consumers actually search for pornography content on the internet [18]

The lack of studies is surprising since pornographic websites, because of the way they are structured, contain a wealth of information that can be analyzed to shed light in this dark aspect of internet activity as we can see below.

- The first feature is that the data is highly structured and rich in information. The websites often also provide statistics related to the content type, number of views, or to the video type and quality. This allows us to get a sense of the different types of pornographic material that are popular in different parts of the world in terms of views and that users are actually *watching*.

- Pornographic websites often employ rating systems similar to the ones employed by video sharing websites such as `youtube.com`.

- The websites allow users to leave comments thus providing even more feedback in natural language. Although this feedback is unstructured and usually it comes in large quantities that would take a human a long time to process it also provides rich analysis data. Both rating and commenting on pornography are a rich form of *feedback* that users leave.

This paper is aimed at studying the behaviors outlined above, watching pornography and pornography feedback. We have built a distributed crawler and extract the information we outlined above. We develop a sentiment analysis based methodology to analyze the relationship between comments and ratings.

Some of the key research questions we aim to answer in this paper are. Is internet pornography related behavior different than other internet related behaviors? What are the different types of pornographic content that are popular around the world? What are the differences across pornographic services? What are the sentiments expressed in comments on pornography items and what is the relationship with the rating that each item has? Our analysis shows there indeed is a difference in the way users search for pornographic items basically by employing more keywords and therefore carrying out more specific searches. We also find that different categories of pornography are popular around the world and we have observed differences in the way users rate items. With regards to the sentiments expressed by users in comments, we have observed a very strong objectivity reflected across the pornography services analyzed.

2. A SHORT OVERVIEW OF INTERNET PORNOGRAPHY SERVICES

There are multiple ways in which users can have access to pornography on the internet. For example some users might search and download it on p2p networks such as BitTorrent or Gnutella. Such behavior is relatively hard to track as can be seen in [12] and also not applicable to our study as it would be hard to determine if the users actually view the downloaded videos. Also not to mention the legal ramifications inherent from carrying out such a study and the fact that user feedback for viewing such videos would be close to inexistent.

Therefore, because of the ease of use with which we can obtain relevant statistics information and also because of the large amount of feedback available on such services, we focus on online internet pornography. By that we understand websites that provide access to pornographic content. In this section we provide an overview, some examples, and some classification criteria for internet pornography services.

Internet pornography websites can be classified according to several distinctive criteria that we outline below.

- A first distinction is based on pricing. For example there are websites that are free such as `xvideos.com` [1] and websites where one pays money but usually for higher definition content such as `realitykings.com` [2] Such paying websites often produce their own content. Multiple pricing models exist, including monthly memberships and pay-per-view.

- Content type is another division. There are websites that contain only pictures such as `pichunter.com`, [3] websites that contain videos such as `youporn.com` [4], websites that link users to live video performers such as `livejasmin.com` [5] and websites that contain a combination of some or all of the above. Some pornography websites, in typical Web 2.0 fashion, even offer users the possibility to rate or comment on content or chat live with performers for live websites.

Several other types of services linked to pornography exist. For example there exist vertical search engines aimed at searching pornographic content such as Booble [6] . Pornographic websites often employ a very complex infrastructure to bring users to paying websites. Some websites are only a combination of image links that redirect users to various websites that might have or not a connection with the images that users click to get there.

By studying the top websites on Alexa we find that there is a hierarchy in terms of how pornography websites are linked among each other. This, although interesting, is however beyond the scope of this paper and we aim to explore it in future work. Some free websites are used just to link users to paying websites and are usually owned by the same entities. Basically users do not pay for low quality content but if they want to increase the quality or interact with the performers they are linked to higher quality or live websites where they pay for watching the content.

In the rest of this paper we analyze video streaming pornography websites for the following reasons. First, they have a low barrier of entry for users. One does not need an account to stream videos and the websites we have selected are free and highly popular according to [1]. Second, the content involved, mainly video, is also a lot more taxing on transporting networks and any result that determines local user preferences towards certain types of content can have large ramifications on content distribution. Third, because of their structure (employing feedback systems and providing statistics information) they offer one of the richest sources of information about this unknown human activity. To the best of our knowledge paying websites do not provide statistics information, one of the gains that users get by paying for the service being in the enhanced privacy.

In the next section we look at search engine behavior related to pornography seeking in an effort to understand how some users arrive at pornography websites.

3. EVALUATION OF PORNOGRAPHY RELATED SEARCHES

Search engines influence the way users access much of the information or content that can be found on the internet. User behavior related to search engines has also been analyzed in great detail in the recent years [9]. Some aspects relate to how specific users are when entering searches, how many keywords they use to identify the information, how distinct and long the searches become.

