Hybrid framework for mitigating illegitimate Peer Nodes in Multimedia file sharing in P2P

Mr. Ramesh Shahabadkar[a], Dr. Ramachandra V. Pujeri[b]
[a]KResearch Scholar, Anna University, India,
[b]KGiSL Institute of Technology, Coimbatore, India

Abstract: Peer to Peer network is one of the frequently used application in terms of file sharing over a global large network. In such types of network, there can exists an illegitimate peer node who will attempt to have an unauthorized access to premium digital content. As it is very difficult to catch hold of the intruder or the illegal client inside the network, the proposed system will at least attempt to prevent access to the digital content available. The proposed system is focused on multimedia file sharing where the protocol will be designed in such a way that whenever any unauthorized peer node will attempt to download any premium digital content, the proposed model will assign a poison chunk of data to be forwarded to the illegitimate client. This phenomenon will result in exponential increase of download time rending discouragement to download the same file by the illegal downloader. While the protocol assures the secure delivery of the cleaner chunk of data to the legitimate client.

Keywords: Peer to peer, security,. content poisoning.

I. INTRODUCTION

The peer-to-peer multimedia-sharing network is recent trend in delivering large amount of file to enormous number of participants [1]. The unauthenticated users are punished by means of poisoning the network, which raises immense legal issues in the field of P2P network. The prominent reason of the illegitimate multimedia file sharing are found as peer nodes which overlook the various rights as well as attempt to mitigate the various illegitimate peer nodes. Conventional content delivery networks [2] will use a huge quantity of the proxy digital content servers spread over globally distributed WANs. The content distributors need to replicate or cache contents on many servers. The bandwidth demand and resources needed to maintain these CDNs are very expensive. A P2P content network significantly reduces the distribution cost [3], since many content servers are eliminated and open networks are used. P2P networks improve the content availability, as any peer can serve as a content provider. P2P networks are desired to be scalable, because more peers or providers lead to faster content delivery.

The legitimate clients are considered as the entity in the network which obeys the digital rights management and other restrictions. Pirates are peers attempting to download some content file without paying or authorization. The colluders are those paid clients who share the contents with pirates. We use identity-based signatures (IBS) [4] to secure file indices. IBS offers the same level of security as PKI-based signatures with much less overhead. We apply discriminatory content poisoning against pirates. We focus on protection of decentralized P2P content networks. Protecting centralized P2P networks like Napster or mp3.com is much simpler than the scheme we proposed.

Our goal is to stop collusive piracy within the boundary of a P2P content delivery network. Our protection scheme works nicely in a P2P network environment. The scheme cannot stop randomized piracy in open Internet using Email attachment or any other means to spread copyrighted contents, illegally. Randomized piracy is beyond the scope of this study.

The remainder of this paper is organized as follows. Section II presents P2P content poisoning followed by related works in Section III. The proposed approach is discussed in Section IV. Section V will highlight the system model deployed in this research work followed by Implementation in Section VI and Simulation result in Section VII and finally Section VIII will conclude the research proposal.

II. P2P CONTENT POISONONG

Currently, content owners only distribute free-of-charge files such as product demo, freeware and open-source software via P2P file-sharing networks. They do not distribute content to paid customers via P2P networks due to the high risk of copyright infringement. If any of the paid content is found in P2P networks, they are pirated. In that case, content owners will use content poisoning to disrupt the illegal file distribution. The disruptions appear as failed distribution, corrupted contents, or repeated frustrations on the part of unauthorized downloading clients. The concept of content poisoning in such a P2P system is illustrated in Fig.1.

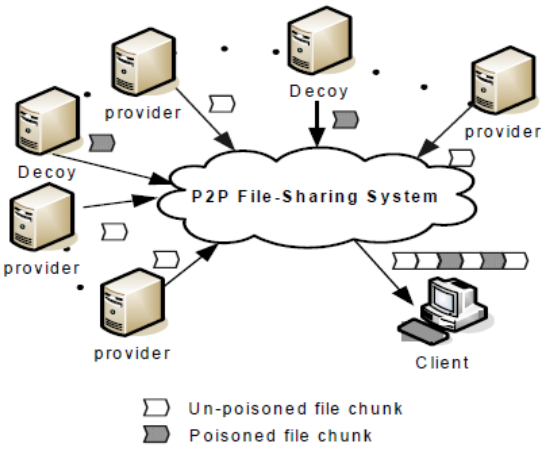

Figure 1: Content poisoning in a P2P file-sharing System, where providers generate clean file chunks and the decoys generate poisoned chunks

The illegal content providers on a P2P network are the copyright violators. They share clean file chunks to other peers. The content owners deploy a small number of decoys to poison the copyrighted files at subdivided chunk level. Unlike providers, decoys send out poisoned chunks in response to client's download requests. In the sequel, we use the term peer to refer to either a provider or a decoy, because both are not distinguishable by clients. The number of decoys divided by the total peer count is the decoy density.

