Measuring Web Feature Impacts in BitTorrent-like Systems

Sirui Yang[1], Hai Jin[1], Bo Li[2], Xiaofei Liao[1], Hong Yao[1], Qi Huang[1]

[1]Service Computing Technology and System Lab
Cluster and Grid Computing Lab
School of Computer Science and Technology
Huazhong University of Science and Technology,
Wuhan, 430074, China
hjin@mail.hust.edu.cn

[2]Department of Computer Science
Hong Kong University of Science and Technology
Clear Water Bay, Kowloon, Hong Kong, China
bli@ust.hk

ABSTRACT

In *Peer-to-Peer* (P2P) file sharing systems, the attributes of resource description can influence the user behavior, especially on resource selection. However, this has been only qualitatively speculated but lacks of quantitative analysis. In this paper, we carry out a systematically quantitative study on the impact of these attributes presented in the form of web features, by measuring the largest BitTorrent website in CERNET. The measurement lasts for 31 days, and there are 168,610 records containing 11,228 distinct resources collected. The result is two-fold. On one hand, it confirms the above qualitative speculation; on the other hand, it shows more significant findings: (1) with the highlight feature on popular items, the downloads of each resource yield to a long-tail distribution however deviating from Zipf Law; (2) publications with attracting titles disseminate 1.9 times faster than others; (3) publisher authority feature does not evidently help the system escaping from malicious resources' pervasion; (4) other features such as taxonomy and size also influence users' choice. We further demonstrate the implications of the web feature impact for system designers and potential attackers.

Categories and Subject Descriptors

C.2.4 [**Distributed Systems**]: Distributed applications

General Terms

Measurement, Performance, Security

Keywords

Peer-to-Peer, File sharing, Measurement, Zipf, Highlight effect

1. INTRODUCTION

Peer-to-Peer (P2P) file sharing has been under extensive study for years [1][2][3][4][5]. The new-fashioned pattern not only accelerates the content delivery by taking full advantage of edge resources in Internet, but also increases the system reliability by eliminating the single failure point.

In some first generation P2P systems, e.g., Gnutella [2], users enter keywords and initiate searches to find resources. While in many second generation systems, e.g., BitTorrent [1][6], user needs to visit a portal page first to view existing published resources, and then decides which one to download. The portal web pages where resource publications are listed and descriptions are presented can help users for information acquisition and resource selection. It is believed that rather than plain texts, the diversified appearance styles of these web sites will more directly affect user behaviors, especially on resource selection. For example, the hot resource (downloaded by many people) being highlighted with a striking color will be noticed and decided to download with a higher probability by following users than normal ones.

There are many measurement researches on P2P file sharing networks. Sariou et al [7] investigated Napster and Gnutella systems and showed there was significant difference among end host capacity. Gummadi et al [8] studied KaZaA traffic trace and found out that the "fetch-at-most-once" property would influence the file request distribution. Pouwelse et al [6] analyzed the components uptime and the flash crowd case in BitTorrent network and pointed out these will affect the performance. Besides above factors, it is also speculated that the attributes of the web sites for resource exhibition (i.e., the web features) also exert impact on the system performance. This is quite significant in BitTorrent-like systems where users can only fetch resources from portal web pages. However, there is little work to quantify the impact in P2P file sharing systems.

Different from prior studies, we focus on the web feature impacts in BitTorrent-like systems and use quantitative analysis to examine the qualitative speculation. We have conducted a measurement of the web pages of "5QZone" [9], from October 19 to November 19, 2007, lasting for 31 days. 5QZone is the largest BitTorrent website in CERNET (*China Education and Research NETwork*). Our crawler has collected totally 168,610 records in which there are 11,228 distinct published resources. Notice that there are many records published before the first day of the measurement. So even though the measurement lasted for 31 days, we collected resources being published over 73 days. With the crawled data, we particularly concern on the following web features: the presence of the detail information of publications including the title, the file size, the genre, the publisher, the numbers of temporal seed providers [1] and downloading users, as well as the aggregate completion times of downloading.