Not surprising, a large part of internet search is also related to pornography. As such in this section we try to contrast searches related to pornography with the rest of searches. As data, we have a week long trace consisting of users accessing web services from mobile devices and laptops connected to the internet using 3G dongles. The data comes from one of the biggest mobile providers in the United States. The trace contains the data of close to 2 million users spread across the entire United States and the collection and information extraction process is more thoroughly described in [19]. Roughly, we extract the URLs that users access in the trace interval and we distinguish the ones that lead users from search engines to pornography websites from the ones that lead to all the other websites. We identify pornography websites by extracting the top adult websites ranked by

---

[1]http://www.xvideos.com

[2]http://www.realitykings.com

[3]http://pichunter.com

[4]http://www.youporn.com

[5]http://www.livejasmin.com

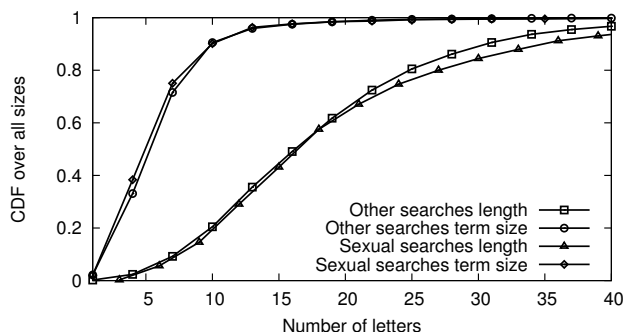[6]Booble-Adult Search Engine http://www.booble.com

Figure 1: Search sizes for searches and terms that end up at pornographic websites compared to all other searches.

Alexa as pornographic in nature [2]. We analyze pornography related searches and normal searches. As search engines we have the top three search engines in our trace: Google, Bing, and Yahoo.

Although we do not have ground truth information related to the devices that people use and only know roughly that the trace contains a combination of feature phones, smartphones, and laptops equipped with dongles, by using one of the parameters that devices pass to websites and namely the resolution we can roughly classify the devices into the classes mentioned above. When looking at pornography interests across device types we can observe that more users that have a higher a higher resolution device are using that device to search for pornography. Another result is that for mobile devices that have capabilities closer to laptops such as advanced smartphones or tablets, the number of users that search for pornography is similar to the number of users that search for pornography using laptops. The number is within 3%.

At first glance, when looking at the length of the search strings related to pornographic content, they are longer than the search strings that users enter for other searches. This is presented in Figure 1. The average is 23.03 characters for pornography searches compared to 20.41 characters for all the other searches. Searches usually consist of several terms however so we take a look at this aspect next. By breaking search strings in terms we observe that pornography searches contain 3.34 terms on average while all other searches contain 2.91 terms on average. This makes sense as by examining the different searches we can see that users searching for pornography tend to be more precise in their searches reflecting the preferences in gender, race (*e.g.* latina), or hair color (*e.g.* blonde) for the performers in the pornography item. However one can note also that pornography terms tend to be shorter since in pornography searches, terms are 6.89 characters on average and the terms in all other searches are 7.01 characters on average. We do not actually detail on the terms used because of the obvious graphical nature.

By examining the distribution of search string sizes and term sizes shown in Figure 1 we can observe the same effect. Most search strings related to pornography websites are indeed larger than all the other search strings. The opposite effect can be observed for search terms related to pornography that are shorter than search terms that lead to all the other websites.

To conclude, search for pornographic material indeed seems different than searches for other types of information in the sense that users tend to be more specific in the types of pornography they are looking for. In the next section we take a look at the types of pornographic content that users are interested in.

4. DATASET

As pornography websites come in many flavors we focus on just a few websites that we select from among the most popular ones. The ones we focus on are shown in Table 1. `xvideos.com` and `youporn.com` are the most popular video sharing websites according to [1]. We select the others for local popularity reasons. `jizzhut.com` is popular in Singapore, `indianpornvideos.com` contains mainly indian videos, and `tubeadulte.com` and `p0rno.com` are popular in France (according to Alexa). Although desirable, we cannot perform any user level analysis as the information we have is aggregate.

One feature of video sharing pornography websites is that the videos they display are categorized. This gives us a wealth of information from which we can study user preferences across the world towards certain categories of pornography. We crawl several websites and show related statistics in Table 1.

URL	Videos	Avg. Len(s)	Avg. View
p0rno.com	13,824	268	20.4
tubeadulte.com	12,563	267	2.7
youporn.com	58,039	549	406
xvideos.com	446,481	636	N/A
xhamster.com	688,349	666	104.2
jizzhut.com	1,220,953	1,139	91.4
indianpornvideos.com	3323	830	36

Table 1: Content and views statistics.

The data we manage to collect is quite large (50 Terabytes in size, comments taking the most part). Statistics about the average number of views per video and the average length of videos from different websites can be seen in Table 1. Most videos are several minutes on average: 4 to 20 depending on the website. We can see that `p0rno.com` and `tubeadulte.com` carry shorter videos on average while `jizzhut.com` carries larger videos. All websites except `xvideos.com` display the number of views and we can see that `youporn.com` carries the most popular videos in terms of average number of views.