For a real example in August 2006, we discovered a piracy case, by which a brand-name word-processing package was illegally distributed over a well-publicized P2P network. We hide the real names and fileID involved in order to protect all the parties involved. We refer the software file by a fake name: XYZ Pro. This file was measured to have 643.63 MB, subdivided into 3571 chunks. The file is identified by a fake fileID =XYZ1234567890Pro. When requesting the file, 56 peers on the P2P network responded, among them 8 peers are immediately available for file transfer. By deploying decoys into P2P file-sharing systems, owners want to discourage the download of copyrighted files by all clients, who did not subscribe the requested file. The major design issue of content poisoning lies in cost-effective deployment of decoys. Effective poisoning should use only a few decoys. Excessive decoying is unnecessary and cost-prohibitive. Content owner should also anticipate that those clients download pirated content off the P2P network will apply various techniques to resist poisoning.

Definition of Content Poisoning Effect

Many technical factors affect the download performance of a P2P file-sharing system, such as network topology, routing scheme, traffic congestions over the network, geographical locations of peers, etc. In order to single out and focus on the effect directly caused by content poisoning, we need to filter out all other effects. Let S be the actual file size (in Mega bytes) and D be the total number of bytes downloaded. We define the poisoning effect by:

$$\text{Poisoning Effect} = 1 - S / D$$

Poisoning Effect isolates the download effort wasted due the existence of decoys providing poisoned chunks. Its value represents the portion of downloaded bytes that are wasted due to the existence of decoys in a P2P file-sharing system. For example, in an ideal P2P file-sharing system where no decoy was present, we have S = D, meaning the client received exactly the same amount of bytes as the actual size of the file. Thus, the poisoning effect becomes zero. The poisoning effect approaches 100%, if D becomes extremely large, meaning most download requests failed. The content owner does not care of the copyright violator's interest. Their goal is to achieve higher poisoning effect. For example, 0.9 poisoning effect implies that the client must download 10 copies of the same file in order to retrieve an un-poisoned version. Since the owner's cost is directly related to the number of decoys deployed, a cost effective approach is to deploy just enough decoys.

III. RELATED WORK

Nicolas e.t. al [5] conduct a measurement study of content availability in four of the most popular peer-to-peer file sharing networks, in the absence of poisoning, and then simulate different poisoning strategies on the measured data to evaluate their potential impact.

Ruichuan e.t. al [6] proposed poisoning-resistant security framework for P2P content sharing systems which is able to defend against the content poisoning attack effectively and efficiently.

Stephanos [7] study current peer-to-peer systems and infrastructure technologies in terms of their distributed object location and routing mechanisms, their approach to content replication, caching and migration, their support for encryption, access control, authentication and identity, anonymity, deniability, accountability and reputation, and their use of resource trading and management schemes.

Dan [8] has proposed a fully functional identity-based encryption scheme (IBE). The scheme has chosen ciphertext security in the random oracle model assuming a variant of the computational Diffie-Hellman problem.

Lei Guo [9] has presented a novel and efficient design of a scalable and reliable media proxy system supported by P2P networks. This system is called PROP abbreviated from our technical theme of "collaborating and coordinating PROxy and its P2P clients" with an objective is to address both scalability and reliability issues of streaming media delivery in a cost-effective way.

Ernesto Damiani [10] propose a self-regulating system where the P2P network is used to implement a robust reputation mechanism. Dumitriu [11] consider the file targeted attacks in current use in the Internet, and we introduce a new class of p2p-network-targeted attacks.

Tom [12] propose a P2P protocol that integrates the functions of identification, tracking and sharing of music with those of licensing, monitoring and payment.

Balachander [13] discuss how CDNs are commonly used on the Web and define a methodology to study how well they perform. A performance study was conducted over a period of months on a set of CDN companies employing the techniques of DNS redirection and URL rewriting to balance load among their servers.

Stefan [14] examines content delivery from the point of view of four content delivery systems: HTTP web traffic, the Akamai content delivery network, and Kazaa and Gnutella peer-to-peer file sharing traffic. Pablo Rodriguez [15] has discussed that although P2P technology has been widely associated with the distribution of pirated content and has been subject to a barrage of attacks (e.g. DoS, spoofing and content pollution), and there are ways to decrease the risks associated with distributing content using P2P technology.