Our contribution is to find following results: (1) The historically aggregate completion times of all the publications follows a long tail distribution which, however, obviously deviates from the Zipf Law [8][10]: median results in the measurement are more than in the law; (2) More than 53.6%

QShine 2008, July 28-31, 2008, Hong Kong, Hong Kong.
 ISBN 978-963-9799-26-4
DOI 10.4108/ICST.QSHINE2008.3906

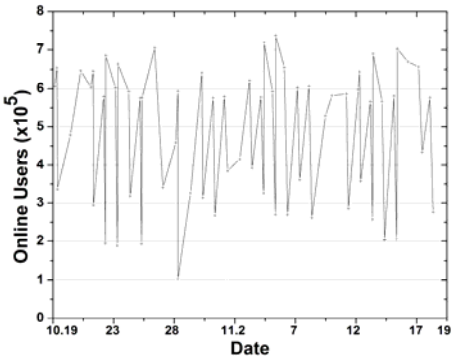

Fig 1: User dynamics

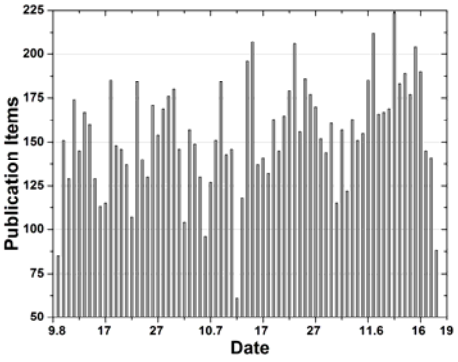

Fig 2: Daily publications

publications with attracting words in the title, such as flirting words and the names of some hot movies, will become wide spread finally. The spread rate of such publications is 1.9 times faster than normal ones; (3) Common publishers can disseminate their resource as fast and wide as authoritative publishers do, i.e., publisher authority feature does not prevent the malicious content from spreading; (4) Besides the downloading status impact, the resource taxonomy and size also take effect on the users' selection.

The remainder of this paper is organized as follows. In Section 2, we present the related works. In Section 3, we introduce the 5QZone website and its web features. The measurement methodology and results, as well as corresponding analysis are brought forward in Section 4. In Section 5, we present the implication of the measurement results. Finally, we draw the conclusion and the future work in Section 6.

2. RELATED WORK

Besides the differences among the various P2P file sharing protocols, the performance of such systems is also influenced by external factors, such as the host heterogeneity, information acquisition manner, user behavior and web features.

For the first generation P2P file sharing protocols in which Napster [11] and Gnutella [3] are the representatives, Sariou et al [7] conducted an in-depth measurement of the end user hosts in the networks. They characterized the bottleneck bandwidths and IP latencies between peers, connection and disconnection rates, shared and downloaded file quantities, etc. Their measurement shows that there is significant heterogeneity and lacks of cooperation across peers.

For the second generation systems in which BitTorrent [1], KaZaA [2] and eMule [4] are the representatives, Pouwelse et al

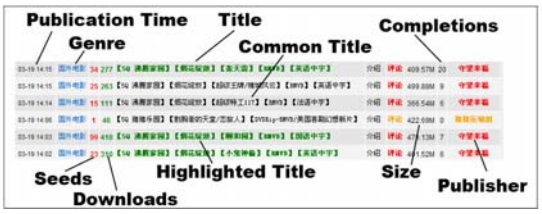

Fig 3: A segment of resource table

[6] measured BitTorrent/Suprnova sites and focused on the moderator system to investigate the system availability, integrity, flash-crowd handling and download performance. Their results show that the global component adopted in BitTorrent is useful to increase system integrity and improve the download performance.

Gummadi et al [8] observed the file sharing behaviors and the impact of key system parameters in KaZaA networks by analyzing the traffic trace. They tried to explore the user characteristics including the "fetch-at-most-once" property, the patience and activity during downloading, and the object characteristics including the multi-workload, object dynamics, and the non-Zipf distribution of user requests on different resources. They explained that KaZaA is not Zipf because that the "fetch-at-most-once" property was different from the HTTP request in WWW (which is Zipf).

Besides above studies, many other works have been done on the second generation P2P systems, i.e., BitTorrent-like ones. Qiu et al [12] built a simple fluid model to analyze the BitTorrent system and studied the steady state network performance. They also performed a series of experiments to validate the model. Guo et al [13] considered the multiple torrent property which is more real in practical systems and analyzed the corresponding performance. Piatek et al [14] concentrated on the "free riding" phenomenon and proposed their analysis on the incentive mechanism.

