Pollution Measurement in Bittorent

G. N. Purohit
Department of Computer Science, Banasthali University, Banasthali, 304022, India

Deepika Sainani
Department of Computer Science, Banasthali University, Banasthali, 304022, India

Urmil Malik
Department of Computer Science, Banasthali University, Banasthali, 304022, India

ABSTRACT
Since the invent of Napster an earlier version of peer–to- peer [P2P] network , P2P has emerged as the most significant technique for sharing content across the network[1].The traffic volume exchanged among the peers and the number of member peers in a P2P network has risen to a significant level .The reason for this significant rise is a absence of a centralized authority and distributed nature of the network .However, these reasons for significant rise of P2P network also makes the P2P networks vulnerable to various pollution attacks such as denial of service attack, content poisoning, collusion sybil etc. Liang et. al [2] has discussed pollution attacks in the Fast Track file sharing network. This network has also been used for distributing copy-righted material illegally, and their legal owners thought that their assets are distributed illegally through P2P network. Since the file distribution through P2P network is very fast, the producers/developers of multimedia data like music, films, television etc are much concerned with the apprehension that their assets are distributed illegally through P2P network. For controlling the illegal distribution of their assets, they fought legal battles in courts, but could not succeed in it .Then they employed the same P2P network, using P2P pollution attack, to stop the illegal distribution of their data files. In this paper, an evaluation is done to measure the damage to the BitTorrent distributing pirated files under certain pollution attacks, specifically Index poisoning and Uncooperative peer attacks, and to assess whether this strategy is successful..

Keywords
P2P networks, Index poisoning, Uncooperative peer, BitTorrent.

1. INTRODUCTION
BitTorrent is a popular file sharing and content distribution protocol. It is an improved P2P mechanism which is achieved by adding a tracker server in each sharing network and thus clubs the advantages of client server in the network .However, BitTorrent protocol, like any P2P network is equally vulnerable to pollution attacks.

An early study conducted between 2008 and 2009[3] has shown, "P2P file sharing accounts for more than 65% of traffic in many countries and traffic of BitTorrent stands as number one among all other protocols. Movies, television programs, MP3 songs and MP4 videos, images documents, software and games could be distributed through P2P file sharing easily, efficiently and quickly."

The popularity of P2P technology for file sharing is employed in two diverse directions: (i) Ethical or Legitimate sharing of files, (ii) Unethical or illegal sharing of copy-righted/pirated documents like films, MP4 videos etc. Simultaneously unethical applications also generate different kinds of cyber crimes, for example phishing seam. For stopping the illegal sharing different types of efforts have been made. Remedy through legal courses /law suits could not help in this matter .The vulnerability of P2P to pollution attacks could be another possibility for this purpose .Index poisoning, which alters the index of files, seemed to be an alternative technique for controlling the illegal sharing of copyrighted documents .This technique was vigorously investigated by many researchers, particularly by Yoshida et.al [4], Kong et.al [3]. The P2P network is flooded with bogus file indices in index poisoning. According to Liang et.al [5], "After altering the index of illegal files (copy-righted contents), these files cannot reach to any peer, which wishes to access them through the P2P."

Besides index poisoning attack, there are some other kinds of pollution attacks, which works for stopping the distribution of BitTorrent files .In fake block attack [6]; "Discussed the implications of fake-block-attack in P2P live video streaming systems." According to Shi & Jhang [7], "Fake-block attack aims to prolong the downloading time of a file at victim peer by wasting their download bandwidth and damaging the strength utility of the swarm. The attacker makes his client to join the target swarm and then advertises that it has a large number of pieces of the file that this swarm is sharing .On receiving a request message for a leacher, the attacker responses with some fake blocks instead of authentic ones." In uncooperative peer attack, an attacker establishes TCP connection with the peers in a swarm and does not provide any content (fake or real) to the peers in the swarm .Thus it stops the downloading of the files, since it wastes the time and a leacher is provided only a fixed time slot ranging 25-50 seconds.