Matthew [16] has analyzed the cost of the implementing the redundant task allocation in order to prevent illegal cases over internet. Kevin Walsh [AR] employed a novel voter correlation scheme to weigh the opinions of peers, which gives rise to favorable incentives and system dynamics by presenting simulation results indicating that system is scalable, efficient, and robust.

Runfang Zhou [17] proposes a gossip-based reputation system (GossipTrust) for fast aggregation of global reputation scores which leverages a Bloom filter based scheme for efficient score ranking. GossipTrust does not require any secure hashing or fast lookup mechanism, thus is applicable to both unstructured and structured P2P networks.

Currently, many reputation models have been proposed to address the problem of content poisoning in P2P content sharing systems. In general, these reputation models can be grouped into three categories: peer-based models, object based models and hybrid models. In peer-based reputation models, e.g., EigenTrust, PeerTrust and Scrubber, genuine users collectively identify content poisoners by computing a reputation score for each user, and then isolate these poisoners from the system. However, the studies in old research work implied that these peer-based models are insufficient to defend against the poisoning attack.

To address such problems we are going to propose proactive content poisoning system In contrast, our scheme detects unpaid pirates and use discriminatory content poisoning to deter on-line piracy. Legitimate clients can still enjoy the flexibility and convenience provided by open P2P networks. Our scheme stops pirates from illegal download of copyrighted files, even at the presence of many colluding peers. We developed a reputation-based method to detect peer collusion in piracy process.

IV. PROPOSED SYSTEM

The aim of the research work is to design such a protocol that should consist of reduced delivery cost, maximized digital content availability and copyright compliance while exploring P2P network resources. The proposed scheme should also be capable enough to mitigate the collusive piracy within the peer to peer content delivery network. The main objectives of the proposed research work can be briefed as following:

- Error-free identification of colluders and pirates.
- Creation of secure protocol for forwarding poisoned chunk of data to illegitimate client in network.
- Creation of protocol for preventing colluders to come in contact with legitimate clients in the network.

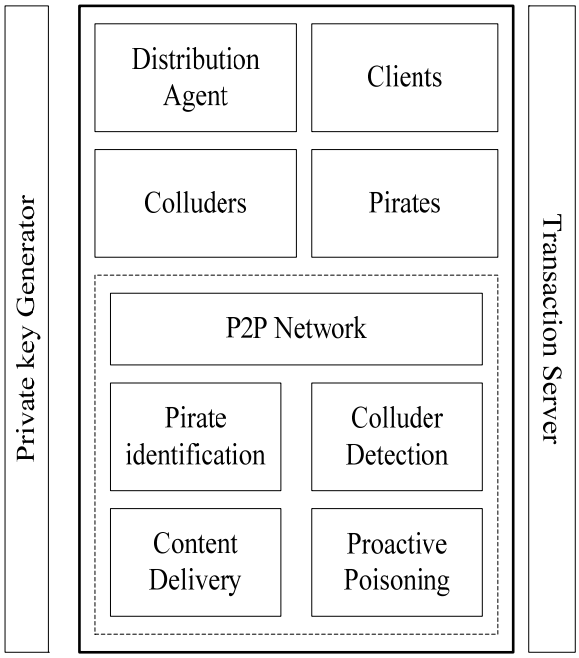

Fig.2 Overall Architecture of the proposed model

Fig 2 shows the overall architecture of the proposed "proactive content poisoning" system where the main components will be Distribution agent, Clients, Colluders, and Pirates. The next phase of architecture will consist of identifying the pirates and colluders, content delivery, and proactive poisoning. The technique used will be to initiate a design of methodology for authenticating peer with IP address and definitely port number. These entities will evaluate the component peers when the download or the upload of the digital content takes place where each of the peers will attempt to identify the illegitimate peers. The proposed approach will seek authentication for the peer looking to download or upload the file and in case the authentication fails for showing existence of illegitimate client in the network, our proposed system model will direct poisoned chunks to the illegitimate peer. The proposed technique can be suitably used in order to identify colluders. The flow of the proposed model is shown in Fig 3.