From the categories that are common across most websites we take a look at: *anal, blowjobs, cumshots, lesbian* - categories that are based on the pictured sexual act, *blonde, brunette, latina* - categories that are based on performer characteristics, and *amateur* a category that contains amateur recorded videos. Although some titles contain actor names, the websites we examined do not differentiate content depending on actors.

Category	Label	Category	Label
Amateur	A	Brunette	E
Anal	B	Cumshots	F
Blonde	C	Latina	G
Blowjob	D	Lesbian	H

Table 2: Categories and the associated labels.

In the rest of the paper we will employ the labels in Table 2. The table also summarizes the main categories that we focus further on. Figure 2 shows statistics related to views and duration of videos for several representative categories. The results are as follows.

5. PRELIMINARY ANALYSIS

First, regarding the duration of the videos one can observe differences across websites. For example, for `xhamster.com` the curves tend to be smoother (Figure 2a), while for `xvideos.com` and `youporn.com` there are significant steps in the curves corresponding to a large number of videos of a certain size (Figures 2c and 2e). This is because `xvideos.com` and `youporn.com` have different other partner websites posting sample videos on them. in an effort to attract users of the free website to a paying website to view the full length content. Many of these videos (being samples) are limited to a fixed size by the posters hence the steps. No such behavior was observed for `xhamster.com`.

Second, with regards to the duration of the different categories we can notice some obvious differences across categories. What is surprising though is that these differences carry across different websites. For example categories A, D, and F tend to be shorter in length across all the analyzed websites (also shown in the figures). This is not surprising as category A tends to involve low quality movies sometimes shot with mobile phones or with low quality cameras. As for categories D and F this is not surprising given the nature of the sexual act involved which tends to be shorter. Categories B and E (and other similar ones such as C, G, and H) tend to be larger in size across the analyzed websites again due to the nature of the sexual act involved in the movie and they involve higher quality videos compared to the previously mentioned categories.

Third, regarding the number of views, several results can be noted. Lower quality videos such as those in category A tend to consistently get a lower number of views across websites. This can be noticed in Figures 2b, 2d , and 2f. This seems to imply the fact that even though they come to a free service, pornography seekers tend to find and watch higher quality content. Another result shown is that there seems to be a correlation between the size of the videos watched and the number of views a video will get. For example categories A, D and F that have lower sizes also have a lower number of views compared to categories B and E. Note also that we do not show the number of views for `xvideos.com` as they are not displayed on the website. We therefore display the view related information for the website popular in France, `p0rno.com` and note the similarity with the other services for views across categories.

One exception to the lower quality, lower video size therefore lower number of views rule was observed as follows. Several websites offer categories involving celebrity sex scandals containing for example leaked celebrity movies. Such categories are named as: sex scandal, sex tapes, bollywood (for `indianpornvideos.com`), *etc.* The videos involved are most times low quality, low size. However because of the obvious scandalous nature of the videos they tend to be popular and heavily viewed.

6. FEEDBACK SYSTEM

The feedback that users leave when interacting with pornography websites is broadly represented by two components: (*i*) ratings as most websites provide a way to rate the content either through a grade from 1 to 5, or with a percent value, or a binary **yes** or **no** value, and (*ii*) comments, informal feedback that users provide regularly in natural language. Table 3 provides some statistics regarding the number of comments we extracted and the types of rating systems installed by different websites.

URL	Comments	Rating	Average type
p0rno.com	N/A	1-5	Rounded
tubeadulte.com	N/A	1-5	2 digits
youporn.com	40,329	1-5	2 digits
xvideos.com	5,659,310	Yes/No	2 digits
xhamster.com	1,500,802	1-5	Rounded
jizzhut.com	213,127	1-5	2 digits
indianpornvideos.com	1,569	1-5	2 digits

Table 3: Number of comments and rating systems for the analyzed websites.

With regards to the rating systems implemented by the different websites, some websites display five stars allowing the users to select a number of stars. `xvideos.com`, differently from the above, displays **yes** and **no** buttons allowing the user to say if they liked the video or not. This difference is captured in the column Rating of Table 3. Another difference is in the way they display the average rating. Some websites display a more precise average by showing 2 digits while others display an integer number of stars. `xvideos.com` just divides the number of **yes** answers to the total number of answers and displays a rating.

6.1 Investigating Ratings

Figure 3 shows the rating distributions for different categories for the websites `youporn.com`, `jizzhut.com`, `xvideos.com`, and `indianpornvideos.com` . As we explained above, the rating systems for all websites except for `xvideos.com` are exactly the same in nature. We normalize the ratings for comparison purposes and we therefore bring them to a scale from 1 to 100 with 100 corresponding to the highest possible rating.

One can observe the similarities with regards to ratings observed across the different websites shown in Figures 3a, 3b, and 3c. As in the previous section with regards to views, low quality amateur videos are rated lower across all websites (Category A). Again, the same effects as for views can be observed for category B.

