Security, Privacy and Anonymity in Legal Distribution of Copyrighted Multimedia Content over Peer-to-Peer Networks: A Brief Overview

Amna Qureshi, Helena Rifà-Pous, David Megías
Estudis d'Informàtica, Multimèdia i Telecomunicació, Internet Interdisciplinary Institute (IN3),
Universitat Oberta de Catalunya (UOC),
Rambla del Poblenou, 156, 08018, Barcelona, Catalonia, Spain
{aqureshi,hrifa,dmegias}@uoc.edu

***Abstract*—Recent decades have witnessed a swift development in network structures like Peer-to-Peer (P2P) systems, offering advantages such as significantly reducing the setup running cost for the content providers and providing end users data access within very short time. However, today's P2P distribution systems are severely abused by illegal re-distributions. The enforcement of copyright protection mechanisms in these systems poses serious privacy threats to end users. Various researchers have examined the challenges characterizing these systems from diverse viewpoints, proposing strategic solutions. This paper, to the best of our knowledge, for the first time conducts a review analysis focused on security, privacy and anonymity in P2P systems, describing the challenges and solutions that are associated with legal content distribution in them.**

Keywords: copyright protection; privacy; anonymity; security; peer-to-peer networks

I. INTRODUCTION

Peer-to-Peer (P2P) is often described as a type of decentralized computing system in which nodes, referred to as peers, use the Internet to communicate with each other directly. These systems are attractive because they do not require any special administrative arrangements unlike centralized facilities, and their decentralized nature makes them scalable, bandwidth efficient and fault tolerant.

P2P applications account for approximately 60% of Internet's Traffic. However, a major part of this traffic is generated by P2P content distribution systems. In particular, today's P2P content distribution applications (eDonkey [1], BitTorrent [2]) are extremely popular among millions of users. These applications allow users to contribute, search and obtain a digital content, ranging from relatively small-sized pictures or music files to large-sized contents like complete software packages, movies or similar types of multimedia content, in a distributed manner. Content distribution in P2P has also received considerable attention in the research community [3].

The P2P technology for content distribution systems is beneficial to both content providers and end users. From media companies and e-commerce vendors point-of-view, this technology enables them to make valuable content available to a large amount of people at minimal cost and better performance. For example, Red Hat Inc. uses Bit-Torrent to distribute Red Hat Linux. Similarly, Pando Networks [4], a managed P2P content distribution application, enables content owners to publish, distribute and track their multimedia contents at reduced delivery costs. Also, from end users perspective, multimedia files can easily be accessed and downloaded within a short time.

Despite the potential of P2P content distribution technology to revolutionize the Internet in numerous respects, it has often been surrounded with the copyright controversy. On the one hand, these systems offer content providers an opportunity to achieve global exposure with low distribution costs. On the other hand, these providers argue that users of these systems get copyrighted work for free. They fear losing control of content ownership and worry about promotion of illegal activity. P2P content distribution itself is not illegal; however the act of sharing copyrighted contents without permission is illegal. But this illegal activity is not only onerous to content providers but also to the end users. Users may receive or re-distribute files that may make them accountable to civil or criminal liability under copyright infringement laws.

In addition to security, another concern among users is whether the presence of copyright protection mechanism in P2P distribution systems can violate their privacy interests. The seriousness of the effects that protecting copyright has on the privacy interests of users is significant [5]. Moreover, within P2P distribution systems, a collection of identifiable personal data should be limited to the minimum since anonymity is one of the basic functions of privacy.

There is an inherent conflict of interest between copyright protection supporters and privacy advocates. Currently, this is a hot research area among researchers who are focusing on preservation of content providers ownership properties and content receivers' privacy and anonymity. Cryptographic and fingerprinting mechanisms [6-8] are utilized to distribute contents within P2P systems such that the provider neither knows the receiver's copy nor his identity. Only on finding an illegally re-distributed copy, the provider identifies the re-distributor. Intuitively, there is large scale of interdependence amongst security, privacy and anonymity aspects within legal content distribution in P2P systems, thus favoring the combined approach taken in this study. This paper surveys current P2P distribution systems in terms of technological aspects of security, privacy and anonymity, aiming to provide a comprehensive account of the implemented mechanisms and their important features.