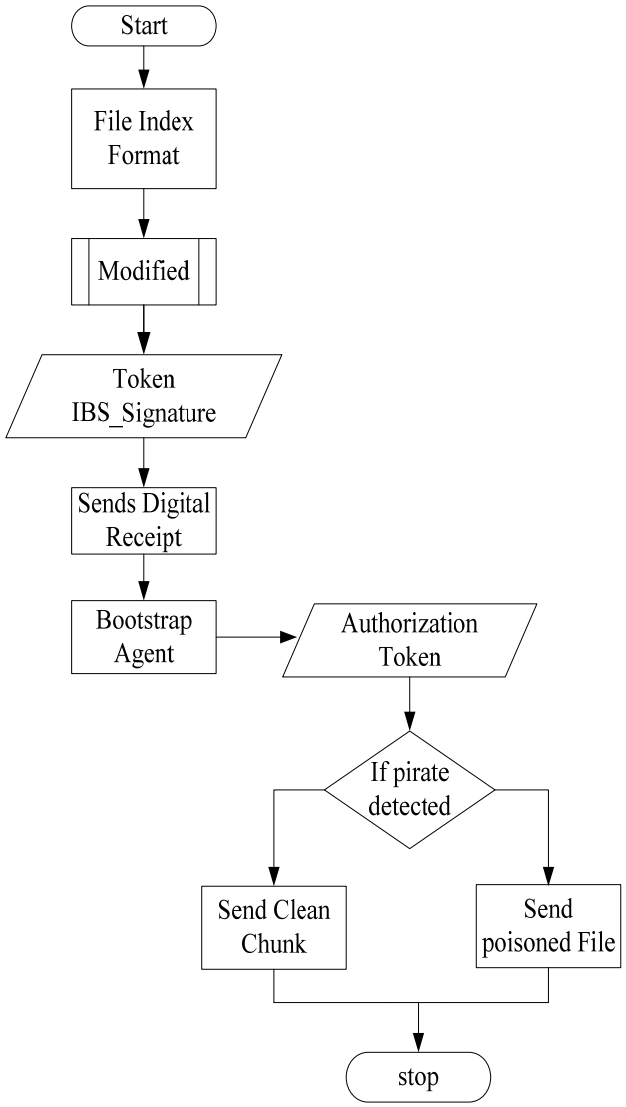

Fig.2 Proposed flow of system model

The above fig.3 shows the pro process flow diagram for our proposed system model of proactive content poisoning. The proposed technique uses the concept of forwarding poisoned chunks to discourage the illegitimate peer existing in the P2P network, where the restricted poison identification potential is exploited and enforce the pirate to reject the clean

chunks downloaded along with the poisoned chunks. Creation of such technique will induce the highly increased download time for the unauthorized client thereby discourage their interest and allow secure transmission to the service provider to legitimate client.

V. SYSTEM MODEL

The proposed system is designed with 4 categories of peer as:

- Clients: This type of the peer are normally the authorized peer within the network
- Colluders: This type of the peers are paid peers who are sharing contents with other legitimate peers
- Distribution agents: These are the set of trusted peers which are controlled by the digital content owners for file distribution.
- Pirates: These are unpaid or unauthorized clients who are always at lookout of some unauthorized downloads of premium contents.

The proposed design also includes a transaction server (see fig 2.) which is assumed to be managing the buying and billing related issues. The transaction server will receive the request from the clients in the network. The communications among the peers are secured by configuring a private key generator in order to generate set of private keys with Identity-Based Signatures (IBS). The private key generator also acts as certificate authority in various secure communications. When the peers will attempt to get themselves connected in a network, the system will deploy the key generator and transaction server. Along with invocation of IBS, the peer communication will not be depended on any public key specially designed as their identity itself will serve as the key of public type.

The system model will initiate with colluders, paid clients, as well as pirate without indexing for their identification. The proposed prototype system model is designed to quantize them involuntarily. The model also consists of a bootstrap unit as an input point of entry which is chosen from one of the distribution agent. The existing network designs involuntarily declare its identification without any type of identification. In order to deal with this, the system model uses port number and IP address instead to any other types of authentication parameters that can be possible used in the same set of the network. A complete connection with the peer is only ensured when it is reachable via any listening port on its host.

