Research and Practice on Blockchain Technology for UGC Content Moderation: Postprint

Authors: Zhang Fucheng, Zhishuai Guo, Xiao Yong

Date: 2023-10-08T00:00:00+00:00

Abstract

As a crucial breakthrough in China' s independent innovation of core technologies, blockchain technology will undoubtedly play a pivotal role in the domain of deep media convergence. This paper analyzes the necessity of risk control for UGC content review, the characteristics of blockchain technology systems, and the current issues confronting UGC content risk control. It elaborates on the construction approach and key priorities of a blockchain-based UGC content review risk control system, integrating blockchain-based content review standards. Furthermore, based on practical system construction experience, it clarifies the functional architecture design of the system, aiming to ensure the safe production of news content through the practice and exploration of blockchain technology.

Full Text

Research and Practice of Blockchain Technology in UGC Content Moderation

Zhang Fucheng, Guo Zhishuai, Xiao Yong
Neusoft Group, Shenyang, Liaoning 110179, China

Abstract: As a critical breakthrough in China' s independent innovation of core technologies, blockchain technology will undoubtedly play a pivotal role in the deep integration of media. This paper analyzes the necessity of risk control in User-Generated Content (UGC) moderation, the characteristics of blockchain technology systems, and the current challenges facing UGC content risk control. Building upon blockchain-based content moderation standards, it elaborates on the construction 思路 and key focuses of a blockchain-based UGC content moderation risk control system. Based on practical system construction experience, the functional architecture design is clarified, aiming to ensure safe

production of news content through the practice and exploration of blockchain technology.

Keywords: Content Moderation; Blockchain; Data Trustworthiness; Full-Process Traceability; Platformization
Classification Code: G712
Document Code: A
Article ID: 1671-0134(2021)11-020-04
DOI: 10.19483/j.cnki.11-4653/n.2021.11.005

1. Challenges Facing UGC Content Risk Control

In 2018, with the development of blockchain technology, a wave of blockchain applications emerged in the media industry both domestically and internationally, as the sector explored how blockchain could add value to media convergence and attempted to leverage the technology to enhance the depth and breadth of integrated media development.

China' s relevant authorities and policymakers have always enforced strict standards for online content compliance. The advent of 5G technology has made content more massive and instantaneous, rendering content supervision highly challenging [3]. Consequently, there is a trend toward increasingly detailed regulatory standards and heavier penalties. UGC platforms, which rely on user-generated content as their core business, naturally become high-risk zones for content violations. Platforms or accounts with violations face consequences ranging from warnings and summons to indefinite shutdowns. For platform managers, especially traditional state news organizations such as newspapers and television stations venturing into the internet to occupy the main battlefield, content risk control is a matter of survival. Decisions regarding user experience and platform operation strategies must be based on content moderation risk control.

The primary risk factors in UGC content include:

- **Illegal and Non-Compliant Content:** This represents the most severe risk category and the bottom line for moderation control, requiring the strictest measures; otherwise, platform survival is at stake.
- **Content Affecting User Experience and Platform Atmosphere:** Including insults, aggressive information, meaningless spam, and advertisements. Such content poses moderate risks.
- **Content Affecting Platform Direction:** Platforms typically have their own content focus, particularly media organizations like newspapers and broadcasting companies with specific requirements regarding political and serious content. Content that deviates from correct public opinion orientation should not be encouraged, as it affects brand image. For instance, excessive entertainment gossip would negatively impact professionalism,

though such content carries lower risks.

- **Content Endangering User Safety:** Such as fraud and pyramid scheme information. As content platforms, they have an obligation to protect user property and safety, which also reflects the platform' s rigor and authority. Such content requires focused attention.
- **Copyright Infringement:** Information technology development has exacerbated digital copyright infringement. Copyright information concerns the legal operation of platform content, and using or disseminating infringing content exposes platforms to compensation claims, making this a key moderation focus.

2. Core Characteristics of Blockchain Technology

Blockchain is a distributed database system characterized by multi-node participation in data storage. Any node stores identical data through consensus algorithms, ensuring secure and reliable storage. Its technical features—decentralization, immutability, high scalability, and flexibility—provide novel solutions and critical support for digital content copyright protection and information encryption. Blockchain is a composite technology encompassing hash algorithms, asymmetric encryption, digital signatures, timestamps, and digital certificates. By storing digital abstracts, it makes digital content non-forgeable and tamper-proof, achieving evidence preservation and providing trustworthy data support for authentication, confirmation, and traceability.

Blockchain brings new possibilities to the media industry in combating fake news, content moderation, copyright protection, and targeted advertising [1]. Through disruptive innovative computing models, blockchain has formed a new generation of converged media technology based on next-generation information technology, guiding the future direction of traditional and new media convergence in China [2].

UGC (User Generated Content) originated in the internet sector, where users produce and share original content on platforms. UGC is not a specific business but rather an interaction mode for internet users, transforming from passive information consumption to active content provision and upload. With China' s mobile internet development, users have become both content viewers and creators, particularly through official and self-media platforms like WeChat subscription accounts, Douyin accounts, Toutiao accounts, and mainstream news client accounts. UGC has significantly impacted content dissemination, services, and e-commerce.