As we noted above, `xvideos.com` (Figure 3b) has a different rating systems compared to `youporn.com` (Figure 3a) and `jizzhut.com` (Figure 3c). For `youporn.com` and `jizzhut.com`, when compared to `xvideos.com` the first two websites do not have a lot of videos rated lower while for `xvideos.com` there are 5% to 10% depending on the category more ratings lower than 70%. This is most probably because the **yes**, **no** rating system tends to have some users who would have given a slightly lower rating pick the No rating because of no other options. One of the websites that displays rounded ratings namely, `xhamster.com` even though the fewer data points available has very close ratings to `youporn.com`.

Also note Figure 3d representing the ratings observed for `indianpornvideos.com`. Its ratings are noticeably lower than the ratings observed for all the other websites. As we have said, `youporn.com` and `xvideos.com` are popular across the world while `indianpornvideos.com` is a website popular in India with most of the videos containing indian performers. The categories observed are however very different than the categories for all other websites mainly because they represent local items and the website seems aimed particularly at that. The rating system used by `indianpornvideos.com` is exactly similar to the one used by `youporn.com` and therefore the significant difference in the ratings distribution can be
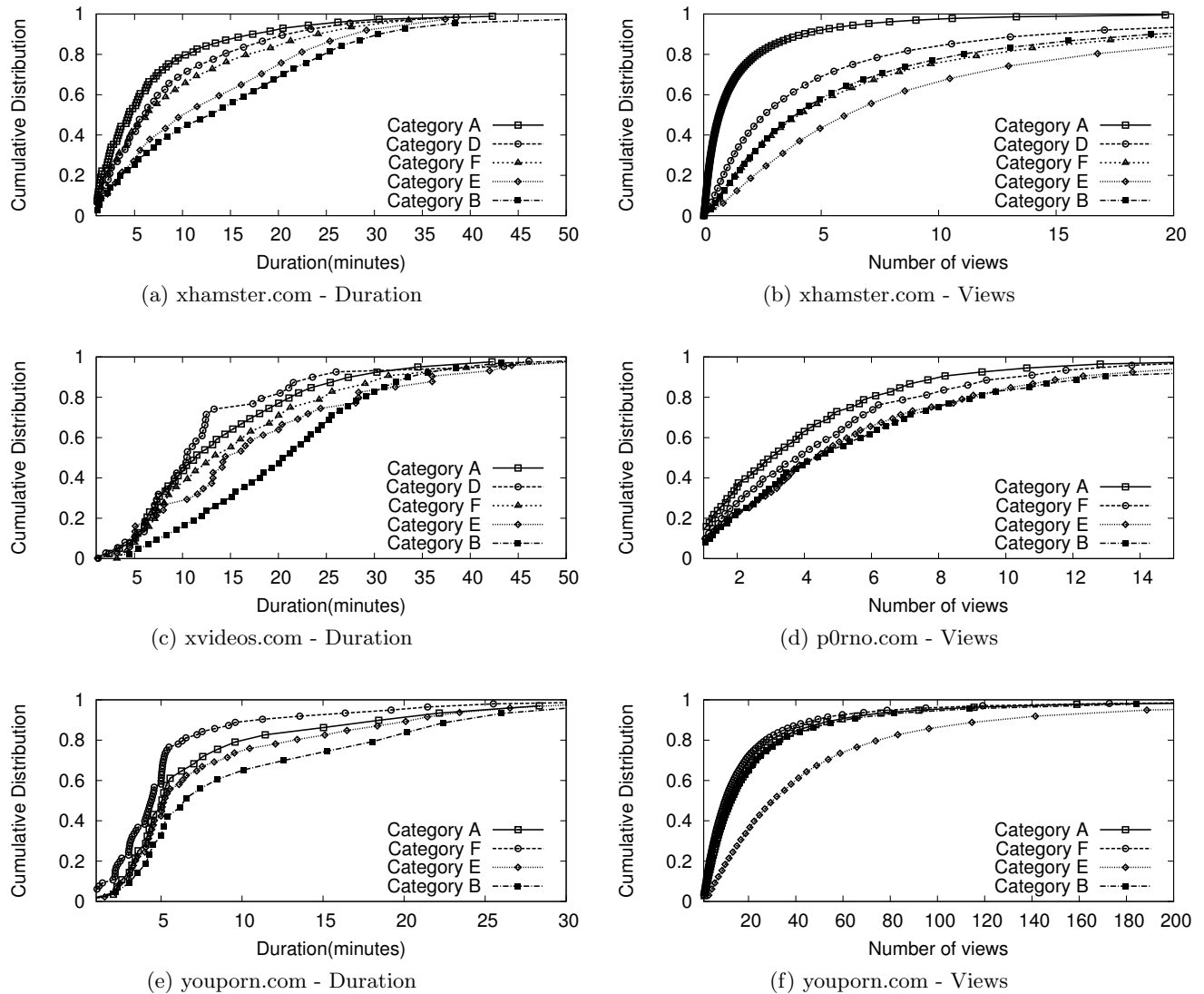
(a) xhamster.com - Duration

(b) xhamster.com - Views

(c) xvideos.com - Duration

(d) p0rno.com - Views

(e) youporn.com - Duration

(f) youporn.com - Views

Figure 2: Duration and the average number of views for different categories across different websites.

explained by a difference among the populations that access the website.