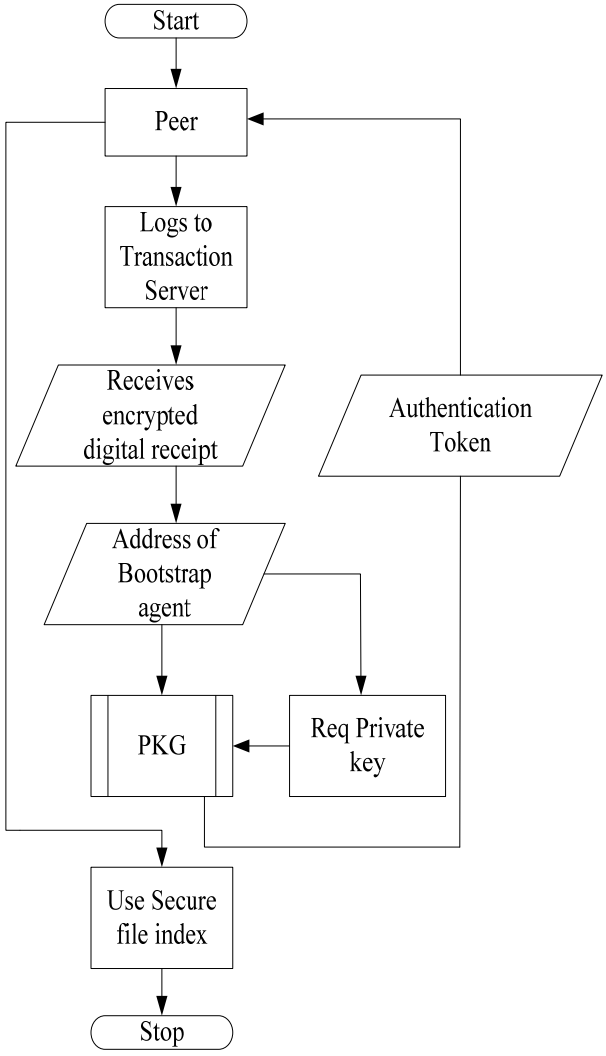

Fig 4. Peer Node Participation Process

The above Fig. 4 shows process flow diagram for peer participating process where it can be understood that each legitimate peer has a valid token (use indication). The token (indication) will be have possession of validity for a very minimized instant of time in order for the peer components to update in constant time duration. The file-index format is modified in order to include a token (indication) and signature, which will be used by the peer nodes for securing the download permission. The system model will use the identification of the listening port as valid identification of the peer. The model also assumes that every peer node posses a better configured listening ports. It was also found that the

majority of the peer to peer clients will linked themselves to the World Wide Web using their home network where the fundamental standard will be to deploy network address translation equipment for sending forward incoming ports. The issue surfaces when a large quantity of the peer is behind a single network address translation are used. The public key to be used in the peer's end will be considered as end point address for which there is absolutely no requirement of encoding the contents of the file thereby minimizing the feasibility of network overhead. The model deploys the bootstrap unit in order to forward the incoming request. The client module will be not disclosed about the identity of the all units apart from bootstrap unit, which prevents the malicious peer node to blacklist or initiate an attack on the distribution agent.

VI. IMPLEMENTATION

The implementation of the proposed research work is carried out in 32 bit windows OS of 1.8 GHz with dual core processor. The programming is carried out in java platform. Although it is a very tough assignment to perform this experiment for real-time peer to peer network for evaluating the copyright infringement. The proposed simulation work is carried out in three phases:

- Estimation of the chunk poisoning rate.
- Estimation of the download time
- Comparative analysis of resistance and overhead in fortification information gather.

The simulation environment of the proposed system is as shown below in fig.5.

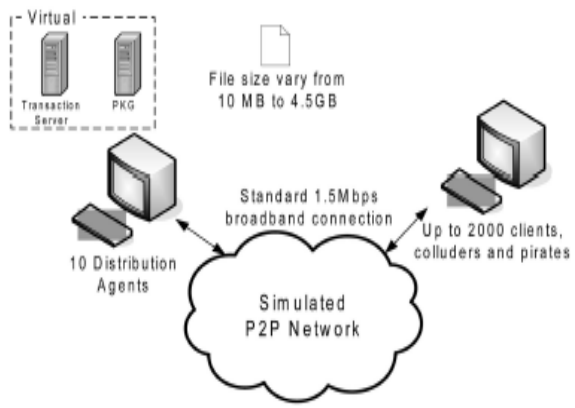

Fig.5 Simulation Scenario considered

The proposed framework will perform in three levels again:

- The top level will work towards emulating the peer to peer transmission.
- The intermediate level will be used for emulating the patterns of the considered peer types (distribution agents, legitimate clients, colluders, illegal pirates.)
- The bottom level will be allocated for data aggregation and information updating.