The remainder of this paper is a survey of research in these areas. In Section II, we discuss the challenges being faced by P2P developers in building secure and anonymous P2P systems. Section III briefly gives an overview of implemented mechanisms and in Section IV, we discuss the conclusions and open problems of implemented systems.

II. SECURITY, PRIVACY AND ANONYMITY CHALLENGES IN P2P CONTENT DISTRIBUTION SYSTEMS

The concepts of security, privacy and anonymity used in this paper are defined in the context of providing a legal content distribution in a P2P system, and are described as follows:

1) Security: A mechanism aimed to protect an intellectual property and provide trustworthiness.

2) Privacy: The protection of user-related information in such a way that no personal information of an end user is revealed, unless a user is found to be guilty of illegal re-distribution. It is also called a conditional privacy.

3) Anonymity: A method to protect the identity of provider and receiever and also to protect the contents of transferred data between them.

A. Security Issues in P2P

The decentralized nature of P2P technology makes copyright holders more resistive to its adoption due to absence of a central authority, which could regulate how and what kind of files get distributed within the system. Xiaosong and Kai [9] studies highlight that the security state for P2P systems is worse because of the absence of a centralized authority that can vouch for security parameters. The diverse nature of the multimedia material presents a severe challenge to the establishment of effective strategies that would foster secure systems [10]. P2P networking renders multimedia distribution channels vulnerable to various forms of attacks, e.g., copyright infringement, the possibilities of downloading files infected with malicious codes, denial of service and susceptibility to attacks [11].

Recent years have witnessed a large number of cases related to copyright infringement, due to P2P file exchange, in many countries. Apart from being a source for pirated content, P2P content distribution systems share files that pose severe security risks to end users. With millions of connected users and even more available files, there is no way to verify the legitimacy and safety of shared files. The downloadable files could contain malicious codes that can attack users′ computer with worms, malware, spyware, viruses and more. For example, a severe virus known as Antinny, affected the Japanese-based P2P content distribution system "Winny". This virus led to the disclosure of a large amount of U.S. military base security codes along with private documents of a police investigator [12].

Moreover, tracing a copyright violator is an immense task which requires providers to work in conjunction with watermarking and fingerprinting technology provider firms as well as P2P content distribution developers. These controversies have been analyzed in [13] concluding that, in order to design a secure content distribution system, P2P developers need to identify the gaps in security.

B. Privacy Issues in P2P

Privacy is largely an access control issue. Legal approaches are insufficient, because they can easily be circumvented. There is a need to design systems in such a way that no personal information of end users is revealed, unless a user violates the terms of service. Only in that case, his or her personal details and how he or she violated the terms would be revealed to the relevant authorities. Moreover, only the identity of a copyright violator should be revealed without affecting other innocent users of the system. Many existing P2P content distribution systems with copyright protection mechanisms monitor activities of the users. These systems are more concerned about security of data rather than protection of user privacy. Few P2P systems have been proposed that address both security concerns of providers and users' privacy concerns [7, 8, 14].

Cryptography, Digital Rights Management (DRM), and watermarking are digital content protection techniques against piracy. The hidden information, watermark, is used for copyright protection, covert communication, source tracking of

leakage, and so forth. Two significant issues in these schemes need to be considered, i.e., the provider's security as well as the privacy of an end user. The demand to protect privacy information of end users is increasing along with the rise of privacy leakage.

C. Anonymity Issues in P2P

Anonymity is an important factor to account for in P2P content distribution systems. The major anonymity issue within these systems is that the users' identities and actions can be revealed by other users. To provide anonymity to users in these systems, the personal and sensitive information, such as user identification and IP address of the user, must be hidden during communication with others.

From the P2P distribution system's perspective, there exist three types of anonymity: Receiver, provider anonymity and mutual anonymity. In receiver and provider anonymity, identity of the receiver and provider is hidden, respectively. Mutual anonymity hides the identities of both receiver and provider from each other and also hides the comunication between these two entities.