In this paper an attempt has been made to measure the damage in BitTorrent by pollution attacks. The evaluation of the impact on BitTorrent of some pollution attacks is done to measure the effectiveness of pollution on BitTorrent to stop the peers from downloading of the files. Two types of pollution attacks have been considered, (i) Index poisoning attack and (ii) Uncooperative peer attack. It was observed that these attacks were not much successful while attacking individually compared to their combined attacks.

The remaining of this paper is organized as follows: since attacks on BitTorrent are considered only, the BitTorrent architecture is described briefly in Section II. The two pollution attacks are described in Section III. Whereas combination of these two pollution attacks is considered in Section IV. Experimental setup and impact evaluation is given in Section V. At the end Section VI includes the concluding remarks.

2. BITTORRENT ARCHITECTURE
Since the inception by Cohen [8], BitTorrent has become the most popular file sharing and content distribution protocol. Its architecture comprises of three components, a tracker, a seeder and a leacher. Seeders are peers with entire file and

leachers comprise of non-seeder peers. The role of tracker server is to forward the arriving peers, after identifying them, to the specific sharing network, designated as swarm. A swarm is a group of peers comprising of identifiers of the same file content. In case a seeder wishes to share a file, it provides a torrent file including the metadata for whom it is desired, and announces to the tracker about the shared file [9].

In sharing a file through BitTorrent, the supplier develops a torrent and sends it to the internet. BitTorrent clients (peers) download it and distribution process starts. Before distribution of the content, the content provider splits the content into multiple chunks or pieces, normally each piece of size 256KB. Further each piece is again subdivided into sub pieces, each of size 16KB. A metainfo file is created which contains information like URL of the tracker, data file names, their length etc. by the content provider. A willing peer for joining the swarm retrieve the metainfo file,[10].

After receiving the announced message from the seeder, the tracker searches for an exactly matching swarm and if it is unable to trace such a swarm it generates a new swarm for the file. If a leacher requests the tracker for a swarm using torrent file; the tracker provides information about the requested swarm to the leacher. The tracker working as a coordinator/ facilitator provides the requested information to the leacher also. Once a leacher obtains swarm information from the tracker, it communicates with the peers in the swarm directly, without the help of the tracker. The leacher requests for the pieces of the file through sending interest message to the peers. Further it also announces the information about the pieces held by it by sending HAVE and BITFIELD messages. Using these messages the leacher finally downloads the desired file completely.

There are interval chunk selection mechanisms also, such as rarest first, strict priority and others [8]. In addition, choking mechanisms such as optimistic unchoking and antisnubbing also run on these protocols. With these mechanisms BitTorrent peers can share files quickly and get improved quality of interconnection among peers [11].

The BitTorrent is driven by two protocols; "Tracker HTTP Protocol (THP) and Peer wire Protocol (PWP)" [10]. The former provides the BitTorrent service between peers and a tracker by using HTTP and the later is used for the exchange of pieces described in the metadata between the peers. As such the BitTorrent is vulnerable to various attacks. Trackers face threats similar to threats faced by an ordinary HTTP server. If an attacker has the authority to control a tracker, then by misusing the vulnerability of the THP it can disrupt the swarm information so that the peers are faced with server attack situations such as denial of service [DOS]. In addition peers can face attacks from other peers by exploring the PWP by the attackers. The flow chart for BitTorrent protocol exchange is given in Fig. 1.

