ISSN: 2633-4828 Vol. 4 No.2. September. 2022

International Journal of Applied Engineering & Technology

Blockchain-Based Music Metadata Copyright Protection Using Fuzzy Logic

[1]Vitri Tundjungsari and [2]Nindyo Artha Dewantara Wardhana

[1,2]Magister of Computer Science, Universitas Esa Unggul. Jl. Arjuna Utara No.9, Duri Kepa, Kebon Jeruk, Jakarta, Indonesia

[1]vitri.tundjungsari@esaunggul.ac.id and [2]megative.music@student.esaunggul.ac.id

Date of Submission: 10th June 2022 Revised: 03rd October 2022 Accepted: 14th October 2022

How to Cite: Vitri Tundjungsari et.al..,(2022). Blockchain-Based Music Metadata Copyright Protection Using Fuzzy Logic. International Journal of Applied Engineering &Technology 4(2), pp.101-105.

Abstract: **Copyright has been a big problem recently with false positives, especially in the music industry where it is a gray zone. While it does not sound that bad, in reality these false positives may cost millions of dollars for unsuspecting producers. It may incur an insurmountable loss for the big guns of the industry, but it may also potentially end the career of some independent producers of the industry. To help prevent copyright strikes for producers, the author proposed a block chain-based music metadata copyright protection as block chain is recently popular with the research of copyright protection. The protection will analyze the metadata of the music in question with the alleged plagiarist and determine how similar it is. With the help of Smart Contract, the block chain will help storing and associating the metadata of the music such as melodies, time signatures, lyrics, and other data for the rightful producer with confidentiality, integrity, and non-repudiation. It does not solve copyright infringement in general, but this helps reduce the number of false positives infringements dubbed as "accidental plagiarism" before publicizing. This helps with the efficiency of legal lawsuits against music that would take months and numerous musicologists to be much more distributed and not tamper-able by third parties.**

Keywords: **Copyright protection, anti-plagiarism, music industry, block chain, smart contract**

1 INTRODUCTION

Copyright is an instrument to signify someone truthfully own something. That something ranges from literature to artistic works. It may be a book, a movie, a music, etc. This means only the copyright holder may have the right to duplicate the thing being copyrighted (Intellectual Property Office of Singapore, 2021).

However, copyright does not prevent unauthorized duplication. It only signifies who are able to duplicate. That means, technically anyone was able to duplicate despite no legal backing. This leads to copyright infringement of a thing (Intellectual Property Office of Singapore, 2021).

Copyright infringements are a widespread occurrence. Lawsuits happening from time to time, costing insurmountable of monetary loss for people. This monetary loss may end someone's career (Copytrack, 2019).

Music is a complex problem. The structure of music may be similar, but the outputs are different. The opposite also reigns true. There are only few able to evaluate music copyright infringement. One such people are called musicologists. A musicologist helps determining whether music truthfully violates copyright or not (Berklee, 2022).

Not all the copyright lawsuits are true positive. Some of them are just false accusations that are considered to be positive. These false positives render everyone overly scared towards creating something. As someone may target them for copyright infringement lawsuit (Rolling Stone, 2020).

To help reducing false positive copyright lawsuits, the author proposes a framework model. The author will start by describing related works for this article, followed by the methodology used in creating the structure of the framework. Afterwards, the author will explain the process flow of the proposed framework. At the end, the author will do a summary conclusion and potential improvement for the proposed framework.

2 RELATED WORKS

2.1 MUSIC INFORMATION RETRIEVAL

Music Information Retrieval (MIR) is technologies that recently arise as a study on extracting information from music (Cheng, 2020). MIR includes, but not limited to Pitch and Melody, Music Rhythm, Music Harmony, and Singing Information Processing.

Pitch is a singular component inside a music that defined the sense of high sounds and low sounds. Sequences of pitches are called melody. Melodies also depict rhythm and harmony inside music.

There are multiple ways to detect pitch or melody. One way to extract pitch is (1) by analyzing it in-time using methods such as zero-crossing method, adaptive filter, maximum likelihood, etc; called Time-domain detection. (2) The other one is by analyzing the peak of the waveform periodically using algorithm like Short-time Fourier Transform (STFT); called Frequency domain detection. (3) The next one is by simulating human perception of pitch and using a replica model of an ear to receive the input of the music in order to process all of the information; called Auditory model detection (Cheng, 2020).

Extending the pitch detection into melody detection, there are quite a few method used to extract melodies, namely (1) pitch importance method, (2) singing voice separation method, and (3) data-driven note classification method (Cheng, 2020).

