Expert Systems with Applications 42 (2015) 1391–1408

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Framework for preserving security and privacy in peer-to-peer content distribution systems

Amna Qureshi *, David Megías, Helena Rifà-Pous

Estudis d'Informàtica Multimèdia i Telecomunicació, Internet Interdisciplinary Institute (IN3), Universitat Oberta de Catalunya (UOC), Rambla del Poblenou, 156, 08018 Barcelona, Spain

ARTICLE INFO

Article history:
Available online 16 September 2014

Keywords:
Privacy
Security
Anonymity
Collusion resistant fingerprinting
Peer-to-peer networks

ABSTRACT

The use of peer-to-peer (P2P) networks for multimedia distribution has spread out globally in recent years. The mass popularity is primarily driven by efficient distribution of content, also giving rise to piracy. An end user (buyer) of a P2P content distribution system does not want to reveal his/her identity during a transaction with a content owner (merchant), whereas the merchant does not want the buyer to further distribute the content illegally. Therefore, there is a strong need for a content distribution mechanism over P2P networks that do not pose security and privacy threats to the copyright holders and end users, respectively. The existent systems for copyright and privacy protection employ cryptographic mechanisms at a cost of high computational burden which makes these systems impractical to distribute large sized files, such as music albums or movies. In this paper, we propose and analyze a P2P content distribution system which allows efficient distribution of large-sized content while preserving the security and privacy of merchants and buyers, respectively. Our proposed framework is able to resolve the problems of piracy tracing, buyer frameproofness, collusion resistance, dispute resolution and buyer's anonymity. We have carried out simulations to evaluate the performance of our framework in terms of imperceptibility, robustness, throughput and content delivery costs. The experimental results confirm that the proposed framework provides an efficient solution to copyright infringement issues over P2P networks, reducing the multimedia file sizes as much as five times on average, while protecting the end users' privacy and anonymity.

1. Introduction

P2P systems are attractive because they do not require any special administrative arrangements, unlike centralized facilities, and their decentralized and distributed nature make them scalable, bandwidth efficient and fault-tolerant. P2P applications account for approximately 60% of Internet's traffic (García-Dorado, Finamore, Mellia, Meo, & Munafó, 2012). In particular, P2P content distribution applications (eDonkey2000, 2000; gtk-Gnutella, 2000) are extremely popular among millions of users. These applications allow users to contribute, search and obtain a digital content in a distributed manner. Content distribution in P2P has also received considerable attention in the research community (Passarella, 2012; Theotokis & Spinellis, 2004). The P2P technology for content distribution systems is beneficial to both content providers and end users. From the media companies point-of view, the P2P technology enables them to make valuable content available to a large number of people at minimal cost and better performance. These benefits are the attractive features for media companies towards the adoption of P2P systems, e.g. BitTorrent (BitTorrent, 2001) is one of the most popular P2P distribution systems used on the Internet and it accounts for a significant amount of traffic on the Internet. Similarly, Internap (Internap, 1996), a managed P2P content distribution application, enables content owners and media companies to publish, distribute and track their games, video and software at reduced delivery costs. Besides, from the end users perspective, audio, video and software 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. The copyright holders argue that they provide copyright content to the end users of the systems and that these end users are involved in illegal re-distributions. They apparently fear losing control of content ownership and worry about the illegal activity promotion. Moreover, tracing (Chor, Fiat, Naor, & Pinkas, 2000) a copyright

* Corresponding author. Tel.: +34 934505405.
E-mail addresses: aqureshi@uoc.edu (A. Qureshi), dmegias@uoc.edu (D. Megías), hrifa@uoc.edu (H. Rifà-Pous).

http://dx.doi.org/10.1016/j.eswa.2014.08.053
0957-4174/© 2014 Elsevier Ltd. All rights reserved.

violator is an immense task which requires content providers to work in conjunction with watermarking (Bianchi & Piva, 2013; Cox, Miller, Bloom, Fridrich, & Kalker, 2007; Hartung & Kutter, 1999) and fingerprinting (Barg, Blakley, & Kabatiansky, 2003; Voloshynovskiy, Farhadzadeh, Koval, & Holotyak, 2012) providers as well as P2P content distribution service developers. However, this illegal re-distribution (Von-Lohmann, 2003) act is not only onerous to content providers but also to the end users. The major concern among end users is whether the presence of copyright protection mechanisms (Lian, Kanellopoulos, & Ruffo, 2009) in P2P distribution systems can violate their privacy interests. The fact that a tracing mechanism makes use of a record which details what multimedia files are shared through a specific IP address, or a list of the peers with whom a user has interacted, disrespects the privacy of the user. Therefore, there is an inherent conflict of interest between copyright protection supporters and privacy advocates and thus there is a need to balance security and privacy needs when developing P2P content distribution systems. Similarly, the conflict between privacy and security within P2P content distribution system manifests itself in a debate between anonymity and accountability, i.e. decreased anonymity (less user privacy) is proportional to increased accountability (more security to provider). Currently, security and privacy in P2P systems is a hot research area among researchers who are focusing on the preservation of content providers ownership properties, content receivers' privacy and accountability. However, most of the existing P2P content distribution systems provide security and privacy at a cost of high computational burden at the merchant's and/or at the user's end.