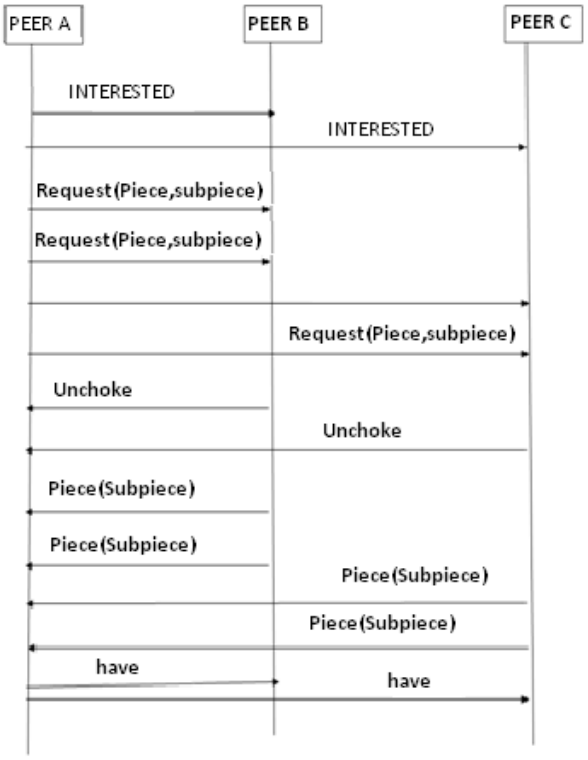

Fig.1. Flow chart for BitTorrent File sharing

3. TWO POLLUTION ATTACKS IN P2P

In this section, index poisoning attack and uncooperative peer attack are briefly discussed. The description is limited to the need of this paper.

3.1 Index Poisoning Attack

As described in the introductory section I, the purpose of index poisoning is to control / damage the file distribution in a P2P network. File indices are used in BitTorrent file sharing by the peers for obtaining information about the downloadable files. Index of a file contains information like the URL of the tracker, data file name, their length, location of desired content etc, so that it can be retrieved easily. On a query from a neighboring peer, if the requested peer has a copy of matching index, it forwards the desired index to the questioner. The recipient peer can hold a prescribed number of other files also. In addition , "each copy is given a lifetime and is periodically updated by the file owner, so that the index of popular files will be held by many peers, where the way of distributing and retrieving indices depend on the underlying file sharing software". [14]."

In P2P network the tracker works as a coordinator/facilitator and provides necessary information to the peer about other peers on request. An interested peer for downloading a file in BitTorrent initially informs the tracker about it. Then the tracker provides the information about the peers downloading the same file or the swarm. But here is a problem, while providing the information to the peer the tracker or seeder does not authenticate the peer request and does not verify the availability of the requested content. Making use of this, "An attacker deliberately advertises large quantity of invalid peer information of the targeted content. This invalid information could be random content identifiers

that do not correspond to any existing content, IP addresses that do not correspond to any peer participating in the file sharing system or unavailable service port numbers at participating peers. So when a user attempts to download the content corresponding to the task, his BitTorrent client always fails to establish connection with other peers, due to high probability of connecting to invalid peers". [5]

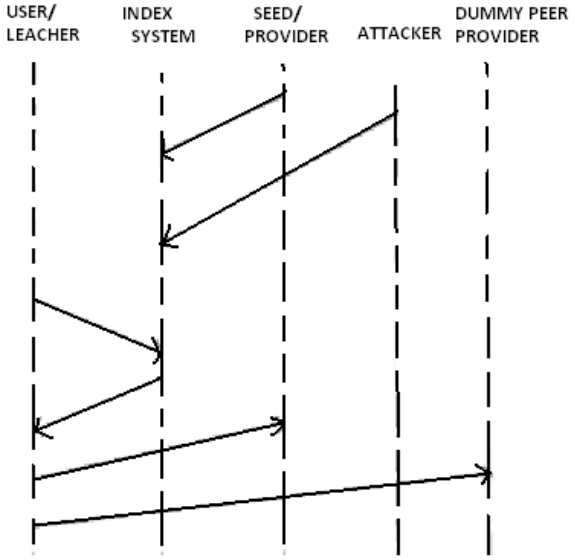

Fig.2. Flow chart for index poisoning attack

Once the majority of the peer information in the tracker has been poisoned, the information provided by tracker about targeted task is invalid. However, the index poisoning cannot completely control the illegal distribution of copy righted content in BitTorrent file distribution. For instance, "If any peer establishes connection with a seeder, it will make the other peers to establish connection with benevolent peers rapidly through interconnecting. During the index poisoning, if one of the downloading peers connects to a benevolent peer, most of the peers in the swarm will start downloading rapidly and index poisoning will fall."[3].