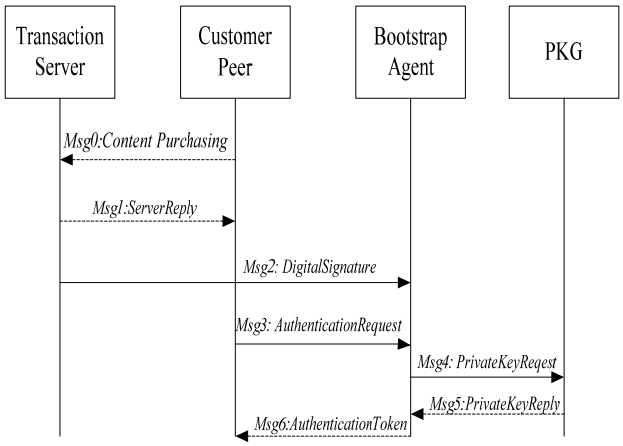

Fig.6. Sequence diagram of the simulation considered

As shown in Fig.6.the simulation is conducted in the same procedure. When a peer node will participate in the network, it has to initially access the transaction server in order to complete the payment process for possessing the digital content. Once the transaction is completed, a digital acceptance slip consisting of the basic information of the content and identification of the clients will be in possession of the client. This digital acceptance slip will be lock in such a procedure that only the legitimate owner of the digital content will be able to unlock it. The address of bootstrap unit is possessed by the authorized client. The newly created digital acceptance slip will be considered for authenticating the newly participating client with bootstrap unit. As the bootstrap unit is configured by the actual owner of the digital content, therefore it unlocks the acceptance slip and authenticate its verification parameters. The transaction server will assign a session key for securing the privacy in the communication channel. Design of a legitimate token takes place when the bootstrap unit request key from the key generator. The existence of the pirates are scrutinized by the peers by evaluating the legitimacy of the supplementary signatures in the index of the files. This scheme of security is deployed by the legitimate peer nodes to distribute clean digital

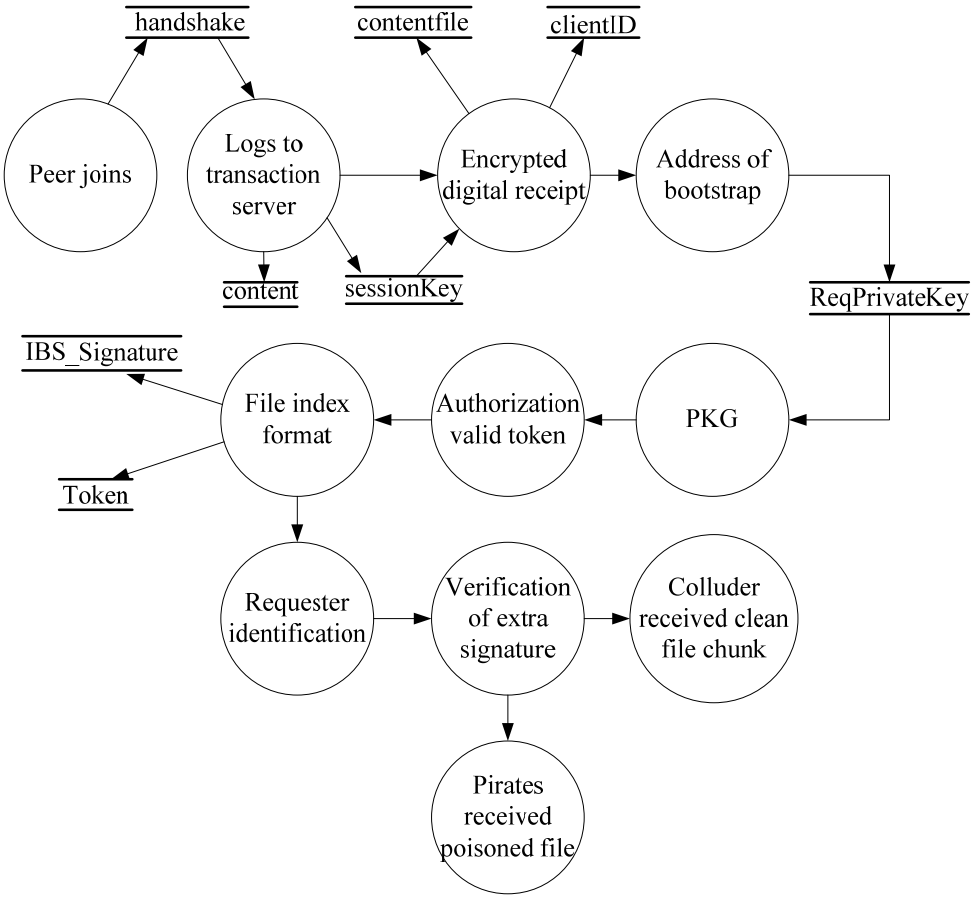

Fig.7.Data Flow Diagram of the proposed model

contents explicitly among the peer nodes and use the forwarding of the digital content poisoning procedure to the unauthorized clients in peer to peer network. The considered tokens which are time-stamped makes sure that identified colluders should not be able to posses newly generated token once the old token perishes.