Grodzinsky and Tavani [15] emphasize the absence of anonymity in P2P systems by noting that the user reveals his or her details, such as plain-text queries and IP addresses, to provider of services when downloading files. Furthermore, Wang et al. [16] highlights that a great deal of information regarding the user preferences can be collected in content distribution by tracking the user activities at the provider side, thus compromising the user's anonymity.

The conflict between privacy and security within P2P distribution systems vouch for a debate between anonymity and accountability, i.e. decreased anonymity (less user privacy) is proportional to increased accountability (more security to provider). However, both accountability and anonymity properties must exist side-by-side within these systems.

Similarly, the open nature of P2P systems makes data privacy a major challenge. Since there is no central authority in these systems that can authenticate and protect against malicious end users, it is up to the user to protect itself and be responsible for its own actions. Consequently, each user needs to evaluate the information received from another user in order to determine the trustworthiness about the information as well as the provider. This can be achieved by employing trust management techniques [17] within these systems. However, most trust models in P2P systems are identity-based; hence there exists a trade-off between trust and anonymity.

III. MECHANISMS TO COUNTER CHALLENGES IN P2P CONTENT DISTRIBUTION SYSTEMS

This section discusses implemented mechanisms that strive to overcome security, privacy and anonymity issues in P2P content distribution systems.

A. Mechanisms for Security

Various researchers have devised techniques and mechanisms to counter security attacks.

***1) Encryption*:** Encryption is a core technique for content protection. Before distribution, the content is encrypted by the provider and the decryption key is provided only to those users who have permission to access the legitimate copies of the content. Despite the fact that such techniques can protect the contents during their transmission, they cannot prevent a user from re-distributing the data illegally once they have been received and decrypted. However, this problem can be solved by establishing a DRM capability.

2) DRM: DRM is a technology that prohibits illegal copy and distribution of contents and permits only authorized users to access them. These systems were originally intended to help content providers in secure distribution of digital media to a large number of users in a manner that protects the interests of the owner only. Literature review shows that there are few DRM solutions available in P2P networks. Kalker et al. [18] describe a solution based on DRM for the problem of copyright infringement in P2P networks for music sharing. Li et al. have proposed a DRM enabled P2P architecture which provides secure distribution of copyright-protected music contents and efficient tracing of unauthorized users [9]. In [19], Chen et al. proposed a DRM mechanism for a Bit-torrent like P2P system which provides end-to-end content secrecy and access control.

3) Digital watermarking: Watermarking is recognized as a promising technique developed to address the problems of copyright protection, content authentication, etc. There are two forms of watermarking, copyright watermarking and fingerprint watermarking (fingerprinting). In copyright watermarking, a copyright message (watermark) is embedded into the content which indicates copyright holder's identification. However, it is only used to declare the copyright and cannot be used to trace the copyright violator. Tsolis et al. [20] have proposed a P2P distribution system which not only allows digital content

exchange but also supports copyright protection and management through watermarking technologies. In second type of watermarking i.e. digital fingerprinting [21], a user specific identification mark is embedded into content so that it can be used to track an illegal re-distributor. Once an unauthorized content copy is found, the embedded fingerprint can uniquely identify the person responsible for this violation.

***4) Trust Management*:** Little efforts have been done to prevent malware spread within P2P distribution systems. To prevent this spread within these systems, trust mechanisms can be used to find the reputed users to exchange with, or to avoid malicious nodes. In [17], Ding et al. have proposed a model based on trust management scheme that can mitigate the malware proliferation in P2P distribution systems. Zhou et al. in [22] investigated the worms taking advantage of P2P weakness and proposed several countermeasures based on individual users to carry out detection and post-detection mitigation.

B. Mechanisms for Privacy

Privacy in P2P distribution systems with copyright protection can be granted using techniques such as anonymous fingerprinting and DRM with privacy.

1) Anonymous fingerprinting: In the aforementioned fingerprinting type, the goal is to protect content providers against illegal re-distributors. However, to protect user's privacy from content providers, another type of fingerprinting, i.e., anonymous fingerprinting [23] is used. This technology preserves users' rights for privacy by providing anonymity and the unlinkability of their P2P activity. However, there is significantly lesser research work done on privacy preserving mechanisms in P2P systems with legal content distribution.