[Figure 1: see original paper] Blockchain Technology Core Characteristics

The core characteristics of blockchain technology include:

- **Security and Reliability:** Single node failure does not affect system services; cryptography protects data security.

- **Decentralization:** No central server; data is stored distributedly, providing high redundancy.
- **Time-Series Data:** Data is timestamped and chronologically constitutes the blockchain.
- **Immutability:** Modifying data requires altering most nodes; data structure is directly related to content, and modifying historical data affects the chain structure.
- **Data Transparency:** Data and execution rules are transparent to all relevant parties.
- **Smart Contracts:** Programs execute business terms simultaneously across nodes, with contract content public to relevant parties.

3. Development Status of Blockchain-Based Content Moderation

On April 2, 2021, the National Radio and Television Administration officially released the *Blockchain-Based Content Moderation Standard System (2021 Edition)*. This standard system' s core objective is to promote healthy and sustainable media content development through blockchain-based moderation standards. The system standardizes the entire process of blockchain-centered content moderation, covering business processes, management procedures, and security strategies to guide technology application and platform construction.

The standard system elaborates detailed criteria across four aspects: system, business, security, and management.

- **System Standards:** Specify technical requirements for blockchain content moderation system architecture, data formats, and system interfaces, including overall technical specifications, media content classification and marking, storage and sharing modes, block data formats, inter-system interfaces, and cross-chain interaction protocols.
- **Business Standards:** Define standards for content moderation, traceability, and re-review processes from a business workflow perspective.
- **Security Standards:** From a moderation security perspective, focus on system security technical requirements and security management audit specifications to standardize blockchain content moderation security.
- **Management Standards:** From a regulatory perspective, detail standards for supervision, content moderation evaluation, system and node assessment, aiming to normalize regulatory requirements for blockchain-based content moderation.

With these standards, media organizations and technology companies have gradually begun experimenting with blockchain-based content moderation platform construction, providing practical foundations and valuable experience for standard implementation.

Through blockchain-based moderation platform construction, the entire moderation process and results can be recorded, providing credible evidence for subse-

quent infringement disputes and content accountability. This imposes stronger constraints on both content producers and moderators, thereby ensuring and enhancing moderation rigor and continuously improving content production safety.

4. Key Considerations in Blockchain-Based UGC Content Moderation Platform Construction

4.1 Technology Architecture Selection

Blockchain is both a technology and an architectural application, making architecture implementation the core of the technology system. The blockchain technology system includes the network layer, consensus layer, data layer, smart contract layer, and application layer. UGC content moderation risk control systems primarily build upon the application layer, which includes smart contracts for blockchain interaction and corresponding frontend WEB applications. The WEB application system calls smart contracts through interfaces to store and retrieve data according to business rules.

Blockchain infrastructure can be built in two ways: public cloud services or private chain construction. Public cloud services suit organizations with average development capabilities, allowing direct utilization of cloud blockchain for business applications, though fees are charged based on usage duration or volume. Organizations with strong R&D and maintenance capabilities can build their own blockchain platforms, offering greater controllability and autonomy. For the underlying blockchain of a UGC content moderation risk control system, flexible selection based on the builder' s characteristics is necessary.

4.2 On-Chain Encrypted Storage of Moderation Data

UGC content is characterized by massive volume, fragmentation, and diverse types. The moderation process requires storing UGC content information, moderation process information, and moderation results on-chain to ensure data integrity and support full-process traceability. Specific data includes:

- **Content Information:** System unique identifier, publisher, publication time, channel, column, and node.
- **Process Information:** Moderation type, start/end time, moderator, and duration.
- **Result Information:** Moderation result, risk flag, level, and warning classification.

Given the numerous associated information elements, storing everything on-chain would consume excessive storage space and computational time. Therefore, each component can be hashed, with only the hash codes stored on-chain for encrypted storage.

4.3 Decentralized Consensus Mechanism Ensuring Tamper-Proof Full-Process Traceability

"Blockchain' s decentralization and distributed data storage can make news publication authentication more rigorous; each node stores independently and cannot be artificially erased" [4]. Similarly, blockchain' s decentralized consensus mechanism ensures that data at each stage of content moderation cannot be maliciously tampered with. For any content item, the business system can retrieve corresponding information from the chain based on a unique identifier and parse it to display both the content and associated moderation information. If malicious tampering occurs, the hash code stored on-chain will differ from the hash generated from the altered data.

At the system implementation level, moderation objects and information are stored within the business system, with corresponding hash codes stored on-chain. Therefore, any tampering with business system data will cause hash value mismatches with on-chain data, revealing the tampering. The moderation platform focuses on recording a trustworthy full process, using multi-stage review processes to determine UGC content violations and storing each moderation instance on the blockchain to achieve immutability.

4.4 Establishment of Multi-Type Content Moderation Business Processes

Traditional content moderation processes are relatively simple: data is reviewed, results are generated, and records are stored in relational databases. After integrating blockchain technology, comprehensive moderation processes must be established for different data types to clarify blockchain' s role and application methods at each business stage.

