

文章编号: 2095-2163(2020)09-0025-04

中图分类号: TP311

文献标志码: A

# 区块链技术在原创作品保护系统中的应用

王春霏, 董 剑, 谷博文

(哈尔滨工业大学 计算机科学与技术学院, 哈尔滨 150001)

**摘要:** 由于区块链技术所具有的不可篡改性与时特性,使其天然适合于在作品的登记与保护领域发挥其作用,而网络上的大量原创作品又是当前普通版权保护技术覆盖不高的领域。针对这些原创作品保护不力的背景下,本文介绍了区块链技术在版权及原创保护领域存在的优势和意义,提出了一套能够广泛适用于原创作者的基于区块链技术的作品保护系统设计方案,并提出了一些关键的原创性保证技术。

**关键词:** 区块链; 版权; 原创保护

## The application of blockchain technology in the protection system of original works

WANG Chunpei, DONG Jian, GU Bowen

(School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China)

**[Abstract]** Due to the immutability and time characteristics of the blockchain technology, it is naturally suitable for playing its role in the registration and protection of works, and the large number of original works on the network are not covered by the current ordinary copyright protection technology field. Against the background of weak protection of these original works, this article introduces the advantages and significance of blockchain technology in the field of work protection, and proposes a set of design solutions for work protection systems based on blockchain technology that can be widely applied to original authors and put forward some key originality guarantee technology. Finally, the shortcomings of the system and the plan are prospected.

**[Key words]** blockchain; copyright; original protection;

### 0 引言

随着互联网日新月异,在其上运作着的数字内容也逐渐成为一个庞大的数字。据估计,到2020年网络上存在着的数据量将达到44 ZB,而其中不乏存在着大量的原创性内容及作品。目前对于著作或知识产权等,只能进行出版或申请专利保护。而对于很多自由撰稿人,如现在互联网上大量出现的网络作家而言,他们大多使用假名或匿名写作,使得他们的权利通常无法得到有效的保护。另外,知识产权的归属一直是网络上乃至社会上公认的容易存在纠纷的问题。此类事件的发生大多是由于当前对于知识产权保护的力度不够,版权信息不够透明。

区块链技术是新兴的技术之一,随着比特币的出现才逐渐进入人们的视线,并迅速获得了全球范围内的关注。区块链技术不仅仅是用于数字货币的特定技术,还是一种能彻底解决信任问题、去中心化问题的重要工具。由于其不可篡改、公开透明等特点,目前已经被广泛使用在物联网、金融、保险、物流、公益等领域。

本文设计并实现了一种面向原创作者的基于区

块链技术的数字知识产权保护方法与系统,对于无论是何种形式、何种类型的原创内容生产者来说,都能通过这种方法使得自己的知识产权得到保护<sup>[1]</sup>。相对应的,已经拥有著作权、出版物或是商标权的所有权人,也可以加入到本方法保护的范围内,使其能够更方便地行使自己的权利。

### 1 现状与需求分析

由于传统知识产权保护存在自主申请不便、保护范围较小、惩罚难度较大等缺点,现将视线瞄准向了区块链技术。目前,区块链技术能够有效解决以下几类传统知识产权保护存在的问题:

(1) 权利主体的认定存在不确定性<sup>[2]</sup>。传统书面或实体多媒体作品在流转和传播过程中,通常会存在出版社水印、防伪等标识,以及相关的版权和作者信息。但是网络作品就难以附加此类信息,如作者的身份信息等,即使通过特殊加密方式隐藏在数据中,也有被篡改或破解的可能。通过引入区块链技术,在作品生成时,利用特殊算法对作品及作者、权利要求等产生一个数字指纹信息,将该指纹信息通过区块链保存<sup>[3]</sup>。该指纹信息一旦上链,就不可再更

**作者简介:** 王春霏(1997-),男,硕士研究生,主要研究方向:区块链、分布式计算;董 剑(1978-),男,博士,教授,博士生导师,主要研究方向:分布式系统、容错计算、区块链;谷博文(1998-),男,本科生,主要研究方向:区块链。