Our work is different from all prior works in that: (1) we focused on the relationship between the web features adopted in BitTorrent web sites and the user behavior, especially resource selection; (2) we understand the deviation of the downloading request distribution from the Zipf Law by inducting the impact of highlight feature on popular resources; (3) we consider the implications of these web feature impacts both from the performance improvement and potential system risk perspectives; (4) our data are collected from 5QZone pages which are gathered by the site's centralized components at background, thus the error is much less than initiating crawls from the client/end-host side.

3. 5QZONE AND WEB FEATURES

Our measurement is based on 5QZone web site for following reasons: (1) its scale is large; (2) it has the direct exhibition of seeds, downloads and completions of all resources. These statistics data are obtained by global servers which lead to fewer errors than the trace-based methods; (3) it has more web features than most (if not all) other BitTorrent sites in China.

For clarity, we define a publication of a piece of resource (i.e., a file or a directory containing several files) as a "publication item", or "item" for short; and a row in the resource table as a "record". A record is sourced from an item but containing more temporal detail information and can be repeatedly crawled during the measurement. We denote the publication date of the item as the "date", and the time we get the record as the "crawling time". The number of seeds, downloading users and aggregate

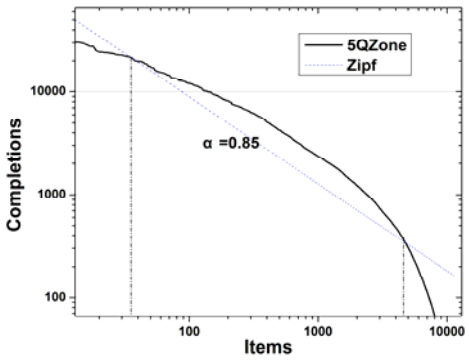

Fig 4: Item completion distribution

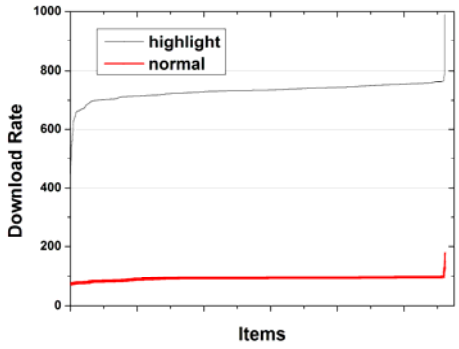

Fig 5: Download rate in highlight and normal periods

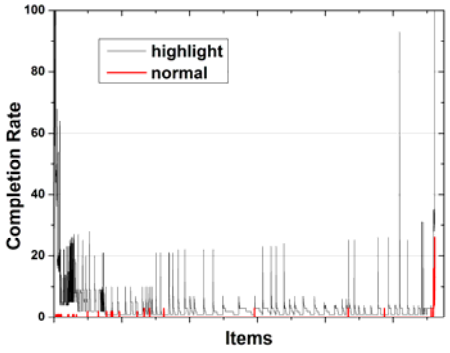

Fig 6: Completion rate in highlight and normal periods

completion times are respectively called "seeds", "downloads" and "completions" for short.

5QZone was created in October, 2003, has become the largest general BitTorrent site in CERNET especially for its user amount. Figure 1 shows the user amount dynamics and Figure 2 presents the published items in each day during our measurement.

When the user tries to publish his resource, he needs to first register a user ID. Then he needs to give a title for this publication which is not necessarily the name of the resource (e.g., a movie's name). Further, he should select a genre for his resource. Finally, the resource is published in this site and available for other users.

A "resource table" is used to organize all published resources in 5QZone, as shown in Figure 3. The title, the genre, the size, the publication time and the publisher's ID will be listed in the resource table. Moreover, the temporal number of seeds and downloading users, as well as the aggregate completion times of downloading the resource, will be directly shown in this table. Another important web features in 5QZone is the highlight features. Popular items are highlighted with green and authoritative publishers are highlighted with red or orange. A popular item is an item being downloaded and seeded by more than 300 users. An authoritative publisher is either employed or invited by 5QZone. Notice that for unknown reasons, large resources (larger than 2.8GB) are also highlighted with green. We treat this case separately in our measurement.