A similar tendency for users rating items lower despite of similar rating systems can be observed across the two websites popular in France `tubeadulte.com` and `p0rno.com`. The result is not shown in a figure. The tendency is not as strong as for `indianpornvideos.com`, ratings tending to be lower only by 15%-20% on average than the globally popular websites. We can note the similarity of the curves that we have extracted for `xvideos.com` and `youporn.com` implying that although the systems are different they arrive at similar results in the pornography case. Also one can note that the majority of pornography videos have ratings above 70% implying that people are fond of such content.

6.2 Investigating Comments

Similar to regular video websites, one can leave comments on video items. Some statistics regarding the number of comments per video can be seen in Figure 4a. As expected, the distributions of the comments follow the zipfian pat-

tern, characterized by having most of the energy contained within the first elements as well as a long tail, which is very expected for human produced data. As can be seen, `xhamster.com` and `youporn.com` are the most active communities in terms of comments with `xhamster.com` being closer to the distribution observed for `youtube.com` in [17]. a sample of videos while our data is a complete crawl of the studied services.

Figure 4b shows the distribution of comment sizes for different websites. The steps observed for `jizzhut.com` are caused by the large spam activity observed on the website. Most of the comments that make up the step are just spam links to various services. All the other curves are expectedly smoother. We filter out all spam links in the analysis that follows by taking out any comment that contains a URL inside. That filters out the steps for `jizzhut.com`. One can also observe the relative shortness of comments as a majority of comments for most services we analyze are lower than 50 characters in size.
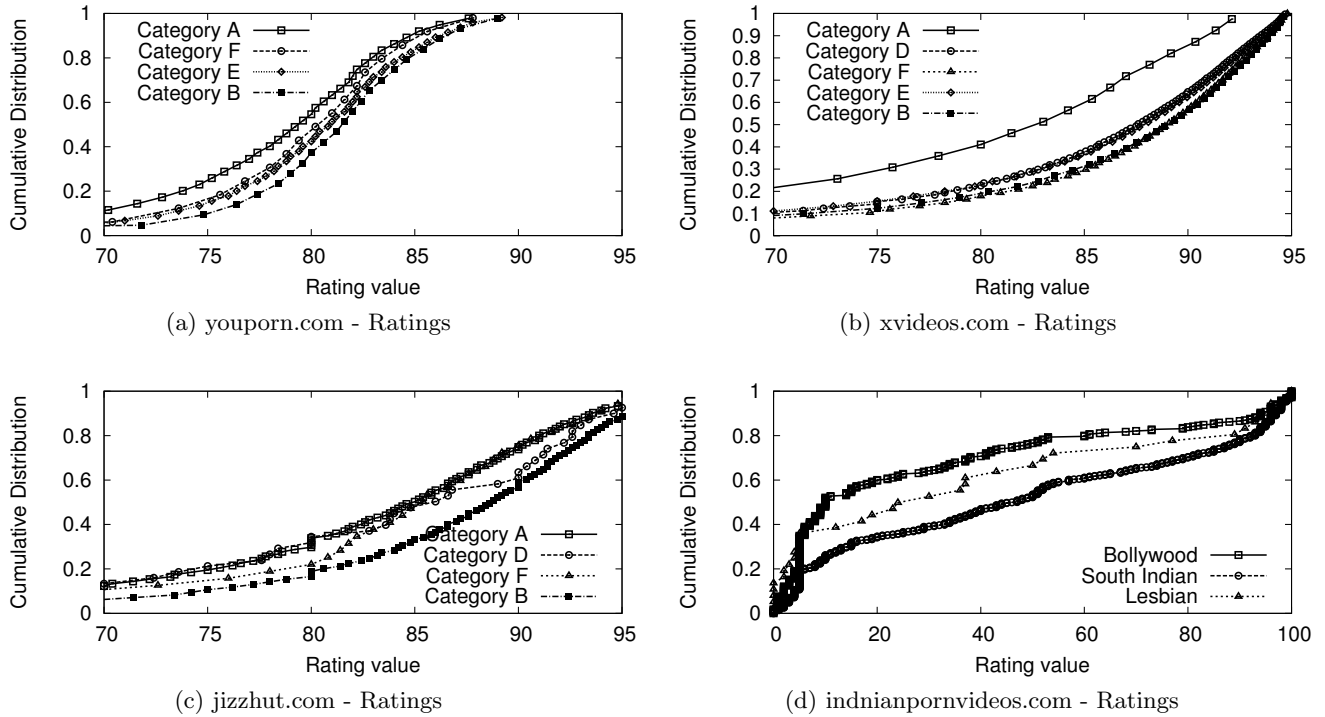
(a) youporn.com - Ratings

(b) xvideos.com - Ratings

(c) jizzhut.com - Ratings

(d) indnianpornvideos.com - Ratings

Figure 3: Ratings distribution for different websites.