收稿日期: 2020-06-06

改,保证了真实作者一定会拥有对作品的所有权。

(2)作品的利用方式不合理。以往的知识产权作品在使用方式上通常是授权的模式,即作者授权出版商出版或是传播。然而,不合理的授权许可模式不仅有碍于优秀作品的自由广泛传播,还会有悖于著作权法鼓励创作、促进作品传播的初衷。例如,腾讯旗下的QQ音乐与多家国际唱片公司签订了独家授权合同,在国内占据了近90%的总曲库授权,其他音乐公司若想要使用这些音乐,均需要向QQ音乐付费,几乎成为了实际意义上的音乐版权管理组织,这极大降低了作品传播的范围。如果将作品通过区块链保存,每当需要使用该作品时,由使用者向区块链网络提出请求。这一过程突破了原有的授权管理模式,跨越了管理组织、代理平台等中介机构,实现了著作权人与消费者直接的接触,实现了智能许可、自动化的报酬分配等目的,所有过程会被区块链所记录,形成了透明化的产业链。

(3)侵权成本较低且难以认定。当前网络上广泛流传着诸多盗版或是侵权的作品,通常是由盗版者手工复制或翻录发布在网络上,这一过程非常难以追溯。当发现有侵权行为时,很多网络著作权人无法准确提供自己网络作品的创作证明或是所有权证明,因此也就无法对侵权行为予以打击。如果借助区块链技术来记录作品的发布、转让或许可等过程,根据这一透明的价值链,就可以很方便地追溯到作品的来源以及每一次授权的信息,从而解决了著作权人难以举证的问题<sup>[4]</sup>。

本文所述系统是为了保护越来越多的网络普通

创作者的知识产权,维护所有人的合法权益,保护网络空间对知识产权的尊重。本项目由前端,后端,本地数据库,区块链部分组成,基本的流程是:用户注册登录,用户账户密码等数据保存在本地数据库,登录后用户可上传将自己的作品上链,通过特征提取算法将作品压缩,将压缩后的作品转换为字符串进行哈希运算,得出一个定长哈希值,将这个哈希值传到区块链上的同时将作品原件存储到IPFS分布式数据库中<sup>[5]</sup>。这个上传过程需要查重,检查该上传的作品是否已经被保存过版权,防止盗版侵权行为。

根据系统的功能需求,可以主要划分为以下功能:

(1)用户交互:用户可视化的交互界面。

(2)作品保存:将用户上传的作品原件做备份保存,以备提取使用。

(3)特征提取:提取作品的唯一指纹信息,作为原创检测及区块链存证的证据。

(4)原创检测:防止系统中存在有恶意修改原创作品并试图上传为自己作品的行为。

(5)区块链交互:与区块链网络交互,以保证上传登记与查询。

本系统设计有3种不同的用户登录权限,分别是管理员、原创作者与查询者。不同用户可以使用的权限不同。

## 2 系统设计

### 2.1 模块设计

根据系统设计要求,系统功能共有5个模块构成,如图1所示。

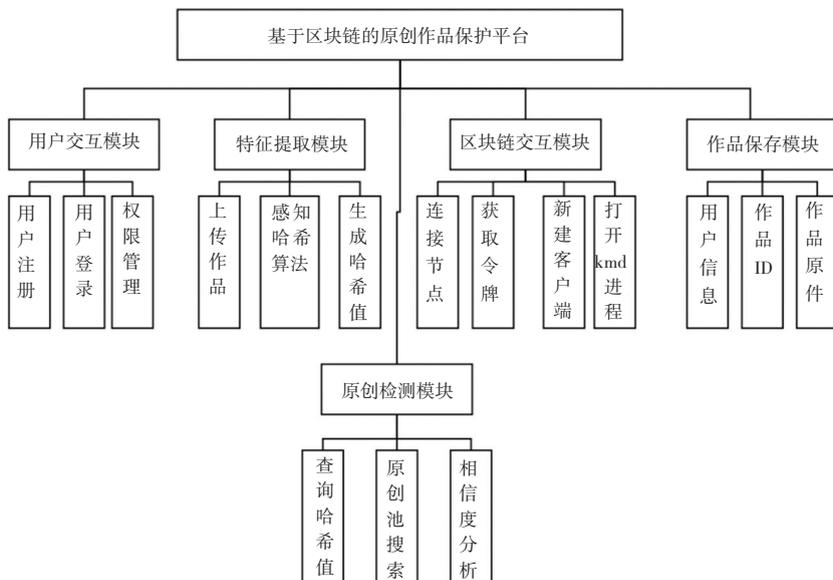


图1 系统功能设计

Fig. 1 System function design

(1)用户交互模块。对用户进行管理,未注册的用户可以账号注册,通过不同权限区分了普通用户和管理员。同时提供给用户可视化的操作界面,是其余模块与用户的桥梁。

(2)特征提取模块。用户上传文件到后台,进行特征提取。将文件原件通过感知哈希算法处理,提取过的文件对小幅的修改如旋转、拉伸等具有抵抗性,最后文件转化为一个64位的哈希值,即为作品所对应的唯一特征值。