The flow chart for index poisoning is given in figure 2.

3.2 Uncooperative Peer Attack

In this type of attack the downloading time is extended and in turn the downloaded bandwidths at peers end are wasted. Further it damages swarm strength and its robustness. In this attack, "The attacker exploits the BitTorrent message exchange protocol and joins the targeted swarm establishing TCP connections with many peers in the swarm. The purpose is to obstruct the downloading client. Depending on the selected victim client, these peers keep on simply sending BitTorrent handshake messages without ever sending any file (Fake or Genuine) or they continue to send continuously keep alive message without delivering any blocks. Since the peer connections is for limited duration (Often set for 50 sec.), connecting to numerous chatty peers can drastically increase the download time of contents. The effectiveness of this attack increases if a significant fraction of peers has been victimized."[15]

The flow chart for uncooperative peer attack is given in fig. 3.

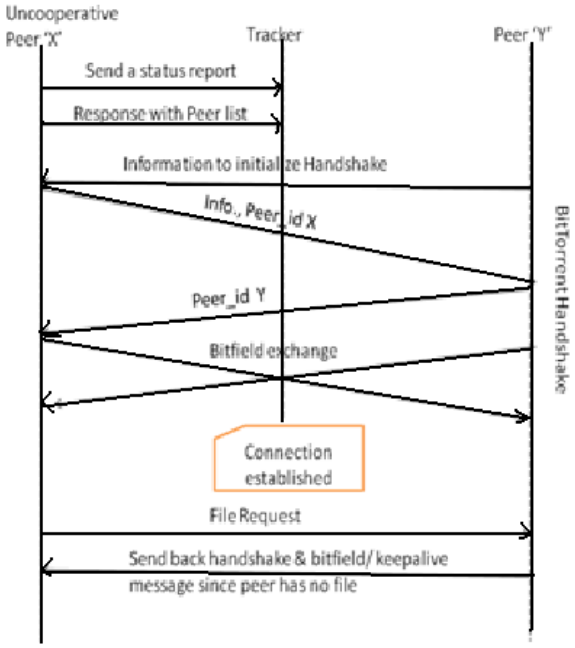

Fig.3. Flow chart for Uncooperative Peer Attack

4. HYBRID POLLUTION ATTACK

In this section amalgamation of two previously described pollution attacks is considered. The purpose of doing this is to improve the pollution and ultimately to control the file distribution in BitTorrent. Under this scheme, the two pollution attacks act in a complimentary manner; the index poisoning denies connection among peers and uncooperative peer attack damages the swarm strength and its robustness. Further it wastes the resources available with the peers. In this hybrid scheme, "the attacker initially starts to pollute the BitTorrent swarm. Simultaneously it infuses a substantial quantity of invalid peer information in the tracker and also the peer information of the hybrid pollution attacker. When a leacher joins the swarm for downloading, it initially obtains the information of other cooperative and genuine peers from the tracker. But under the influence of this hybrid pollution, the peer fails in establishing connection with cooperative and genuine peers due to the index poisoning. Then the peer spends plenty of time in establishing connection with invalid peer and thus the average download speed of such BitTorrent swarm is reduced."

There is a possibility that after much struggle, the downloading peers (leachers) may be able to join a genuine peer, overcoming the effect of index poisoning. However, its probability is very low. If some genuine peers falsely considers the hybrid pollution attacker as a genuine peer and establish connection with it (attacker) will download fake data from it. The flow chart for this hybrid pollution scheme is given in Figure 4.