(a) Comments per video
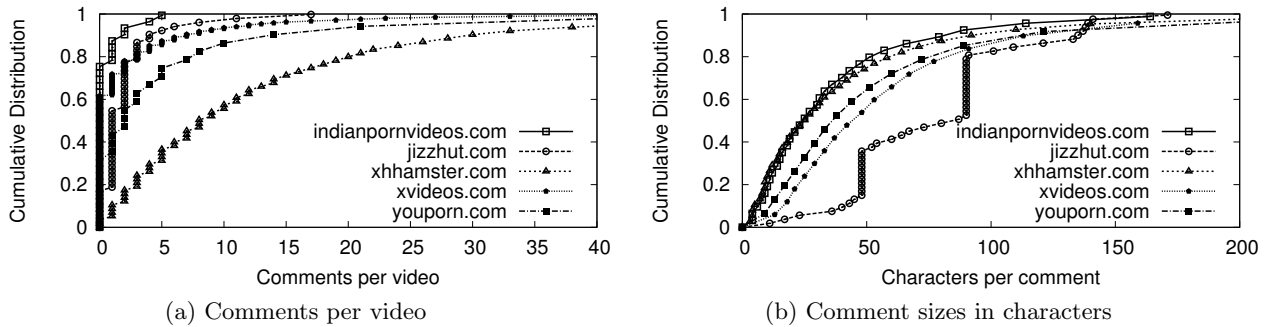

(b) Comment sizes in characters

Figure 4: Statistics regarding the number of comments per video and comment sizes.

6.3 Sentiment Analysis

For sentiment analysis we use the available SentiWordNet collection to study the sentiment scores observed across services. SentiWordNet [7] is an enhanced lexical resource for sentiment analysis and opinion mining. SentiWordNet uses WordNet [4], a lexical database that contains items that are grouped into sets of cognitive synonyms (synsets) that express a distinct concept.

SentiWordNet assigns a tripe of sentiment values (positive, negative, objective or neutral) to each synset and hence to the comprising words. The three values sum up to 1 and are between 0 and 1. We use the methodology employed in [17] by assigning a sentiment value to each comment by computing the averages for positive, negative and objective over all words in the comment that have an entry in SentiWordNet. In order to analyze the relationship between video ratings and sentiments expressed in comments we di-

vided video comments into three classes depending on the rating of the video they were extracted from:

- **Low rating.** In this class we have comments that come from a video with a rating between 1 and 33 on a scale of 1 to 100.

- **Middle rating.** In this class we have comments that come from a video with a rating between 34 and 66 on a scale of 1 to 100.

- **High rating.** In this class we have comments that come from a video with a rating between 67 and 100 on a scale of 1 to 100.

We compute sentiment value triplets for all websites except p0rno.com and tubeadulte.com that do not have comments. We filter comments that are not in english by using an online language detection service.

(a) xhamster.com negativity   (b) xhamster.com positivity   (c) xhamster.com objectivity

(d) youporn.com negativity   (e) youporn.com positivity   (f) youporn.com objectivity