VII. SIMULATION RESULTS

The simulation is conducted in Java which estimates the shortest feasible download time by a pirate or by any normal clients. Practically, the download time should be highly increased for the case of unauthorized clients present in the network. The simulation is initiated by considering distributing agent to have a possession of clean chunks. The proposed simulation test bed has consideration of various challenging scenarios of piracy for understand the efficiency of the framework designed.

Fig 8 shows the simulation results considering average path length, network diameter, and algorithm for identity based signature and non-indexed system.

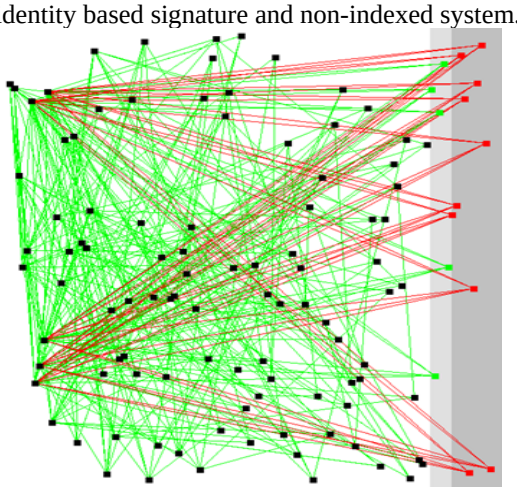

Fig 8. Simulation output showing peers

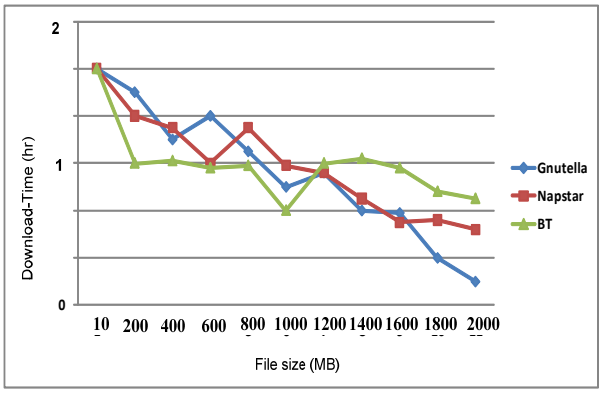

Fig. 9. Simulation result for authorized clients

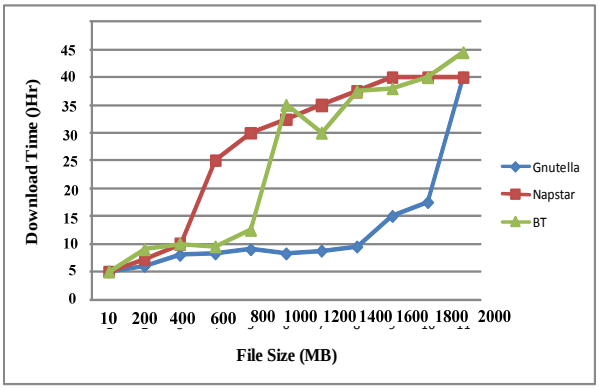

Fig 10. Simulation result for un-authorized clients

Any frequently used P2P protocol like BitTorrent or NapStar could be used for the purpose of simulation. The distribution is evaluated for 2 GB of a multimedia file downloaded from youtube.com. The proposed system is also evaluated with the other different size of the file which have different contents (other than multimedia). Understanding the security implementation of the proposed protocol, it has been seen that when the download request is originated from any client, the proposed framework will first attempt to evaluate the secure confidential identity of the peer nodes and once the authentication is positively accomplished, the cleaner chunk of the digital content requested by the authorized client starts downloading whereas the unauthorized clients will end-up either downloading a poisoned chunks of digital data which will render the unauthorized client to increase the download time to higher extent.

It is clearly observed in fig 9. that when the experiment is performed with 3 prominent p2p protocols e.g. BitTorrent, NapStar , and Gnutella, it can be seen that download time is reduced with every bytes of download of clean chunk of digital contents. While reverse event is witness in case of unauthorized clients as shown in fig 10, where it can be seen that the download time increase with every bits of download of poisoned chunk of data.

VIII. CONCLUSION

The proposed system highlights a novel approach for discriminating authorized and unauthorized peer client in the network. The main intention is to discourage the attempt of any illegitimate client to download the legal and premium content of the digital file on which only the legitimate client have the rights to download. The approach discussed identifies the legal and illegal client and once successful identification is accomplished, the proposed algorithm sends the poisoned chunk of data to the unauthorized client and clean chunk of the digital data only for the genuine and legal peer node. Simulation results claims to highlights the efficiency of proposed algorithm.