Megías and Domingo-Ferrer [7] have proposed a novel concept of automatic fingerprint recombination designed for P2P content distribution systems. The proposed scheme utilizes the fingerprinting concept to provide identification to the copyright owner and detect illegal content re-distributors. Furthermore, the users can preserve their privacy as long as they do not get involved in illegal re-distribution. In [8], Domingo-Ferrer and Megías have proposed a P2P protocol for distributed multicast of fingerprinted content. In the proposed framework, cryptographic primitives and a robust watermarking technique have been utilized to produce different marked copies of the content for the requesting user such that it can help the provider to trace re-distributors without affecting the privacy of honest users.

2) DRM with privacy: In literature, only a few DRM mechanims in P2P can be found, which address both content provider security and end user privacy. In [24], Sun et al. proposed an identity-based DRM system with privacy enhancement. Their DRM system retains user privacy by hiding the relationship information between users and the contents the users own.

C. Mechanisms for Anonymity

Various P2P anonymous mechanisms have been proposed and implemented, aiming to provide protection to end users within the system.

1) Anonymous communication: Onion routing [25] is a distributed P2P mechanism that allows two users to communicate anonymously over the network. It protects its communication against traffic analysis. The main aim of onion routing is to prevent intermediary nodes from knowing the source, destination and contents of the message. Onion Routing has been adapted in several anonymous P2P systems such as Anonymous P2P File Sharing (APFS) and Tor [26], to provide anonymous communication between users. Yu et al. [27] have proposed a P2P protocol, Nemor, which not only allows a requesting user and a provider to communicate anonymously with each other and from other participating users, but also protects the identity of the content being exchanged. Another P2P protocol, Peer-to-Peer Personal Privacy Protocol (P5) [28], uses a hierarchical broadcasting technique to achieve mutual anonymity between users. For different levels of the hierarchy, different levels of anonymity are provided.

2) Anonymous authentication: It is impractical to pursue user anonymity without taking accountability into consideration. Accountability has traditionally been achieved through authentication mechanisms. In order to preserve anonymity within these accountable systems, trust mechanisms are utilized, e.g. the solution proposed by Lu et al. [29] uses an anonymous zero-knowledge authentication protocol to support trust management such that users can use unforgeable and verifiable pseudonyms instead of their real identities. Similarly, Wang et al. in [30] have proposed an anonymous collaboration signature authentication protocol in which each user, instead of using his or her real identity, owns an un-forgeable and verifiable identity signature and this identity signature is signed by a trusted party through a collaboration signature method.

Table I gives a comparison of the presented P2P systems with respect to security, privacy and anonymity properties.

TABLE I. COMPARISON OF THE PRESENTED P2P SYESTEMS

P2P Systems	Guaranteed Properties		
	Security	***Privacy***	***Anonymity***
Ref. [9]	DRM Encryption	No	No
Ref. [20]	Watermark-ing	No	No
Ref. [17]	Trust Mechanism	No	No
Ref. [7]	Encryption Fingerprint-ing	Fingerprint-ing	Anonymous communi-cation
Ref. [8]	Encryption Fingerprint-ing	fingerprint-ing	Anonymous Communi-cation
Ref. [24]	Encryption	DRM	No
Ref. [28]	No	No	Anonymous communi-cation
Ref. [29]	Encryption Trust Mechanism	No	Anonymous authentication

IV. CONCLUSIONS AND OPEN PROBLEMS

The field of P2P technology presents a number of interesting challenges which include new methods for providing security, privacy and anonymity. Considerable amount of research work has been carried out by researchers to provide an appropriate balance between distributing content on a large-scale and preserving the right of copyright owners.

Much of the work has been done by using applications of watermarking, fingerprinting and DRM mechanisms. However, most of the research work involving fingerprinting protocols for copyright protection incurs high computational and communicational burdens due to the use of public-key encryption of the contents, secure multiparty protocols amd other techniques. Also, research work for developing robust and secure watermarking scheme is still in progress. The tradeoffs between robustness, capacity and imperceptibility of watermarking schemes are yet to be achieved. Also, there is a need to develop an efficient traitor tracing scheme, such that the anonymity of honest users is not affected.