2.2 COPYRIGHT

Copyright is a term used to describe that a property is owned by a certain copyright holder and have the right to duplicate said property. For music producer, this is a way to generate revenue from musics produced. It also meant that a record label that sheltered these music producers would also gain monetary profit. But copyright infringements lawsuit happened everywhere and every time (Copytrack, 2019), either accidental or intentional. Music as a form of art is too complex to define into objective list, hence these lawsuits ended up being too subjective even when musicologists are involved (Rolling Stone, 2020). With the progression of the recent copyright infringement false positive cases, companies are working to solve copyright protection issues using blockchain and smart contract technologies (Kim & Kim, 2020).

2.3 BLOCKCHAIN

Blockchain is used to develop a decentralized distribution system. The blockchain treated its signed transaction as a block and secures it cryptographically. Each block may be linked with another block after being verified. After being linked, the previous block would be immutable and hard to modify (Qureshi & Jiménez, 2021; Shang & Sun, 2020).

With the rapid innovation blockchain brings to the table, technologies have become more decentralized (Qureshi & Jiménez, 2021). Various instruments have emerged from blockchain, such as smart contract.

3 METHODOLOGY

3.1 METADATA STORING IN BLOCKCHAIN

Storing music metadata in blockchain can be classified into two different cases. (1) Case of storing singular music; and (2) Case of storing multiple music at a time / *en bloc* (Album / EP).

Each blockchain will store relevant content and metadata of the music, such as music title, timestamp of copyright date, music artist, copyright organization, and hashed data of the music. Optional data such as MIDI sequence may as well be provided as an alternative to evaluate the melody correctly. This block will be associated with the creator of the music (Kim & Kim, 2020).

For storing a collection of music into the blockchain, the process is similar to the one that is working for singular music. The key difference of this case is that a Merkle Hash Tree will be generated based on list of hashes of each music inside the collection (Kim & Kim, 2020).

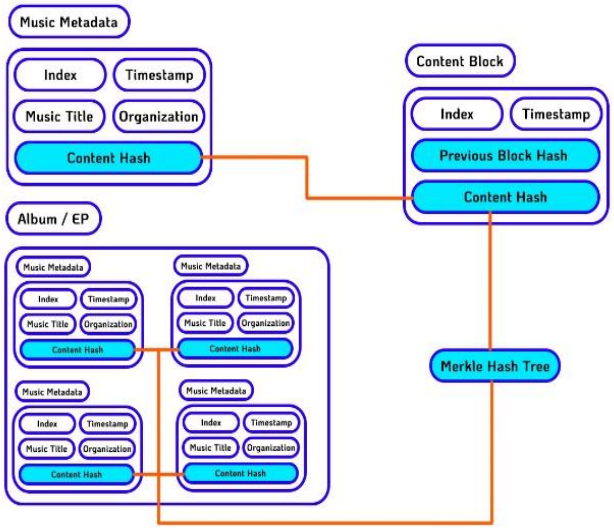

Figure 1: Storing hashes of music inside block for blockchain

3.2 Feature Extraction for Plagiarism Checker

Feature Extraction will retrieve data from each music. The model will create a window consisting of x number of measures. This window will extract a number of features from the selection of melody such as set of melodies, and the time signature. This window will slide into the next set of measure one by one. The window for the counter item will slide by one measure and will overlap each other, whilst the window for the copyrighted item will move without overlapping each of these measures (Park, Kim, & Shin, 2005).

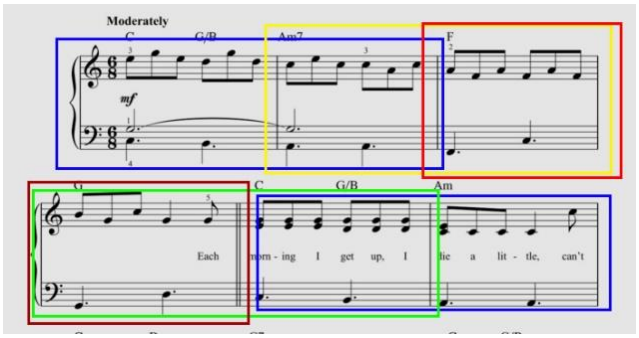

Figure 2: Window overlapping with another window for the suspect music

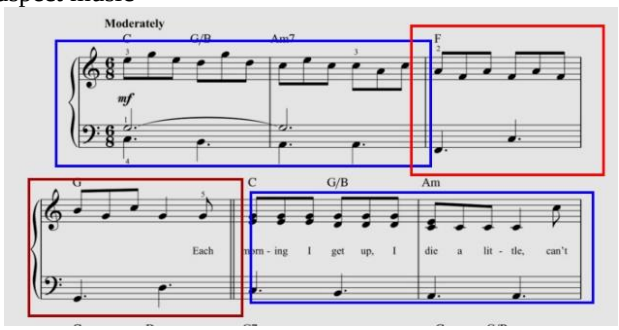

Figure 3: Window non-Overlapping for the copyrighted music

Extracted featured will undergo pre-processing of the pitches into a more meaningful information.