(3)区块链交互模块。系统与区块链网络通信的端口,使用服务器的令牌创建本地客户端,并与远程节点连接,向系统提供打包新交易或查询的功能支持。

(4)原创检测模块。向区块链提交新作品信息前,需要先通过特征提取模块将文件特征提取为哈希值,并在已有存证中检查是否已存在该特征值,如果没有发现重复就认定为是原创作品,上传后将返回的交易ID作为索引保存;若发现具有相似特征值,则拒绝上链。

(5)作品保存模块。作品信息存储到区块链后,还必须把作品原件无压缩的存储到服务器,以保证作为证据使用。本项目采用星际文件系统IPFS,同时将作品的具体信息及备注也存储起来,IPFS为每个作品生成唯一地址进行访问。

### 2.2 架构设计

本项目使用MVC架构设计实现,主要框架如图2所示。

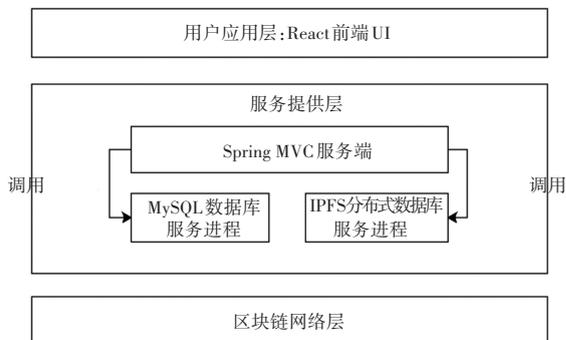


图2 系统架构设计

Fig. 2 System architecture design

图2中自上到下依次为用户应用层、服务提供层以及区块链网络层,用户请求由网页界面开始,经过服务层,最终到达区块链服务。用户应用层与服务提供层通过前后端分离的程序编写方式,React构建的前端界面是一个独立运行的程序,通过fetch与后端交流数据。MySQL数据库作为系统记录作品

历史及存储信息的核心数据库。Spring内置的Web服务器Tomcat为前端用户操作提供相应服务,根据用户请求自动编译代码将页面资源发送给请求地址。

### 3 关键特征提取技术研究

目前已有的基于区块链对原创作品提供保护的成熟方案中,所采用的图像特征提取技术实际上只是对于作品文件内容的摘要提取,即使用密码学哈希函数。这一过程极易被规避,只要被压缩的内容有轻微改动,所产生的哈希值就会有极大变化。对于这种情况就有可能存在恶意篡改作品,并当作自己的作品登记的行为。

本文根据消息摘要过程中信息的损失,提出了一种改进的特征提取方案,其主要原理是图像学中的感知哈希算法<sup>[6]</sup>,图像的特征提取方案如图3所示。

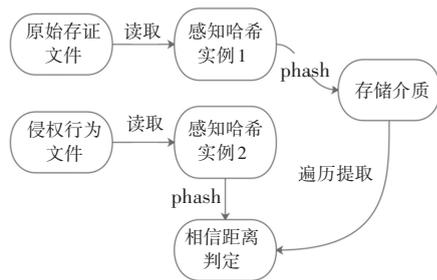


图3 图像的特征提取方案

Fig. 3 Image feature extraction scheme

本方案同样需要先将原始的作品文件提取特征值并保存在存储介质中,若发生侵权行为,就将侵权行为文件也做一次提取,并在已有存储库中和已有记录进行比对。

不同之处在于本文在感知哈希处理过程中,每次处理的过程都会产生一个实例,对于不同的图片大小,需要根据其长度与高度的比例,计算出能够保证最终的phash值长度相同的一组参数。通常情况下,会将该参数设定为8或8的整数倍,以保证最终能够有效产生具有充分意义的16进制的phash值。