4. RESULTS AND ANALYSIS

In this section, we present the measurement results and corresponding analysis. As mentioned, 5QZone presents the statistics of item distribution and updates them periodically collected from a specific server. Thus we directly collect the data from the web pages. This methodology suffers from fewer errors compared to search-based measurement from BitTorrent clients.

4.1 Metrics

Our measurement began from October 19 to November 19, 2007, lasting for 31 days. We conducted one to three crawls each day and 63 times in total. Each two crawls are several hours away.

As described in Sec 3, we investigate metrics as follows: (1) Highlight effect: How is the distribution of the completions of all items? Is the popular highlight feature impacting on the user behavior significantly? What if it is compared with the normal distribution without the feature, namely the Zipf Law? (2) Title wording: The title of an item seems to be important, since users can not preview the content of the published resource before downloading them. Thus, what is the probability for an item with an alluring title to be popular? In opposite, what is the probability for an item with a plain title to be popular? (3) Publisher authority: The authoritative publishers are considered to be more likely to provide clean content and better service (e.g., longer seeding time) than normal ones. However, is it true that users tend to choose authoritative publishers' items? (4) Size effect: Users know that small items are unlikely to cover much content. A larger file may be of a higher quality, e.g., a DVD version of a same movie is clearer than a highly compressed RMVB version. But is a larger item necessarily more popular? (5) Genre interests: Users have different background, such as careers, education and hobbies. Are the users who are interested in the listed 13 genres equivalent in quantity? What is the most popular genre among them? These metrics are measured and analyzed in the following subsections one after another.

4.2 Highlight Effect

Completion distribution: First, we present the completions for all the 11,228 items by the end of the measurement, as shown in Figure 4 (the bow curve) with the log-log coordinates. The completion distribution is shown a long-tail property. It first confirms the speculation that the download request distribution is not even but influenced by user behaviors. Specifically, there are 476 items which have never been finished by any users except their publishers. Only 139 items have been finished for more than 10,000 times, and the most popular item has been completed for 67,497 times.

As we known, the popularity for Internet web pages yields to the Zipf Law. For comparison, we use 28 pairs of data selected from our collected result to obtain a most fitting Zipf distribution (the straight line with α=0.85) with Least Square Method. The two curves are not well matching, which means the distribution deviates from the Zipf Law noticeably. The conclusion that P2P