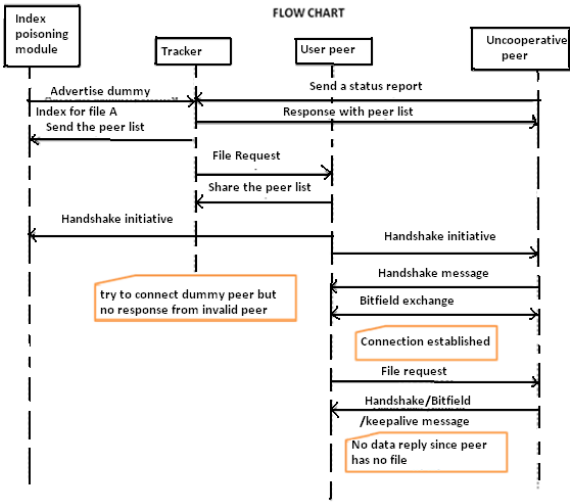

Fig.4. Flow chart for Hybrid Pollution Attack

5. POLLUTION IMPACT EVALUATION

In order to know how effective these pollution attacks, described in previous sections, were able to control the file distribution in BitTorrent, experimental work was carried out. The results show that individual attacks were less effective compared to the hybrid pollution attack, however, none is able to control it completely. Following tools were employed in the experiment:

Bit Comet 1.11 as BitTorrent client,

Bit Comet Tracker 0.5,

File size for distribution 180 MB

P2P Network: 32 normal peers, 10 seeders.

For correct evaluation, each experiment was repeated 10 times. 10 sample peers are selected randomly. Evaluation for each pollution attack was considered separately. The experiment was carried in the institution's laboratory.

5.1 Index Poisoning Attack

The following two indices are considered for evaluation.

1) ACT: Average Connection Time- The time a peer uses to join a genuine peer.
2) TNCF: Number of times that none of sample peers connect to a genuine peer.

The results are shown in Table I. In the Table TNCP=p, 0<=p<=10 means that none of the sampling peers could connect to the genuine peer in all the b-experiments.

The ACT increases with the increase of invalid peer information, the reason being that the leacher continuously asks the tracker about the peer information, since it is unable to contact a genuine peer. However, the TNCF does not increase unless the invalid peer information is very high. Thus index poisoning cannot completely stop file distribution; it can only prolong the connection.

Table 1. Index Poisoning Attack

No. of Invalid Peer information	ACT (in seconds)	TNCF
None	9.0	0
50	15.4	0
100	18.7	0
500	30.6	1
1000	187.3	3

5.2 Uncooperative Peer Attack

For its evaluation following indices are considered:

1) ACST: Average Connection Sharing Time- a genuine peer spends to download data from uncooperative peer.
2) ADF: Average Download Failure-represents number of peers which could not connect to the uncooperative peer
3) ACR: Average Completion Rate-denotes the average downloading completion rate of the peers which fail to complete the downloading.

The results are shown in Table II; Y denoted the ratio of the number of genuine peers to the number of attackers.

Table 2. Ucooperative Peer Attack

Y	ACST (in sec.)	ACR	ADF	ADT(Seconds)
0(No attackers)	109.1	None	0	453.2
1:1	189.2	None	0	1023.7
1:2	209.7	None	0	1350.2
1:10	472.0	18.4	10	Very Large

It is inferred from the above results that ADT increases with the increase of uncooperative peers and same happens with ACST. In case Y=1:10 it is not possible to download, however, Y=1:10 is almost impossible to achieve.

5.3 Hybrid Pollution Attack

For its evaluation the following indices are only considered:

1) ADF
2) ACR which are same as defined in case (b).

The results are given in Table III. As before Y stands for the ratio of the number of genuine peers to the number of uncooperative peers, while X stands for the number of invalid peer information advertised in the tracker.

Table 3. Hybrid Pollution Schemes

X	Y	ADF	ACR
250	1:1	8	23.4%
250	1:2	8	12.6%

500	1:1	8	16.3%
500	1:2	8	3.8%

From the result it is inferred that most of the sampling peers could not complete the downloading. The simultaneous increase in X and Y further makes it difficult for the downloading peer. The impact of hybrid scheme is better than the individual pollution attacks. A limited number of downloading is able to join the genuine peers, but most of these are uncooperative peers. This damages the robustness of the swarm.