In this paper, we propose a P2P content distribution system that provides copyright protection to the merchant at a reduced computational cost and also offers revocable privacy to an end user. In the proposed system, the multimedia file is partitioned by the merchant into a base and a supplementary file. The base file is much smaller than the original file and contains the most important information. Without this information, the supplementary file is unusable. The base file is dispensed by the merchant on payment from the user and a supplementary file is sent to the P2P network to be distributed in P2P fashion. Thus, this scheme enables the merchant to save bandwidth and CPU time. The asymmetric fingerprinting protocol is performed by the merchant and the buyer in the presence of a trusted party in such a way that the merchant does not know the fingerprint and the fingerprinted content, while the buyer receives fingerprinted content with his/her unique identity. Collusion-resistant fingerprinting codes are embedded by the merchant into the content so as to identify an illegal re-distributor(s) from a pirated content. The proposed framework also enables buyers to obtain digital contents anonymously, but this anonymity can be revoked as soon as he/she is found guilty for copyright violation. The buyers are provided anonymity by using dynamic pseudonyms instead of their real IDs. To ensure anonymous communication between buyers, onion-routing is used for an anonymous data transfer. Moreover, to provide accountability, a key agreement protocol has been adopted in our scheme. The simulation results show that the proposed framework yields an effective reduction in the computational overheads for a merchant. Also, the security analysis proves that the proposed system exhibits security and conditional anonymity to the merchant and the buyer, respectively.

The paper is organized as follows. Section 2 reviews the related work on P2P networks, multimedia content and privacy protection schemes and P2P content distribution systems. Section 3 provides the building blocks of the proposed framework. Section 4 discusses the design of the framework. Section 5 presents the results of the experiments designed to evaluate the performance of the framework. Also the security analysis of the proposed framework is discussed in this section. Finally, Section 6 summarizes the conclusions and future research issues.

2. Related Work

This section reviews related work on P2P networks, multimedia content protection schemes, privacy protection mechanisms, and P2P content distribution systems.

2.1. P2P architectures

Peer-to-peer networks may be categorized into the three categories: centralized P2P networks, pure P2P networks and hybrid P2P networks. In centralized P2P network (Napster, 2011), a central server is used which manages the files and user databases of multiple peers that log onto it. These networks provide the highest performance but suffer from lack of scalability and a single point of failure. In pure P2P networks (Freenet, 2000), all the peers have similar responsibilities acting as both server and client. These networks offer inherent scalability and avoidance of a single point of failure but at a cost of slow information discovery and increased overhead traffic. A hybrid network (iMesh, 1999) combines the features of both the centralized and pure P2P networks. Within these networks, some peers on the basis of their resources (storage, CPU, etc.) are elected as super peers. The super peers are assigned with responsibilities like maintaining a central index of the files shared by peers and helping a peer in establishing a relationship with another peer, etc. Hybrid P2P networks provide an efficient search mechanism with no single point of failure. Some hybrid P2P systems (Rodriguez-Perez, Esparza, & Muñoz, 2008) can be found in the literature that select super peers on the basis of their reputation among the peers.

2.2. Multimedia content protection mechanisms

Piracies of multimedia contents are increasing with the pervasive usage of content distribution systems. Mechanisms must be deployed to ensure that the multimedia content can be used safely by legitimate users who have appropriate usage rights of that content. In this section, we provide a brief overview of state-of-the art content protection technologies, i.e. digital watermarking, fingerprinting and buyer–seller watermarking protocols.

2.2.1. Digital watermarking and fingerprinting

Encryption can be used to package the multimedia content securely and enforce all access rules to the protected content. However, once the content is decrypted by an authorized user, it does not provide any protection to the content (Grangetto, Magli, & Olmo, 2006). Thus, encryption alone is not enough to prevent an authorized user from illegal re-distribution. Similarly, classic Digital Rights Management systems (Apple iTunes, 2001; Microsoft DRM, 2008), which are considered as a second line of defence against copyright violation, do not prove to be an effective access control against a user with the knowledge and determination to violate it. For content owners, digital watermarking proves to be a more effective anti-piracy solution. Digital watermarking has become a significant area of research and development, and the usage of these techniques is now being considered a requisite to address the issues faced by the proliferation of digital content. Watermarking consists of embedding a watermark into the content that can later be used to check the source of the content. There are two forms of watermarking, copyright watermarking and fingerprint watermarking (fingerprinting). In copyright watermarking, a watermark is embedded into the content which indicates the copyright holder's identification. This is used to