3.3 DEFINING THE MELODIES FOR PLAGIARISM CHECKER

These set of melodies in each window will be defined into its vectorial form. This vector form of the melodies will include the note, the length, the octave, and the interval of the note (Prisco, Malandrino, Zaccagnino, & Zaccagnino, 2017).

Most common number of octaves are usually 7 octaves, which is derived from a full-sized piano. Each of these octaves have 12 different notes, which consists of seven white keys and five black keys. They are C, C# (or Db), D, D# (or Eb), E, F, F# (or Gb), G, G# (or Ab), A, A# (or Bb) and B. The notes without sharps (#) are the white keys, and the notes with sharps (#) are the black keys. To denote a note inside an octave, write the note followed by the octave it is in. For example, C5 meant note C in octave 5, G#6 meant note G# in octave 6.

Melodies incorporate numerous of different notes in succession. The amount of leap done from one note to the next note is called Intervals. For example, the interval from E5 to F5 are 1 note, the interval from C6 to D6 are 2 notes (leaps over C#6).

To extract the interval of the note, the formula $interval(n_i, n_{i+1}) = n_{i+1} - n_i$ is used. n_i is the current note and n_{i+1} is the next note. If one of these notes are rest note, $interval(n_i, n_{i+1})$ would equals to 0. The absolute position of the pitch will be put into M vector, whilst the interval of the notes will be put into V_M vector (de Prisco et al., 2017; Prisco et al., 2017).

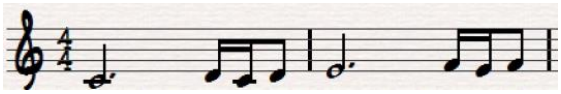

Figure 4: Example of a melody in music sheet

Take an example of "Will you be there" by Michael Jackson. The sequence of the melody are C4, D4, C4, D4, E4, F4, E4, F4 (by absolute pitch position it is 48, 50, 48, 50, 52, 54, 52, 54). This resulted with vector M = (48, 50, 48, 50, 52, 54, 52, 54) and V_M = [+2, -2. +2, +2, +2, -2, +2] (de Prisco et al., 2017; Prisco et al., 2017).

There is also the aspect of duration of the notes in mind that needed to be included as well. Common duration of notes are 4/4 (whole note), 2/4 (half note), ¼ (quarter note), 1/8 (eighth note), 1/16 (sixteenth note), and so on. With the example of the "Will you be there" by Michael Jackson, the durations would be [3/4, 1/16, 1/16, 1/8, ¾, 1/16, 1/16, 1/8] (de Prisco et al., 2017; Prisco et al., 2017).

3.4 PLAGIARISM CHECKER

Upon receiving the data of the sequences for the melody, it is processed using a fuzzy framework using Jaccard coefficient. Each window from both of the music will be evaluated one by one to deduce if a plagiarism is evident in one part of the music towards the copyrighted music. The fuzzy degree consisted of 0 and 1, with 0 as completely different while 1 as completely identical (Prisco et al., 2017).

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|}$$

Equation 1: Jaccard Similarity Coefficient

Assume A is vector of melody from copyrighted music, and B is vector of melody from suspect music. By applying Jaccard similarity coefficient for both of the vectors, we can get the amount of melodic similarity of the measures being calculated.

We have both melodic vector and rhythmic vector available from previous step. Both of these may be applied into the Jaccard similarity coefficient for better similarity coefficient. Assume Vm_A is melody vector of measure A, Vm_B is melody vector of measure B, Rm_A is rhythmic vector of measure A, and Rm_B is rhythmic vector of measure B. We have the following equation for both measuring melodic and rhythmic similarities. Note that these equations need to be applied for every available measure that are selected by the sliding windows in previous step.

$$J(Vm_A, Vm_B) = \frac{|Vm_A \cap Vm_B|}{|Vm_A \cup Vm_B|}$$

Equation 2: Implementation of Jaccard Similarity Coefficient for Melodic Similarity

$$J(Rm_A, Rm_B) = \frac{|Rm_A \cap Rm_B|}{|Rm_A \cup Rm_B|}$$

Equation 3: Implementation of Jaccard Similarity Coeffient for Rhythmic Similarity

After we found the result of the similarities, the average of the similarities needs to be calculated.

$$\bar{x} = \frac{\sum_a^b x}{n(x)}$$

Equation 4: Simple Average Equation

By Applying the Jaccard coefficient into the averaging formula, we have the following equation.

$$\overline{J(Vm_A, Vm_B)} = \frac{\sum_a^b J(Vm_A, Vm_B)}{b - a + 1}$$

Equation 5: Implementation of Simple Average Equation for Melodic Similarity Equation

Where "a" equal the minimum number of notes in Melody Vector A and B. "b" equal the least number of comparation from Melody Vector A and B.