6. CONCLUSION

In this paper three pollution schemes: (i) index poisoning scheme (ii) uncooperative peer scheme and (iii) hybrid pollution scheme are discussed in BitTorrent File Distribution. The impact of pollution in each of these three schemes is evaluated. It is inferred that hybrid pollution scheme is more effective compared to other schemes individually, but none is able to stop the distribution completely.

The recent scheme is scalable. Though this scheme has a lot of advantages, it has disadvantages too. It is still not a perfect one because there are issues like trust and certification, anonymity, security which are yet to overcome. In future we can combine some more approaches to make an efficient hybrid pollution scheme.

7. REFERENCES

[1] L. Hu and Z. X. Lu, "Downloading Trace study for BitTorrent P2P performance measurement and analysis," Peer-to-Peer Networking and applications, Vol. 5, Issue 4, pp.384-397, 2012.

[2] J. Laing, R. Kumar, Y Xi and K. W. Ross, "Pollution in P2P files sharing systems," 24th IEEE International conference INFOCOM Miami, Florida, USA pp. 1566-1590, 2005.

[3] Jie Kong, Wandong Cai, Lei Wang and Qiushi Zhao, "A study of pollution on BitTorrent", The 2nd International Conference on Computer and Automation Engineering (ICCAE), Vol. 3, 2010.

[4] M. Yoshida, S. Ohzahata, A. Nakao, and K. Kawashima, "Controlling file distribution in winny network through index poisoning" in Proc. of International Conference on Information Networking (ICOIN), 2009.

[5] J. Liang, J.N.Naoumov and K. W. Ross, "The Index Poisoning Attack in P2P File sharing systems," Proceedings of IEEE INFOCOM-2006.

[6] Lian, Q., Peng, Y., Yang, M., Zhang, Z., Dai, Y., Li, X., "Robust incentives via multi-level tit-for-tat," research articles in Concurr. Comput. Pract. Exper, pp. 167–178, 2008.

[7] Shi, J, Zhang H., "A protocol based countermeasure to BitTorrent Fake-Block Attack," J. Computational Inform. Sys. Vol. 12, No.8, pp 5211-18, 2012.

[8] B. Cohen,"Incentives Build Robustness in BitTorret," Proc. of the Ist workshop on the Economics of Peer-to-Peer systems, Berkley, CA, June 5-6, 2003.

[9] S. Kim, S. Choi, B. Roh., "A survey of attacks on the BitTorrent protocol from its operational viewpoints," research notes in Information science (RNIS) Vol. 14, June 2013.

[10] Hyunggon Park, Rafit Izhak Ratzin and Mihaela van der Schaar, "Peer-to-Peer Networks: Protocols, Cooperation and Competition," Streaming Media Architectures, Techniques, and Applications: Recent Advances, 2010.

[11] V. atlidakis, M. Rousspoulos and A. Dellis, "Changing the unchoking policy for an Enhanced BitTorrent," Parrallel processing Springer Berlin Heider berg, 2012.

[12] K. C. Sia, "DDOS vulnerability analysis of BitTorrent protocol," University of California, Las Angeles, USA, 2007.

[13] K. E.Defrawy, M. Groka and A. Markupoulou, "BitTorrent misusing BitTorrent to launch DDOS attack," USENIX, June 2007.

[14] Yuusuke Oookita & Santoshi Fujita, "Index Poisoning Scheme for P2P file sharing system with Low spatial and Network cost," Bulletin of networking, computing systems and software, Vol. 4, No.1, pp. 27-33, 2015.

[15] Marti Ksionsk, Ping Ji, Weifeng Chen, "Attacks on BitTorrent – An Experimental Study," e-Forensics-2010 LNICST 56, pp. 79-89, 2011.