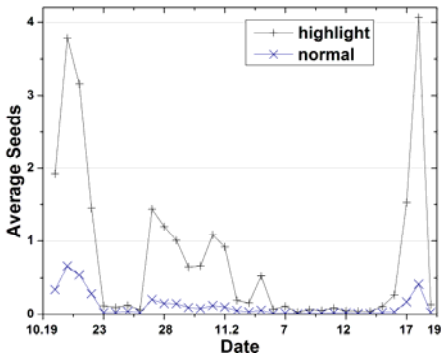

Fig 7: Average seeds for highlight and normal items

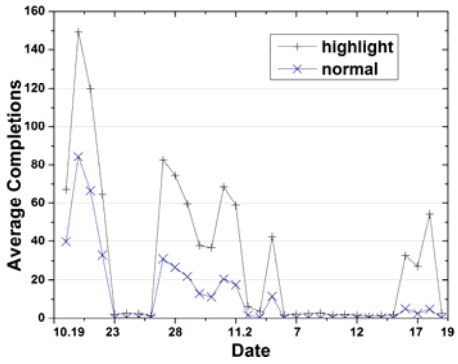

Fig 9: Average completions for highlight and normal items

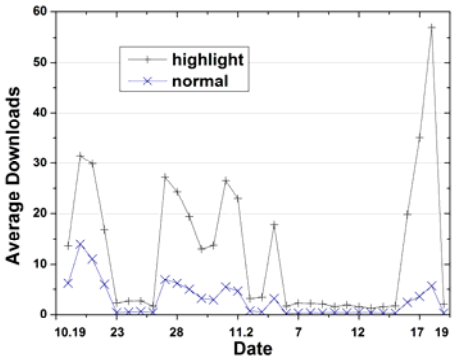

Fig 8: Average downloads for highlight and normal items

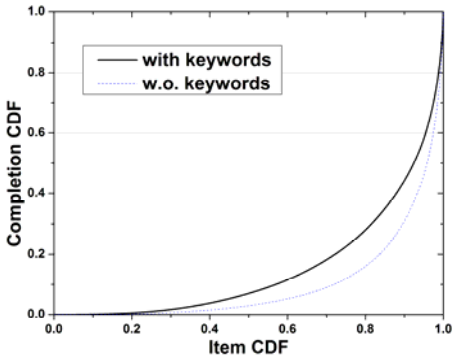

Fig 10: Completion of items with and without keywords

file sharing request is not Zipf is similar to [8]. However, our result differs from [8]: the curve in [8] has a nearly flat beginning, for the "fetch-at-most-once" property reduces the values of popular items; our result does not have such beginning, while it simply increases the median values. This is because the highlight feature takes effect. We notice that users have two kinds of attitudes before joining the BitTorrent system: having desire on specific recourses previously and having no such demands. Users with the former attitude will definitely download the on-demand item once he sees it; and users with the latter attitude just visit the portal pages to see if there is something interesting. Without the highlight feature, both of them may be attracted by really popular resources based on their personal judge. However, the highlight feature makes items with thousands of downloads and those of hundreds of downloads appear in the same way (i.e., green colored) and thus, help some median popular items become attractive. As the proportion of the median popular completions increases, the other two parts of completions is lessened.

Download and completion rate: An item may be popular and highlighted for one period but turn to normal for another period. To quantify the predominance of highlight feature impact, we compare the spread rates for both highlight and normal items. For example, an item has been highlighted during two periods: [Oct. 11, Oct. 14] and [Oct. 16, Oct. 18], and the download during the two periods are 600 and 400, respectively. Then its download rate during highlight periods is (600+400)/[(14-11)+(18-16)]=200 times per day. The completion rate is similarly obtained for both highlight and normal periods. Figure 5 and 6 show the results. For observing convenience, results in Figure 5 are sorted on the download rate of the items, and the item indices in Figure 6 are accordant to Figure 5.

The graphs also show that the download rate of highlight items is about 6~8 times as the rate of normal ones. Besides, the curves in Figure 5 are sorted to be mono-increasing while the curves in Figure 6 are fluctuating. It indicates that the completion rate is heterogeneous to the downloading rate both for highlight and normal periods. Namely, more temporal downloading users do not imply a larger completion amount for a single item. The reason of this phenomenon is that there are many "free-riders" [14] [15] in P2P file sharing systems.

Daily global study: Besides investigating each single item, we also conduct a global measurement of the proportion of highlight items and corresponding seed, download and completion distribution on each day. The results are presented in Figure 7, 8 and 9. Figure 7 shows that on average, there are globally only a few seeds every day. However, the figures also present that the three figures have homogeneous trends. This looks inconsistent with the result in Figure 5 and 6, where the free-riding phenomenon is observed. This means that free-riding case is dependent on single items. However, from a global perspective, the free rider phenomenon is neutralized, as more downloads leads to more completions. This is accordant to the fact that a few contributors in the system support for most services as expressed in [14][15].

4.3 Title Wording

In fact, many users have the motivation and intelligence to attract more downloads by giving an alluring title. We call this

Table 1: Completions for some hot programs

Title	Genre	Total	Highlight
Lust Caution	CN Movie	20	18 (90%)
Prison Break	Foreign Movie	196	112 (57.1%)
Guess Guess	Enter. & Sport	29	16 (55.2%)
Bleach	Cartoon	49	7 (14.3%)

Table 2: Publisher authority and item popularity

Authoritative & Popular	Authoritative & Unpopular	Ordinary & Popular	Ordinary & Unpopular
1572 (14.0%)	3671 (32.7%)	3368 (30.0%)	2617 (23.3%)

case as "title attraction" or "title trap". There are two kinds of title traps: the hot keyword case and the hot program case.