This equation will be applied to the rhythmical similarities as well.

$$\overline{J(Rm_A, Rm_B)} = \frac{\sum_a^b J(Rm_A, Rm_B)}{b - a + 1}$$

Equation 6: Implementation of Simple Average Equation for Rhythmic Similarity Equation

After the average of the similarities calculated, a correlation analysis need to be performed. Within the measure of A and B, a sub-vector need to be acquired. Assume sub-vector of melodies A as "g". For each of the notes in the sub-vector there is a distance between each note that is denoted with $d(s_A, s_B)$. This is the equation of correlation between sub-vector melodies A with Measure B (De et al., 2015; Prisco et al., 2017).

$$\lambda(g_A, m_B) = 1 - \prod_{n(Vm_B)} (1 - d(s_A, s_B))$$

Equation 7: Fragment-to-melody correlation equation

After the correlation is determined, a Fuzzy number may be calculated. The following equation is the final equation of fuzzy for the plagiarism checker (Prisco et al., 2017).

$$F(m_A, m_B) = \frac{\sum_{n(Vm_A)} \lambda(g_A, m_B)}{n(Vm_A)}$$

Equation 8: Vector-based Fuzzy Implementation for melody similarity

This Plagiarism Checker will be stored as a Smart Contract. The reason is to prevent tampering from one party when a plagiarism case is surfacing. The characteristic of Smart Contract is that it is Immutable, which means no one can change it when it's already registered in blockchain.

4 PROPOSED FRAMEWORK

In the previous section, author has explained the approach of the copyright protection suite is taking. Starting from registering the music or the collection into blockchain and how it defines what is plagiarism in scope of music entertainment. In this section, the author will explain how it integrates with each other as a framework.

4.1 REGISTERING MUSIC INTO BLOCKCHAIN

The music producer will need to finish the music or collection that are about to be copyrighted. The producer then submits the music to copyright organization. The organization will then prepare the music to be registered into blockchain (Kim & Kim, 2020; Zeng, 2020).

First, the organization will create a content block based on the submitted music. Afterwards, the block will be verified first whether it is ready to be registered or not. As soon it's greenlit, block will be registered into blockchain (Kim & Kim, 2020; Zeng, 2020).

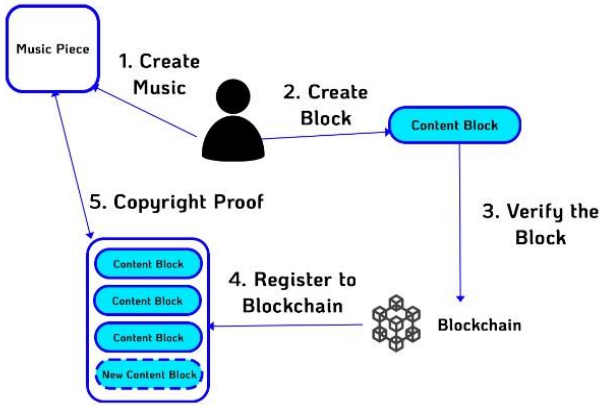

Figure 5: process flow of registering copyright in blockchain

4.2 COPYRIGHT INFRINGEMENT CASE

Whenever there is a case of copyright infringement report, blocks that are registered in the blockchain will be compared with the counter item using the Plagiarism Checker module. The Plagiarism Checker will determine if the counter item strikes a similar characteristic as the one in the blockchain.