(g) jizzhut.com negativity   (h) jizzhut.com positivity   (i) jizzhut.com objectivity
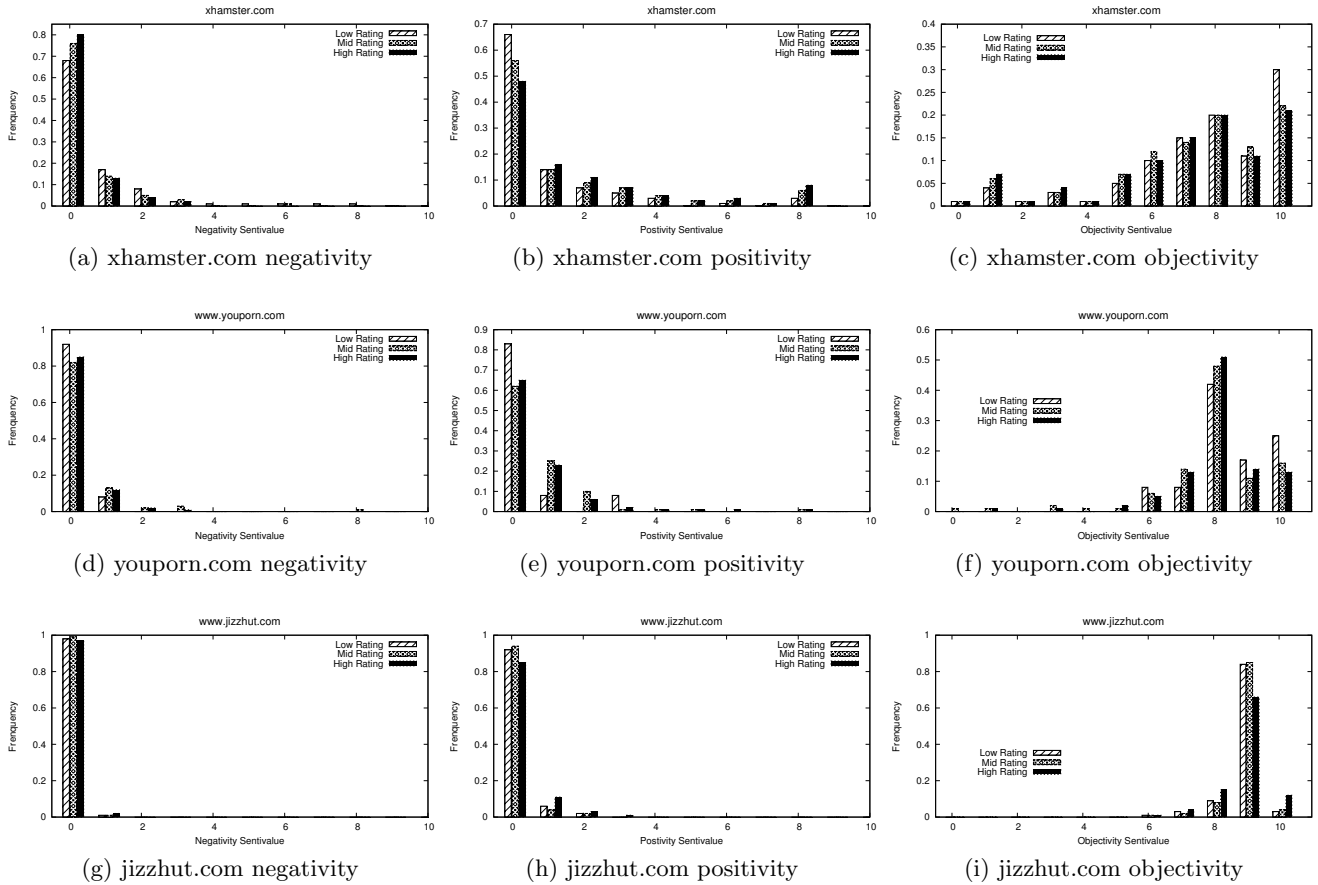
Figure 5: Sentiment positivity, negativity, and objectivity for three websites.

Figure 5 displays the histograms for objectivity, positivity, and negativity sentiment values across comments for `xhamster.com`, `youporn.com`, and `jizzhut.com`. The different bars correspondt to the categories we identified above and the scales of 1 to 10 are scaled from the sentiment values from 0 to 1 that we computed in the previous step. A few things can be mentioned here. One can note the similarity of sentiment values across the three websites. Most comments are low in negativity across the three websites concurring the overall attitude towards pornography that we previously observed in the ratings. A majority of comments are low on positivity also but we can observe some comments that are high in positivity, less for `xhamster.com` and `jizzhut.com` and more for `youporn.com`. Moreover, for `xhamster.com`, `youporn.com`, and `jizzhut.com` we can see that the higher rated videos have generally associated higher positivity values. No such behavior is observed for `indianpornvideos.com` (not shown) however which would make sense considering the lower ratings we previously observed across this service.

Second, we can note the high objectivity observed for pornography comments. This is opposite to the polarizing behavior observed for some `youtube.com` videos by the authors of [17]. Puzzled by this, we decided to investigate and we manually examined several hundred randomly selected comments across all services. With regards to this, our results can be explained as follows. A majority of comments can be described as corresponding to information seeking

behavior. Basically users are interested in either finding related videos, related materials and they are asking other users that reply where such videos or related content can be found. Some other questions include queries regarding the name of the performers in the videos, again to locate perhaps other material containing the same performers. Other comments include mostly neutral observations regarding performer body characteristics or regarding the different acts displayed.

In light of this behavior outlined in the comments we can further hypothesize that a more powerful recommendation system could be built by analyzing more thoroughly the natural language of the comments and separate the links to the items that users recommend to each other. A new recommandation system could combine heuristics built from understanding the natural language employed by users to recommend themselves content with analysis of items frequently watched.

Further, one interesting fact is that other than the observed positivity for higher rated content, we cannot find other signs of correlation between video ratings and comment sentiment values. This suggests that users prefer using ratings to express their satisfaction/dissatisfaction and comments for informative purposes as we have shown above. Also, with regards to differences across different categories of videos, none can be observed with regards to sentiment values across comments. Basically the same high objectivity as observed in Figure 5 can be observed across comments

belonging to all categories. This is in contrast with the differences across categories that we observed earlier in the paper for views and ratings. This strengthens our point above that users employ ratings to express their satisfaction with the content and comments just for informative purposes.