在对侵权行为检测时,根据产生的侵权文件的phash值去和已有存证中所有的记录比对。该比对过程使用的是汉明距离(Hamming Distance, HMD)检测方案<sup>[7]</sup>。汉明距离代表着两个字符串之间字符的差异数,是用于衡量字符串相似程度的指标,采用式(1)计算到两个图片的相似性。

$$Dissimilarity = \frac{HMD}{(phash.length)} \quad (1)$$

即汉明距离/感知哈希值长度,这个比例值记作

不相似度(Dissimilarity)。在比较过程中,若发现与某个已有作品的不相似度较低,则代表着它们之间的差异可能较小,就有较高概率发生侵权。

#### 4 分析与展望

与现有技术相比,本文所提出的基于区块链的原创作品保护系统的优点在于:

(1)提出了一种基于区块链的作品登记的方法,实现了快速作品登记、自动防抄袭原创检测、通过区块链固化证据等重要功能,并能利用区块链的特性为原创作者提供强有力的创作证据,为改善创作环境作出很大贡献。

(2)提出的原创图片登记与保护系统,使用便利,使广大原创作者足不出户就能自主、快速地将图片登记存证,不再需要经过著作权等流程的处理,同时减轻作者与线下操作人员的工作负担。

(3)通过高效的特征提取算法配合重复检测算法,不仅为作者将作品提交到区块链上提供了便利,还提供了强有力的保护,确保了原创作品的唯一性和价值,有效避免抄袭、盗用等恶意行为。

(4)系统不仅适用于网络作者的小规模使用,同时能够适用于批量作品的处理,内部模块设计完善、稳定性与性能良好,登记效率高,能够为日后第三方机构接入提供支持。

但是,本文所述系统及其所提供的高效特征提取算法与原创检测算法目前还均不适用于除图片形式之外的原创作品,如文本、视频等。文本内容目前只能通过密码学哈希函数解决,而视频文件可以通过将其按照关键帧拆分,进而形成一组相关联的图片作品,再利用本方案中的特征提取方案,就能够形成同样有效的保护<sup>[8]</sup>。另外,对于区块链网络还远

远没有利用最大化,后期系统可以进一步建立在区块链智能合约上,形成完全分布式的运行环境。

#### 5 结束语

本文提出了一个利用区块链技术构建的原创作品保护系统,设计了较为新颖的特征提取技术与原创检测方案,使作品在上传到区块链网络存证之前有了一次筛选,避免了恶意抄袭行为的产生,有利于原创作品的发展。由于区块链的不可篡改性与时戳特性,使得系统构建出了一个具有可信力的原创作品保护环境,这也恰恰体现出了区块链技术所希望建立的信任的价值。

#### 参考文献

- [1] MA Z, JIANG M, GAO H, et al. Blockchain for digital rights management[J]. Future Generation Computer Systems, 2018, 89: 746-764.
- [2] SAVELYEV A. Copyright in the blockchain era: Promises and challenges[J]. Computer law & security review, 2018, 34(3): 550-561.
- [3] TSUBOUCHI M, MAEDA T, OKABE Y. Digital Fingerprinting on Executable File for Tracking Illegal Uploaders [P]. Computer Software and Applications Conference Workshops (COMPSACW), 2014 IEEE 38th International, 2014.
- [4] 夏朝羨. 区块链技术视角下网络版权保护问题研究[J]. 电子知识产权, 2018(11): 110-116.
- [5] 高文涛, 张桂芸. 基于联盟区块链和 IPFS 的音乐共享模型[J]. 天津师范大学学报(自然科学版), 2020, 40(2): 68-74.
- [6] Mehta R, Kapoor N, Sourav S, et al. Decentralised Image Sharing and Copyright Protection using Blockchain and Perceptual Hashes [C]//2019 11<sup>th</sup> International Conference on Communication Systems & Networks (COMSNETS). IEEE, 2019: 1-6.
- [7] NOROUZI M, FLEET D J, SALAKHUTDINOV R R. Hamming distance metric learning [C]//Advances in neural information processing systems. 2012: 1061-1069.
- [8] GUAN G, WANG Z, LU S, et al. Keypoint-based keyframe selection[J]. IEEE Transactions on circuits and systems for video technology, 2012, 23(4): 729-734.