Hot keywords case: According to our own experience, we list 19 frequently used attracting keywords in 5QZone, such as "*sexy*", "*beauty*", "*love*", "*funny*", etc. Figure 10 is the CDF diagram for the completions of keyword titled items and normal ones. The figure presents that for keyword contained items, it is much possible to spread wider than normal ones. There are 4,446 items with these keywords and 6,782 items without such keywords. However, the keyword contained items have attracted 5,873,865 completions, while the normal ones have been finished for only 4,692,697 times. With simple arithmetic we can find that keyword contained items have been completed for 1.9 times on average as normal ones do. Also, that the slope of the "with keyword" curve is increasing more slowly than the "without keyword" once. It means the larger completions are not contributed only by some extreme hot items, namely, any item with the keywords has a higher probability to attract a mass of completions.

Hot program case: Besides keywords in the title, the content itself is also important. Thus, we use 4 well-known programs to further study the title wording issue. Table 1 shows the total items related to these programs and the percentage of highlighted versions for them. Something interesting happens with "*Lust Caution*" and "*Prison Break*". Google provides 6,180,000 results about "*Lust Caution*" and 22,200,000 ones about "*Prison Break*" the teleplay. It seems as the "*Prison Break*" is more well-known than "*Lust Caution*". However, the percentage of items being highlighted for "*Lust Caution*" is 90%, while it is only 57.1% for "*Prison Break*". The main reason is that "*Lust Caution*" is a *new arrival* while "*Prison Break*" is a bit older. Another possible reason is that "*Lust Caution*" is a Chinese movie and the users of 5QZone are mostly Chinese, so they are fonder of this movie. Besides, it is also influenced by the number of copies of the item, i.e., "*Lust Caution*" has less copies and users focus on a few choices which increases the percentage of highlight cases.

4.4 Publisher Authority

A user who downloads resources from 5QZone can be anonymous, but he must register an ID if he wants to publish items. To ensure the quality of service, 5QZone has many employed publishers and partner publishers. In the following paragraphs, these publishers are called "authoritative publishers", while others are called "ordinary publishers".

In Sec 2 we introduced that 5QZone has another highlight feature: to distinguish authoritative publishers' IDs with red or orange color in the resource table. The authoritative publishers are more likely to provide true content and better quality of service (e.g., longer seeding time) than ordinary ones. So will the highlight on publisher authority affect the selection of users for item downloading? Table 2 presents the statistics of the percents of popular items provided by the two kinds of publishers. It shows that there is no direct relationship between the publisher's authority and the item's popularity. Namely, an authoritative publisher's item does not necessarily become popular. It is because users just want to get required resources, and concern on the item itself much more than its publisher. Oppositely, since the authoritative publishers do not use alluring words in the title in most cases, their resources may receive less downloads than ordinary ones (as the first column indicates).

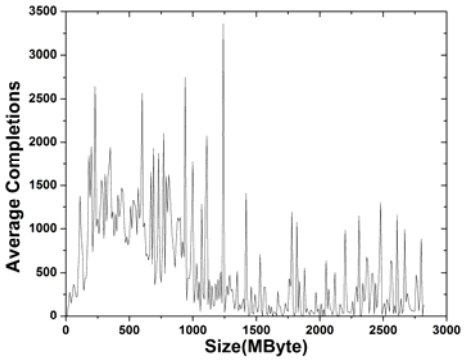

Fig 11: Average completions with different sizes

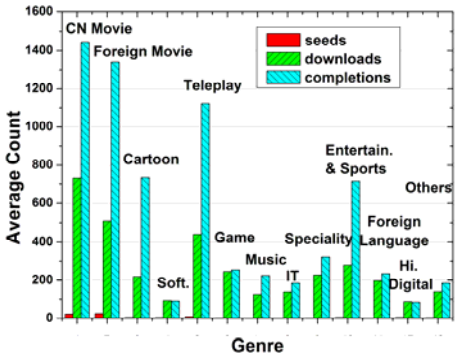

Fig 12: Delivery situation for 13 genres of items

4.5 Size Effect

A user seems to be less interested in the size of the resource compared with other features, as the hard disks are getting larger now. However, some careful users also take the size as a reference when selecting items. We measure the average completions for items in different sizes. So, we divide the size into 280 groups from 0 to 2.8GB (2800MB) every other 10MB. All sizes larger than 2.8GB are considered as in the 281th group. Then, we calculate the total completions and item amounts in each group and get the corresponding average completions.