Reference

[1] http://en.wikipedia.org/wiki/File_sharing
[2] http://en.wikipedia.org/wiki/Content_delivery_network
[3] http://en.wikipedia.org/wiki/Peer-to-peer
[4] Xin Xiangjun; Xing Peixu;, Efficient Certificate-Based and Randomized Signature from Pairings, Information Engineering, 2009. ICIE '09. WASE International Conference on Issue Date: 10-11 July 2009
[5] Nicolas Christin, Andreas S. Weigend and John Chuang, Content Availability, Pollution and Poisoning in Peer-to-Peer File Sharing Networks, Proceedings of the 6th ACM conference on Electronic commerce, 2005
[6] Ruichuan Chen, Eng Keong Lua, Jon Crowcroft, Wenjia Guo, Liyong Tang, Zhong Chen, Securing Peer-to-Peer Content Sharing Service from Poisoning Attacks, Peer-to-Peer Computing , 2008. P2P '08. Eighth International Conference on Issue Date: 8-11 Sept. 2008
[7] STEPHANOS ANDROUTSELLIS-THEOTOKIS AND DIOMIDIS SPINELLIS, A Survey of Peer-to-Peer Content Distribution Technologies, ACM Computing Surveys, Vol. 36, No. 4, December 2004, pp. 335–371.

[8] Dan Boneh, Matthew Franklin, Identity-Based Encryption from the Weil Pairing, SIAM J. of Computing, Vol. 32, No. 3, pp. 586-615, 2003.
[9] Lei Guo, Songqing Chen, Shansi Ren, Xin Chen, and Song Jiang, PROP: a Scalable and Reliable P2P Assisted Proxy Streaming System, Distributed Computing Systems, 2004. Proceedings. 24th International Conference on 2004
[10] Ernesto Damiani, Sabrina De Capitani di Vimercati, Stefano Paraboschi, A ReputationBased Approach for Choosing Reliable Resources in PeertoPeer Networks, Proceedings of the 9th ACM conference on Computer and communications security ACM New York, NY, USA ©2002
[11] D. Dumitriu,_ E. Knightly,_ A. Kuzmanovic, _ I. Stoica, and W. Zwaenepoel_,, DenialofService Resilience in PeertoPeer File Sharing Systems, Proceeding SIGMETRICS '05 Proceedings of the 2005 ACM SIGMETRICS international conference on Measurement and modeling of computer systems ACM New York, NY, USA ©2005
[12] Ton Kalker1, Dick Epema2, Pieter Hartel3, Inald Lagendijk4, Maarten van Steen5, Music2Share – Copyright-Compliant Music Sharing in P2P Systems, Proceedings of the IEEE Issue Date: June 2004
[13] Balachander Krishnamurthy, Craig Wills, Yin Zhang, On the Use and Performance of Content Distribution Networks, ACM SIGCOMM INTERNET MEASUREMENT WORKSHOP 2001
[14] Stefan Saroiu, Krishna P. Gummadi, Richard J. Dunn, Steven D. Gribble, and Henry M. Levy, An Analysis of Internet Content Delivery Systems, ACM SIGCOMM INTERNET MEASUREMENT WORKSHOP 2001
[15] Pablo Rodriguez, SeeMong Tan, Christos Gkantsidis, On the feasibility of Commercial, Legal P2P Content Distribution, ACM SIGCOMM Computer Communication Review Homepage archive Volume 36 Issue 1, January 2006
[16] Matthew Yurkewych Brian N. Levine Arnold L. Rosenberg, On the Cost Ineffectiveness of Redundancy in Commercial P2P Computing, Proceeding CCS '05 Proceedings of the 12th ACM conference on Computer and communications security ACM New York, NY, USA ©2005
[17] Kevin Walsh, Emin G¨un Sirer, Fighting PeertoPeer SPAM and Decoys with Object Reputation, Proceeding P2PECON '05 Proceedings of the 2005 ACM SIGCOMM workshop on Economics of peer-to-peer systems ACM New York, NY, USA ©2005
[18] Runfang Zhou, Kai Hwang, GossipTrust for Fast Reputation Aggregation in Peer-to-Peer Networks, IEEE TRANSACTIONS ON KNOWLEDGEMENT AND DATA ENGINEERING (TKDE-0003-0107R1, FINALIZED FEB. 11, 2008)