Similarly, the proposed works on DRM mechanisms in P2P systems have not been able to effectively prevent copyright infringement. Moreover, in order to develop an effective DRM mechanism to prevent large scale copyright infringement, the developers may often undermine users' fair rights.

It is worth noting that in achieving anonymity in P2P systems, there is a performance overhead. This overhead is due to encryptions and decryptions, insertion of fake traffic and increasing the routing path to provide anonymity between two communicating users. Therefore, better anonymity and efficiency tradeoffs are of primary importance for these systems to be deployed and gain user acceptance. Another challenge faced by P2P systems is the harmonization between anonymity and accountability. There is a need to devise security mechanisms to ensure anonymity for honest users and traceability for misbehaving users. Similarly, another problem is that the users of these systems require anonymity but the government, media and software industry want some monitoring tools to be incorporated within the software in order to track illegal re-distributors. Thus, there is a need to balance the anonymity of users and security of copyright holders.

This brief review illustrates that P2P systems face serious challenges in terms of combining security, privacy and anonymity properties for legal content distribution. Efforts in addressing these concerns are still unsuccessful because of the intricacy of each other. Although these systems are deployed and used on a large scale, there remain many open issues that need to be addressed by researchers and P2P developers.

ACKNOWLEDGMENT

This work was partly funded by the Spanish Government through projects TSI2007-65406-C03-03 ''E-AEGIS'', TIN2011-27076-C03-02 ''CO-PRIVACY'' and CONSOLIDER INGENIO 2010 CSD2007-0004 ''ARES''.

REFERENCES

[1] eDonkey2000 [Online]. Available: http://*www.emule-project.net/*

[2] BitTorrent [Online]. Available: http://www.bittorrent.com/

[3] S. A.Theotokis, and D. Spinellis, "A survey of Peer-to-Peer content distribution technologies," ACM Comput. Surv, 2004, vol. 36, no. 4, pp. 335-371.

[4] Pando Networks [Online]. Available:http://www.pandonetworks.com/

[5] K. W. Hamlen, and B. Thuraisingham, "Secure Peer-to-Peer networks for trusted collaboration," Proc. Int. Conf. Collaborative Computing: Networking, Applications and Worksharing, Nov. 2007, pp. 58-63.

[6] D. Megías, and J. Domingo-Ferrer, "Privacy-aware Peer-to-Peer content distribution using automatically recombined fingerprints," Multimedia Systems, (In press)

[7] J. Domingo-Ferrer, and D. Megías, "Distributed multicast of fingerprinted content based on a rational Peer-to-Peer community," Comput. Commun., 2013, vol. 36, no. 5, pp. 542-550.

[8] J. S. Li, C-J. Hsieh, and C-F. Hung, "A novel DRM framework for Peer-to-Peer music content delivery," J. Syst. Softw, 2010, vol. 83, no. 10, pp. 1689-1700.

[9] L. Xiaosong, and H. Kai, "Collusive piracy prevention in Peer-to-Peer content delivery networks," IEEE Trans. Comput., 2009, vol. 58, no. 7, pp. 970-983.

[10] X. Fan, M. Li, J. Ma, Y. Ren, H. Zhao, and Z. Su, "Behavior-based reputation management in Peer-to-Peer file-sharing networks," J. of Comput. Syst. Sci., 2012, vol. 78, no. 6, pp. 1737-1750.

[11] B. Lipinski, and P. Macalpine, "A security review of anonymous Peer-to-Peer file transfer protocol," Rice University, Houston, Texas, Tech.Report.

[12] M. Ingram (April 2006), "66,000 Names and Personal details leak on Peer-to-Peer," [Online]. Available: http://www.slyck.com/news.php?story=1169.

[13] P. Deewan, and P. Daasgupta, "P2P reputation management using distributed identities and decentralized recommendation chains," IEEE Trans. Knowl. Data Eng., 2010, vol. 22, no. 7, pp. 1000-1013.