Example Process Planning:

- **Occupationally Generated Content (OGC):** Content production → Content moderation (three reviews and three proofreads) → Hash codes of moderated content, process, and nodes stored on-chain → Post-production review → Hash codes of post-review results stored on-chain.
- **User Generated Content (UGC):** Content production → Machine moderation → Hash codes of content and moderation results stored on-chain → Manual re-review → Hash codes of reviewer and results stored on-chain.

Additionally, for diverse data types such as images, videos, and text, different process strategies should be implemented based on business needs—for instance, storing keyframes for videos and keywords for text—to control on-chain information volume and improve query performance.

4.5 Establishment of Cross-Chain Sharing Mechanism for Moderation Information

Among blockchain' s challenges, poor interoperability significantly limits application scope. Cross-chain technology is crucial for achieving blockchain interconnectivity and enhancing interoperability and scalability, whether for public or private chains [5]. Cross-chain applications must address "how to confirm," "where to confirm," and "who will confirm" issues, typically using Notary Schemes, Sidechains/Relays, Hash-locking, Distributed Private Key Control, or Atomic Swaps.

The core of cross-chain information sharing for content moderation platforms lies in who performs mutual trust authentication between different chains and shares moderated data and results. This can be addressed by establishing a cross-chain service center with core functions including:

- **Permission Authentication:** All organizations must register with the service center and obtain authentication keys; all inter-system data sharing interfaces require key-based authorization.
- **Data Decryption:** Provides decryption operations for blockchain-shared data, retrieving complete information from corresponding business platforms using on-chain hash codes.
- **Information Verification:** Provides external verification interfaces to validate complete information against on-chain hash codes, ensuring shared data authenticity.
- **Data Standardization:** Supports standardized format conversion according to service center data interaction templates for sharing between heterogeneous systems.

4.6 Platformization Construction and Operation

The core of content moderation is credibility, and blockchain is precisely the powerful technical system for enhancing and demonstrating credibility. Therefore, blockchain-based moderation systems are more readily accepted. Platformization thinking is a crucial focus during system construction, providing foundational support for subsequent moderation business expansion.

Platformization is achieved through data separation and multi-tenant architecture, storing data for each user and business type in logically separated spaces on the blockchain. Functions are modularly configured to tailor moderation capabilities for platform users. Permission and role settings support data isolation and interoperability between different platform users. Platformization is the foundation of services, while the service process embodies platform value and drives platform iteration and capability enhancement. Thus, leveraging blockchain' s advantages for platformized construction and operation is a key focus for content moderation systems.

5. Functional Architecture of Blockchain-Based UGC Content Risk Control System

The blockchain-based content moderation system comprises several main functional modules: blockchain basic services, cross-chain service center, platform basic function support, multimedia content moderation, manual review, infringement detection and analysis, and multi-tenant platform services.

The system' s data core consists of the business database and blockchain distributed storage. Business data stored in the database works in conjunction with hash codes stored on-chain to achieve on-chain storage while controlling data volume. The system achieves end-to-end application from moderation to service through multi-type data content moderation, manual review, and external service platforms. Standardized security and interface service systems ensure system reliability, security, and open external services.

As General Secretary Xi Jinping emphasized during the 18th collective study session of the Political Bureau of the CPC Central Committee, integrated blockchain application plays an important role in new technological revolution and industrial transformation. We must treat blockchain as a critical breakthrough in independent innovation of core technologies, clarify main directions, increase investment, tackle key core technologies, and accelerate blockchain technology and industrial innovation development.

Blockchain-supported UGC content moderation builds upon traditional moderation systems by leveraging blockchain characteristics to store moderation processes and content on-chain, thereby enhancing security, openness, sharing, and credibility. With the release of the *Blockchain-Based Content Moderation Standard System (2021 Edition)*, blockchain applications in content moderation will see more exploration, generating more excellent products and solutions to ensure safe content production.

References: [1] Wu Di. Reflections on Blockchain Technology Application in the Media Industry [J]. Office Operations, 2021(14): 158-159.
[2] Yang Mei. Research on Blockchain Technology Promoting China' s Media Convergence Development [J]. Media, 2021(5): 38-40.
[3] Feng Yuxiao. Research on Short Video Content Supervision [D]. Beijing: Beijing Institute of Graphic Communication.
[4] Ji Haoren. Application Prospects of AI+Blockchain in News Reporting [J]. China Media Technology, 2020(8): 32-34.
[5] He Shuai, Huang Xiangnian, Chen Xiaoliang. Review of Blockchain Cross-Chain Technology Development and Application Research [J]. Journal of Xihua University (Natural Science Edition), 2021(3).

Authors:
Zhang Fucheng (1982-), male, from Shenyang, Liaoning, Consultant at Neusoft Group Media Business Center, research direction: Media Technology.
Guo Zhishuai (1976-), male, from Shenyang, Liaoning, Business Director at

Neusoft Group Media Business Center, research direction: Media Technology. Xiao Yong (1978-), male, from Shenyang, Liaoning, General Manager at Neusoft Group Media Business Center, research direction: Media Technology.

(Editor in charge: Li Jing)

Note: Figure translations are in progress. See original paper for figures.