7. RELATED WORK

Social research related to pornography can be classified in two broad areas. One broad class of research work tries to look at the root causes that drive individuals to search for pornographic material. Some interesting results can be found in [13, 14] where the authors try to link election outcomes with pornography seeking behavior. Another broad class of research tries to find the effects of the easy access to pornographic material that the internet provides on individuals [10, 15, 16]. Differently from all of these, we are interested more in how users interact with the service and try to understand the types of content they are interested in.

The authors of [11] study the infrastructure that drives pornography websites. Opposite from this, our study is aimed at understanding the type of content users prefer and the feedback they leave. Unrelated to pornography but related to our work, other studies have looked at how users search for information [9], how users watch videos [5, 8] and the preferences they have, and used sentiment analysis to analyze the comment feedback that they leave on videos from `youtube.com` as in [17]. Things are quite different in the internet pornography context as we have seen.

8. CONCLUSIONS

In this paper we have taken a look at internet pornography consumption. We have analyzed several free video pornography websites. Among them, some are globally popular and some popular in specific regions of the globe. We have analyzed aspects related to two internet pornography related behaviors, mainly watching different types of internet pornography, and rating and commenting on pornography content.

Our findings are as follows: (i) With regards to users watching specific types of pornography, we find a strong user bias towards higher quality, often professionally produced video items. Also longer items in terms of duration are preferred by users. (ii) With regards to users rating items, they give higher quality items higher ratings. A bias in rating items across websites popular in different regions can also be observed. (iii) However, with regards to comments we can notice strong objectivity and similarities across the analyzed websites. To conclude, this paper tries to understand more about internet human behavior and put light on a previously uncharted part of the world wide web.

9. REFERENCES

[1] Alexa Adult Website Statistics. `http://www.alexa.com/topsites/category;1/Top/Adult`.

[2] Alexa Website Statistics. `http://www.alexa.com/`.

[3] Internet Pornography Statistics. `http://www.familysafemedia.com/pornography_statistics.html`.

[4] *WordNet: An Electronic Lexical Database (Language, Speech, and Communication)*. The MIT Press, 1998.

[5] CHA, M., KWAK, H., RODRIGUEZ, P., AHN, Y.-Y., AND MOON, S. I tube, you tube, everybody tubes:

analyzing the world's largest user generated content video system. In *IMC* (2007).

[6] DÖRING, N. M. Review: The internet's impact on sexuality: A critical review of 15 years of research. *Computers in Human Behavior 25* (September 2009), 1089–1101.

[7] ESULI, A., AND SEBASTIANI, F. Sentiwordnet: A publicly available lexical resource for opinion mining. In *LREC* (2006).

[8] GILL, P., ARLITT, M., LI, Z., AND MAHANTI, A. Youtube traffic characterization: a view from the edge. In *IMC* (2007).

[9] HOLSCHER, C., AND STRUBE, G. Web search behavior of internet experts and newbies. *Journal of Computer Networks 33*, 1-6 (2000), 337 – 346.

[10] KRISTIAN DANEBACK, MICHAEL ROXX, S.-A. M. Characteristics and behaviors of sexual compulsives who use the internet for sexual purposes. *Sexual Addiction and Compulsivity 13*, 1 (2006), 53–67.

[11] KRUEGEL, C., KIRDA, E., PLATZER, C., HOLZ, T., AND WONDRACEK, G. Is the internet for porn? an insight into the online adult industry. In *WEIS* (2010).

[12] LIBERATORE, M., LEVINE, B. N., AND SHIELDS, C. Strengthening forensic investigations of child pornography on p2p networks. In *CoNEXT* (2010).

[13] MARKEY, P., AND MARKEY, C. Changes in pornography-seeking behaviors following political elections: an examination of the challenge hypothesis. *Evolution and Human Behavior* (2010), 442–446.

[14] MARKEY, P., AND MARKEY, C. Pornography-seeking behaviors following midterm political elections in the United States: A replication of the challenge hypothesis. *Computers in Human Behavior 27* (May 2011), 1262–1264.

[15] PETER, J., AND VALKENBURG, P. Adolescents' exposure to a sexualized media environment and their notions of women as sex objects. *Sex Roles 56* (2007), 381–395.

[16] PETER, J., AND VALKENBURG, P. M. The influence of sexually explicit internet material and peers on stereotypical beliefs about women's sexual roles: Similarities and differences between adolescents and adults. *Cyberpsychology behavior and social networking 14*, 9 (2011), 511–517.

[17] SIERSDORFER, S., CHELARU, S., NEJDL, W., AND SAN PEDRO, J. How useful are your comments?: analyzing and predicting youtube comments and comment ratings. In *WWW* (2010).

[18] SPINK, A., OZMUTLU, H. C., AND LORENCE, D. P. Web searching for sexual information: an exploratory study. *Information Processing and Management 40* (January 2004), 113–123.

[19] TRESTIAN, I., RANJAN, S., KUZMANOVIC, A., AND NUCCI, A. Measuring serendipity: connecting people, locations and interests in a mobile 3g network. In *IMC* (2009).