[14] C-T. Yen, H.T. Liaw, and N.W. Lo, "Digital Rights Management system with user privacy, usage transparency, and superdistribution support," Int. J. Commun. Syst., (In press)

[15] F. S. Grodzinsky, and H. T. Tavani, "Peer-to-Peer networks and the verizon v. RIAA case: Implications for personal privacy and intellectual property," Ethics and Inf. Technol., Oct. 2005, vol. 7, no. 4, pp. 243-250.

[16] Y. Wang, A. Nakao, A. V. Vasilakos, and J. Ma, "Peer-to-Peer soft security: On evolutionary dynamics of Peer-to-Peer incentive mechanism," Comput.Commun., 2011, vol. 34, no. 3, pp. 241-249.

[17] X. Ding, W. Yu, and Y. Pan, "A dynamic trust management scheme to mitigate malware proliferation in Peer-to-Peer Networks," Proc. IEEE Int. Conf. on Commun.., 2008, pp. 19-23.

[18] T. Kalker, D.H.J. Epema, P.H. Hartel, R.L. Lagendijk, and M. Van. Steen, "Music2Share-Copyright–compliant music sharing in Peer-to-Peer systems," Proc. IEEE special issue Digital Right Management, 2004, vol. 92, no. 6, pp. 961-970.

[19] Y. Y. Chen, J. K. Jan, Y. Y. Chi, and M. L. Tsai, "A feasible DRM mechanism for BT-like Peer-to-Peer system," Proc. Int. Symp. Inform. Eng. Electron. Commerce, May 2009, pp. 323–327.

[20] D. K. Tsolis, S. Sioutas, A. Panaretos, I. Karydis, and K. Oikonomou, "Decentralized digital content exchange and copyright protection," Proc. IEEE Symp. Comput. Commun., July 2011, pp. 1056-1061.

[21] D. Boneh, and J. Shaw, "Collusion-secure fingerprinting for digital data," IEEE Trans. Inf. Theory, 1998, vol. 44, no. 5, pp. 1897-1905.

[22] L. D. Zhou, L. T. Zhang, F. Mcsherry, N. Immorlica, M. Costa, and S. Chien, "A first look at Peer-to-Peer worms: threats and defenses," Proc. Int. Workshop on Peer-To-Peer Systems, Feb. 2005, pp. 24-35.

[23] C.-C. Chang, H.-C. Tsai, and Y.-P. Hsieh, "An efficient and fair Buyer-Seller fingerprinting scheme for large scale networks," Comput. & Security, 2010, vol. 29, no. 2, pp. 269-277.

[24] M-K. Sun, C. S. Laih, H. Y. Yen, and J. R. Kuo, "A ticket based Digital Rights Management model," Proc. IEEE Conf. Consumer Commun. Networking , 2009, pp. 1-5.

[25] P. Syverson, G. Tsudik, M. Reed, and C. Landwehr, "Towards an analysis of Onion Routing security," Int. Workshop on Designing Privacy Enhancing Technologies, July 2000, pp. 96–114.

[26] R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The second- generation Onion Router," Proc. 13th USENIX Security Symp., 2004, vol. 13, pp.303- 320.

[27] F. Yu, V. Gopalakrishan, D. Lee, and K. K. Ramakrishnan, "Nemor: A congestion-aware protocol for anonymous Peer-based content distribution," Proc. IEEE Conf. P2P computing, Sept. 2011, pp. 260-269.

[28] R. Sherwood, B. Bhattacharjee, and A. Srinivasan, "P5: A Protocol for scalable anonymous communication," Proc. IEEE Symp. Security and Privacy, 2002, pp.58-70.

[29] L. Lu, J. Han, Y. Liu, L. Hu, J. Huai, L. M. Ni, and J. Ma, "Pseudo Trust: Zero-knowledge authentication in anonymous Peer-to-Peer," IEEE Trans. Parallel and Distrib. Syst., 2007, vol. 19, no. 10, pp. 1325- 1337.

[30] X. Wang, X. Sun, G. Sun, and D. Luo, "CST: Peer-to-Peer anonymous authentication system based on collaboration signature," Proc. IEEE Int.Conf. Future Inform. Technol., May 2010, pp. 1-7.