Research on Copyright Protection Model of Digital Works Based on Blockchain and Digital Watermark

Yihong Zhao[a], Shuqi Fan[b]

School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China

[a]jinghongyipie9@163.com, [b]1030103385@qq.com

Keywords: Internet, storage, blockchain.

Abstract: The emergence of copyright issues has been around for a long time. With the advent of the Internet era, the ability of humans to store data has greatly improved, and digital storage of valuable information has been realized. Aiming at the problems of copyright registration, piracy and copyright disputes faced by digital intellectual property under the Internet ecology, this paper proposes a management and control model combining blockchain, digital watermarking, smart contract, and perceptual hashing. Through the decentralization, difficult tampering, traceability, extensibility, openness and transparency of the blockchain, all kinds of key information of the digital works recorded in the data storage platform are written into the block file, in the registration and other aspects. The integrity and security of the work can be verified at any time, and the digital works are guaranteed to be credible, the user operation is traceable, and the model is impartial and authoritative. The experimental results show that the model reduces the threshold of digital copyright registration, enhances the authority of copyright authentication, and has better real-time and robustness.

1. Introduction

Infringement and piracy restrict the further development of digital publishing, and all parties are deeply affected. In particular, content producers such as authors have always been in a weak position, lacking corresponding voice and dominance, and their creative enthusiasm has been hit hard. Faced with these problems, the state attaches great importance to it, and various policies and support plans are frequently introduced to solve the problem of copyright protection. However, due to the limited technical level, it is difficult to solve it fundamentally.

This paper introduces blockchain, digital watermarking, smart contract, and perceptual hashing technology into the field of digital copyright, and proposes a new management and control model, which can provide full chain management and control services for copyright registration, copyright inquiry, copyright transaction and protection for Internet digital publications. This paper uses perceptual hashing technology to judge the similarity of original works; establishes the mechanism of smart contracts to automatically execute preset instructions, constructs copyright registration, copyright verification and copyright transaction mode without trusting third parties; these technologies work together to create a credible, authoritative, efficient and transparent control model and solve a series of problems existing in current digital publications.

2. Implementation of Model Technology

2.1. Perceptual Hash Technology

In order for the uploaded work to be considered valid, it must be less similar to any previous document of the model. This article uses Perceptual Hashing to perform this comparison. Perceptual Hashing generates short fingerprints for files, and small changes in the work can result in minor changes in the fingerprint. If the fingerprints are too similar, it is concluded that the files are too

similar. Next, we use the audio file as an example to show the comparison of the works. The implementation of the rest of the publications is similar.

The frequency is converted to notes for processing, so the result is a 12-sound range, with each note corresponding to a range. After some filtering and standardization, we will end up with the following image:

Fig. 1 Filtered and standardized audio results

The easiest way to identify the voiceprint is to calculate the "bit error rate" (Hamming distance) for the voiceprint.

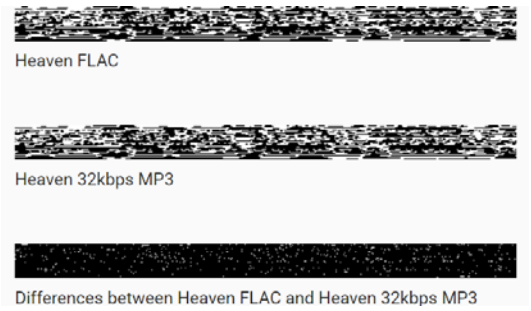

Fig. 2 Comparison of voiceprints of similar musical works (Left)

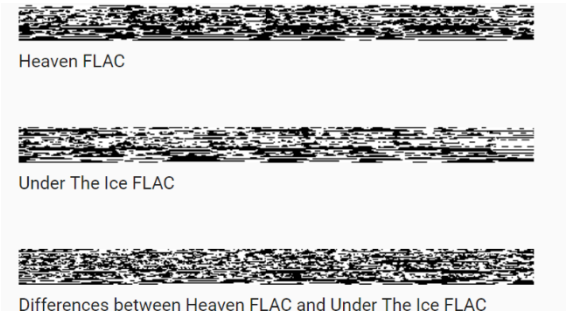

Fig. 3 Comparison of voiceprints of different musical works (Right)

For example, if you upload Heaven music, the difference between FLAC version and MP3 version is not significant. From Perceptual Hashing, it is the same work.

If you upload an unrelated file, you can see that the difference between the two sample audio files of Heaven and Under the Ice is larger, and the error rate is higher (the Hamming distance is larger).