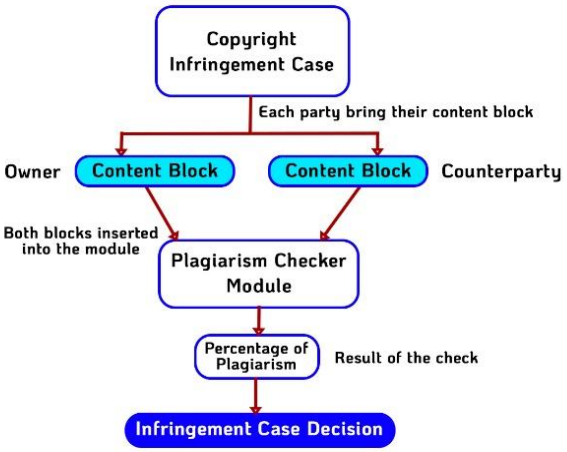

Figure 6: Process flow of a copyright infringement case in this framework

Both of the involved party will provide their content block of the registered music copyright. The blocks will be inserted into the Plagiarism Checker Module to check how similar the music from the suspect with the one that is defending their right. If the result is relatively high, then the suspect may be deemed guilty of copyright infringement (De et al., 2015; Park et al., 2005; Prisco et al., 2017).

5 CONCLUSION AND FUTURE WORK

False positive copyright lawsuit is a big problem not only in literature works, but also entertainment artistic works. One entertainment industry that are currently filled with lawsuits are music industry, either true positives or false positives. This affects people involved in the industry to be paranoid of a potential lawsuit flying towards them.

To prevent false positive copyright infringement cases towards music, author proposed implementation of copyright protection powered by blockchain. This framework consisted of a blockchain instance for storing the metadata of the music copyrighted within and a smart contract-powered music plagiarism detection module to. The blockchain ensures no human able to intervene the process of the copyright protection either in the metadata or the plagiarism detection module.

The author realizes this paper have potential for further studies. As of recently, Non-Fungible Token (NFT) has been in the hype train and is cooling down. Implementation of NFT as a blockchain instrument for storing the music as a whole and collaborating with smart contract of the plagiarism checker may be a potential future research.

Aside from NFT, author also realizes that the proposed framework is in its purely ideal form. It does not observe regulations for copyright and blockchain for each region and / or companies. For future study, this might be potential research

[1] REFERENCES

[1] Berklee. (2022). Musicologists.

[2] Cheng, Y. (2020). Music Information Retrieval Technology: Fusion of Music, Artificial Intelligence and Blockchain. Proceedings - 2020 3rd International Conference on Smart BlockChain, SmartBlock 2020, 143–146. Institute of Electrical and Electronics Engineers Inc. https:// doi.org/ 10.1109/ Smart Block 52591.2020.00033

[3] Copytrack. (2019). Copytrack Global Infringement Report 2019.

[4] de Prisco, R., Esposito, A., Lettieri, N., Malandrino, D., Pirozzi, D., Zaccagnino, G., & Zaccagnino, R. (2017). Music plagiarism at a glance: Metrics of similarity and visualizations. Proceedings - 2017 21st International Conference Information Visualisation, IV 2017, 410–415. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/iV.2017.49

[5] De, S., Roy, I., Prabhakar, T., Suneja, K., Chaudhuri, S., Singh, R., & Raj, B. (2015). Plagiarism Detection in Polyphonic Music using Monaural Signal Separation. Retrieved from http://arxiv.org/abs/1503.00022

[6] Intellectual Property Office of Singapore. (2021). Copyright Factsheet on Copyright Act 2021.

[7] Kim, A., & Kim, M. (2020). A Study on Blockchain-based Music Distribution Framework: Focusing on Copyright Protection.

[8] Park, J.-I., Kim, S.-W., & Shin, M. (2005). Music Plagiarism Detection Using Melody Databases. In LNAI (Vol. 3683).

[9] Prisco, R. de, Malandrino, D., Zaccagnino, G., & Zaccagnino, R. (2017). Fuzzy vectorial-based similarity detection of Music Plagiarism. 2017 IEEE International Conference on Fuzzy Systems.

[10] Qureshi, A., & Jiménez, D. M. (2021, January 1). Blockchain-based multimedia content protection: Review and open challenges. Applied Sciences (Switzerland), Vol. 11, pp. 1–24. MDPI AG. https:// doi.org/ 10.3390/app11010001

[11] Rolling Stone. (2020). How Music Copyright Lawsuits Are Scaring Away New Hits.

[12] Shang, M., & Sun, H. (2020). Study on the New Models of Music Industry in the Era of AI and Blockchain. Proceedings - 2020 3rd International Conference on Smart Block Chain, Smart Block 2020, 63–68. Institute of Electrical and Electronics Engineers Inc. https:// doi.org/ 10.1109/Smart Block52591.2020.00019

[13] Zeng, Y. (2020). Digital Music Resource Copyright Management Mechanism Based on Blockchain. Proceedings - 2020 3rd International Conference on Smart Block Chain, Smart Block 2020, 158–162. Institute of Electrical and Electronics Engineers Inc. https:// doi.org/ 10.1109/SmartBlock52591.2020.00036