The result is in Figure 11. The curve fluctuates in a wide range because in some groups there are only a few items and the average completions are not representative. We analyze the curve according to its trend. It presents that items in median size (about 100MB to 1GB) are most wide spreading, which receive more than 1000 completions on average. Large items (more than 1.5GB)

are not completed for many times. The large items have more bytes to transfer and take longer downloading time. This further makes some users not to select such items as they are not willing to wait too long. The implication is different from [8] where users are patient enough in KaZaA. Small items (tens of MB) are also less frequently complete. The reason is that small items are unlikely to contain attracting content: they are mostly some MP3 records, specialty materials or common software, which are desired by only a few users.

4.6 Genre Interests

In 5QZone, there are totally 13 genres for users to classify their resources, as marked in Figure 12. Items in different genres will attract different user groups. For example, items for leisure (such as movies and teleplays) are more public interesting than specialty materials. Figure 12 also shows the result of seeds, downloads and completions for all genres of items. The most popular genre is the "*Chinese Movie*", "*Foreign Movie*", "*Teleplay*", "*Cartoon*" and "*Entertainment & Sport Program*". In opposite, "*High Digital*" and "*Software*" are the least public interesting genres. This result fits for our impression on people's taste and hobbies in real world. The high digital items, especially movies, are of excellent quality and supposed to attract many downloads. However, they always require some powerful machines to playback. And they are often larger and fewer users are interested in them than anticipated, as analyzed in Sec 4.5.

5. IMPLICATIONS

Similar to the two-fold measurement results, the implications are also two-fold, i.e., for system designers to improve the performance and for malicious users to enhance attacks.

Performance improvement: we believe the designers can meliorate the system global performance from following perspectives. (1) Boosting unpopular items: an unpopular item is not necessarily low quality. To accelerate the spread of unpopular items can increase the information share level and the cooperation between peers. The site manager can introduce score mechanism and highlight not only popular items but also high score items. However, this requires human interventions and user honesty. (2) Increasing publisher authority impact: in current 5QZone, publisher authority has been considered but not employed to play an important role. To help users get better service, the effect of highlighting authoritative publishers should be strengthened. The manager can hash resource content and determine same resources in different publications. Then if one of the resources is published by authoritative publishers, it will be set to the front of other publications or "recommend" marks can be provided. (3) Building incentives: the web features can facilitate the incentives built in BitTorrent. The manager can require users to login first to download (not only provide) resources. For users who contribute much, the manager can promote them to authoritative publishers or publish their user IDs to praise them.

System attack: The impact of these web features can benefit the attackers, including the content polluters [16] and virus injectors, etc. (1) For the title attraction: attackers can give flirting titles to their publications and gain much more downloads rather than giving a normal title. For example, a virus injector can title the item containing his executable virus files as "*funny games*". (2) For the genre interests: attackers may also choose a most public interesting genre to classify the malicious content. For instance, it is better for the content pollution company to publish fake movies than music records since the "*Chinese Movie*" and "*Foreign Movie*" genres are more public interesting than "*Music*". (3) For the publisher authority: our measurement implies that an attacker does not need to acquire a high reputation before conducting the attack. In other words, simply a reputation mechanism is not enough to prevent malicious content from spreading, because even the publisher is suspicious, they will risk themselves to downloading the attracting item. (4) For the highlight effect: if the attacker employs a Botnet [17], i.e., many fake users, to download his resource, the system will highlight his item according to current mechanism which will deteriorate the harm as it leads more users to notice the malicious item.

Therefore, BitTorrent website should pay much attention to specify the web features and provide effective mechanisms to improve performance and avoid possible attacks.

6. CONCLUSION AND FUTURE WORK

In this paper, we present the detailed measurement of a BitTorrent website, 5QZone. The measurement help understand the impact of the web features in BitTorrent-like websites, including the highlight effect, the title wording, the size effect, the genre interests and the publisher authority. From the results, the conclusions are drawn that (1) the completions of all items have been deviated from the Zipf Law mainly due to not only the studied "fetch-at-most-once" property but also the highlight effect; (2) title attraction plays an important role in interfering the users' selection on items; (3) publisher authority does not work currently to enhance the system security; (4) web features such as resource genre and size also affects user behaviors. Besides, we discuss the implications for the results that they may be used in a two-fold way: for performance improvement and system attacks.