2.2. Smart Contract Technology

Smart contracts deployed on Ethereum do not require human involvement, as long as the contract code requirements are met, a fully automated process is implemented. This technology saves time, reduces costs, makes transactions more accurate, and cannot be tampered with. Smart contracts are not interfered with by any third-party agency, which further decentralizes.

On Ethereum, this model deploys two smart contracts: pxCoin.sol, pxAsset.sol and three Ethereum-related standard token protocols, security protocols: ERC20, Safe Math and ERC721 to ensure the auto-execution trustworthiness of the contract and decentralized.

2.2.1. PxCoin Contract.

PxCoin is a smart contract developed using the Solidity language and related to transactions. It stipulates the following behavior:

1) Ethereum network total ETH circulation. Total Supply ()
2) The total amount of ETH owned by the user. Balance of (address _owner)
3) Transaction transfer realization. transfer (address _to, uint _value)
4) Transfer authorization for the transaction. approve (address _spender, uint _value)
5) Transaction reconciliation of transactions. check (address _to, uint _value), etc.

When the client's conditions are met, the function corresponding to Ethereum is automatically executed, thereby realizing the author's work authorization and copyright transfer.

2.2.2. PxAsset Contract.

PxAsset is a contract that maintains the user's own copyright on the chain. It stipulates the following behavior:

1)Add a new digital work for an author, the parameters are: content hash value, price tag, Ethereum resource, metadata, voting number and digital works copyright owner's account address in Ethereum.

New Asset (string _content hash, uint _price, uint _wei, string _metadata, uint _vote Count, address _upload Address)

2) Put the work into the sales area. Sell (uint id, uint price)

3) Bid for the work. Bid (uint _price, address _to, uint _token Id, address _from)

4) Record copyright transfer. Transfer (address from, address to, uint256 token Id)

5) Record copyright authorization. Approval (address owner, address approved, uint256 token Id), etc.

2.3. Digital Watermarking Technology

In this model, the DCT algorithm is used to add a blind watermark to the digital work, and the author information is embedded to facilitate the identification of unauthorized theft and guide the user to obtain legal authorization.

The realization of the DCT watermarking algorithm is as follows:

1) Each frame of the original image and video is divided into 8×8 sub-blocks, and the image is subjected to DCT transform, and the transformed result is the frequency domain.

2) Next, the n most important frequency components are selected from the DCT coefficients of the transformed data D to form a sequence.

3) Take the key to generate a pseudo-random column, that is, a watermark sequence, in which a random number satisfying the Gaussian distribution, and then superimpose the selected DCT coefficients with a pseudo-random Gaussian sequence to generate a watermarked sequence.

4) Convert V to D and finally inversely convert to an image with a watermark. Watermark detection relies on a threshold. When the detection result exceeds the threshold, it is judged that it contains a watermark, otherwise it is judged not.

2.3.1. Digital Watermark Embedding Based on DCT Algorithm.

The DCT transform has many advantages, such as low bit error rate, high information concentration, and good complexity. The two-dimensional discrete DCT transformation formula is as follows:

$$F(u,v) = c(u)c(v)\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\cos\frac{\pi(2x+1)}{2M}\cos\frac{\pi(2x+1)v}{2N}$$

Note: u = 0,1,2, M-1; v = 0,1,2, N-1; M, N are the row and column values of the image block.

The formula for DCT inverse transformation is as follows:

$$F(x,y) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} c(u)c(v)f(u,v)\cos\frac{\pi(2x+1)u}{2M}\cos\frac{\pi(2y+1)v}{2N}$$

2.3.2. Watermark Extraction Process.

The block diagram of the basic process is shown in Fig. 4, which is the inverse of the watermark embedding process:

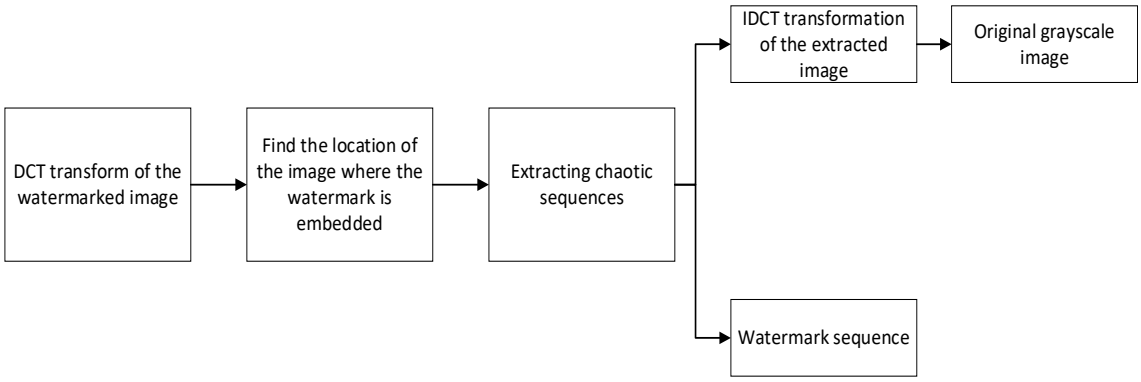

Fig. 4 Flow chart of digital watermark extraction

3. Model Experiment and Analysis

3.1. Security Analysis

In order to ensure that stored digital files can be delivered to users quickly, completely and securely, this model can use "IPFS Distributed Interstellar Storage System". IPFS provides a high-throughput, content-addressable block storage model with content-related hyperlinks. This forms a generalized directed acyclic graph (DAG). IPFS does not have a single point of failure, and nodes do not need to trust each other. Distributed content delivery can save bandwidth and obtain corresponding data from distributed storage systems through hash values. The storage cost is 1/3 of the cost of centralized storage, especially suitable for data storage and security requirements of startups.

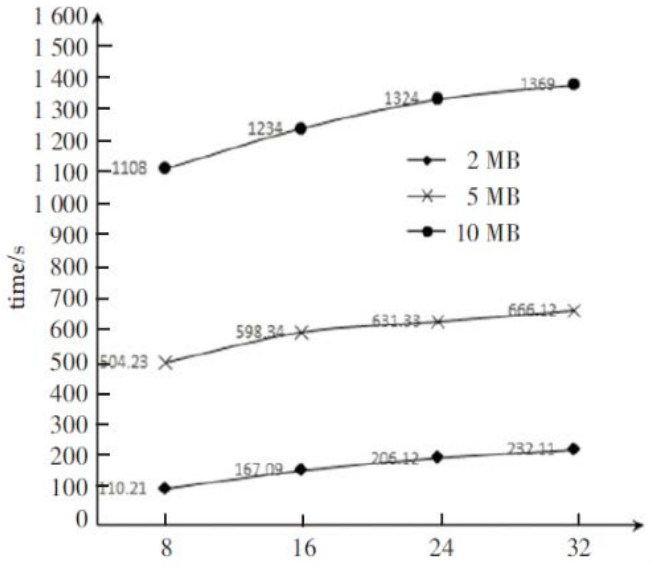

Fig. 5 Model stability test

4. Conclusion

In this paper, a blockchain, digital watermark, smart contract, and perceptual hashing technology are used to design a trustworthy digital work management model. The biggest highlight of the model is that it solves the problem of single and centralized management of the works in the current traditional copyright management and trading system, and introduces the advanced technology of digital watermark to provide traceability method. The use of IPFS distributed storage greatly saves the cost and provides an efficient solution to reduce current electronic product infringement. Blockchain has the decentralized technology trust feature. Combined with blockchain technology and cryptography technology, digital original works can be tamper-proof, irreversible, reliable, open and

transparent. Digital works can be effectively guaranteed in the aspects of confirming, using and safeguarding rights.

However, the open and transparent nature of blockchain technology is a double-edged sword. It can inadvertently reveal some personal information of users while ensuring the transparent operation of the system. In the next step, dynamic transaction signature technology will be combined with blockchain technology to protect users' personal privacy information while making good use of the blockchain technology's features such as non-tampering, openness and transparency.

References

[1] Xu R, Lu Z, Zhao H, et al. Design of Network Media’s Digital Rights Management Scheme Based on Blockchain Technology[C]// IEEE International Symposium on Autonomous Decentralized System. 2017.

[2] Savelyev A. Copyright in the Blockchain Era: Promises and Challenges[J]. Hse Working Papers, 2017.

[3] Meng Z, Morizumi T, Miyata S, et al. Design Scheme of Copyright Management System Based on Digital Watermarking and Blockchain[C]// IEEE Computer Software & Applications Conference. 2018.

[4] Kosba A, Miller A, Shi E, et al. Hawk: The Blockchain Model of Cryptography and Privacy-Preserving Smart Contracts[C]// 2016 IEEE Symposium on Security and Privacy (SP). IEEE, 2016.

[5] Bell T W. Copyrights, Privacy, and the Blockchain[J]. Social Science Electronic Publishing, 2016.

[6] Sicilia M A, Salvador Sánchez Alonso, Elena García Barriocanal. Sharing Linked Open Data over Peer-to-Peer Distributed File Systems: The Case of IPFS[C]// Research Conference on Metadata & Semantics Research. Springer, Cham, 2016.