Some other user behaviors and web features, such as the time in a day when the item is published, and the file format of the items and other systems where users obtain resource information without portal web pages are left as our future work.

7. ACKNOWLEDGEMENT

This work was supported in part by grants from RGC under the contracts 616406 and 616207, by a grant from National Natural Science Foundation of China (NSFC)/Research Grants Council (RGC) under the contract N_HKUST603/07, by a grant from HKUST under the contract RPC06/07.EG27, by a grant from Nokia Research under the contract NOKIA001-BL.06/07, by NSFC grant 60433040, by NSFC and RGC Joint Research Scheme grant 60731160630, and by the Research Fund for the Doctoral Program of Higher Education Grants No. 20050487040.

8. REFERENCES

[1] B. Cohen, "Incentives Build Robustness in BitTorrent", in *Proc. P2P Economics Workshop'03*, Berkeley, CA, June 2003.

[2] J. Liang, R. Kumar, and K. W. Ross, "The FastTrack Overlay: A Measurement Study", *Computer Networks* (Elsevier), Vol.50, No.6, pp.842-858, April 2006.

[3] D. Stutzbach and R. Rejaie, "Characterizing the two-tier Gnutella topology", in *Proc. ACM SIGMETRICS'05*, Alberta, Canada, June 2005.

[4] Y. Kulbak and D. Bickson, "The eMule Protocol Specification", eMule Project, http://sourceforge.net.

[5] I. Stoica. , R. Morris, D. Karger, M. F. Kaashoek, and H. Balakrishnan, "Chord: A Scalable Peer-to-peer Lookup

Service for Internet Applications", in *Proc. ACM SIGCOMM'01*, San Diego, CA, August 2001.
[6] J. A. Pouwelse, P. Garbacki, D. H. J. Epema, and H. J. Sips, "The BitTorrent P2P File-Sharing System: Measurements and Analysis", in *Proc. IPTPS'05*, Ithaca, NY, February 2005.
[7] S. Saroiu, P. K. Gummadi, and S. D. Gribble, "A Measurement Study of Peer-to-Peer File Sharing Systems", in *Proc. SPIE/ACM MMCN'02*, San Jose, CA, January 2002.
[8] P. K. Gummadi, R. J. Dunn, S. Sariou, S. D. Gribble, H. M. Levy, and J. Zahorjan, "Measurement, Modeling and Analysis of a Peer-to-Peer File-Sharing Workload", in *Proc. USENIX SOSP'03*, Bolton Landing, NY, October 2003.
[9] 5QZone, http://www.5qzone.net.
[10] T. Yamakami, "A Zipf-Like Distribution of Popularity and Hits in the Mobile Web Pages with Short Life Time", in *Proc. PDCAT'06*, Taiwan, China, December 2006.
[11] Napster, http://www.napster.com.
[12] D. Qiu and R. Srikant, "Modeling and Performance Analysis of BitTorrent-like Peer-to-Peer Networks", in *Proc. ACM SIGCOMM'04*, Portland, OR, August 2004.
[13] L. Guo, S. Chen, Z. Xiao, E. Tan, X. Ding, and X. Zhang, "Measurements, Analysis, and Modeling of BitTorrent-like Systems", in *Proc. ACM IMC'05*, Berkeley, CA, October 2005.
[14] M. Piatek, T. Isdal, T. Anderson, and A. Krishnamurthy, "Do incentives build robustness in BitTorrent?" in *Proc. USENIX NSDI'07*, Cambridge, MA, April 2007.
[15] M. Sirivianos, J. H. Park, R. Chen, and X. Yang, "Free-riding in BitTorrent Networks with the Large View Exploit", in *Proc. IPTPS'07*, Bellevue, WA, February 2007.
[16] J. Liang, R. Kumar, Y. Xi, and K. W. Ross, "Pollution in P2P File Sharing Systems", in *Proc. IEEE INFOCOM'05*, Miami, FL, March 2005.
[17] M. A. Rajab, J. Zarfoss, F. Monrose, and A. Terzis, "A Multifaceted Approach to Understanding the Botnet Phenomenon", in *Proc. ACM IMC'06*, Rio de Janeiro, Brazil, October 2006.