Journal of Information & Optimization Sciences
ISSN 0252-2667 (Print), ISSN 2169-0103 (Online)
Vol. 46 (2025), No. 2, pp. 485–496
DOI : 10.47974/JIOS-1929

Securing digital certificates with perceptual hashing and JSON web tokens

R. Obulakonda Reddy *
D. Mahammad Rafi †
Department of Computer Science and Engineering (Cyber Security)
Institute of Aeronautical Engineering
Dundigal
Hyderabad 500043
Telangana
India

K. S. R. K. Sarma §
Department of Computer Science and Engineering (Data Science)
Vidya Jyothi Institute of Technology
Himayat Sagar Rd
Hyderabad 500075
Telangana
India

CH. N. Santhosh Kumar ‡
Department of Computer Science and Engineering
Anurag Engineering College
Ananthagiri Suryapet District
Telangana 508206
Telangana
India

* *E-mail:* rkondareddy@gmail.com (Corresponding Author)
† *E-mail:* mahammadrafi0780@gmail.com
§ *E-mail:* sarmaksrk@vjit.ac.in
‡ *E-mail:* santhosh.ph10@gmail.com

© Taru Publications

Elangovan Muniyandy [@]
Department of Biosciences
Saveetha School of Engineering
Saveetha Institute of Medical and Technical Sciences
Chennai 602105
Tamil Nadu
India

and

Applied Science Research Center
Applied Science Private University
Amman
Jordan

N. Sreevani [#]
Deepak Damarla [$]
Department of Computer Science and Engineering (Cyber Security)
Institute of Aeronautical Engineering
Dundigal
Hyderabad 500043
Telangana
India

Abstract

In medical imaging, insurance claims, intelligence systems, forensic investigation, criminal investigation, and journalism, digital pictures are playing a bigger role. Nevertheless, it is tough to determine if photos are real or not in the sense that freely available editing programs could be the only thing that makes changes to them. This study introduces an exclusive technique that utilizes JSON Web Tokens (JWT) and perceptual hashing (P-hashing) to find picture manipulation. The prescribed method passes the picture's P-hash value within the JWT's payload part, which is later encrypted with the ES512 algorithm and secp521r1 curve for securing and is included in the picture data. Using the transmitter's public key, the receiver decrypts the JWT and calculates a new P-hash on the other end to affirm the picture's authenticity. In case the new computed hash differs from the one in the JWT data, the picture has been changed. This method is rapid and can be used without block chain; hence, it brings a safe and offline way of verifying the picture's existence. It is also possible to match the calculated hash against the data of known hashes when it comes to higher security. Furthermore, the detailed analysis of the method considered in the study and its comparison with other picture duplication recognition technologies are also part of

[@] *E-mail:* **muniyandy.e@gmail.com**
[#] *E-mail:* **vanikondareddy@gmail.com**
[$] *E-mail:* **deepak@deepakd.com**

the paper, such as deep learning techniques. The results demonstrate that the combination of P-hashing and JWT is a perfect solution if the customer is tired of the well-known authentication technologies that are implemented in the digital photo and pdf industries.

Subject Classification: *Hashing, JSON, JWT, Forensic.*

Keywords: *Primary 93A00, Secondary 94A00.*

1. Introduction

The internet and mobile devices [1], technology itself, and the presence of high-speed wired and wireless Internet connections have made it possible to manage and distribute a large number of data files accurately and quickly across various devices and also on the Internet [1, 2, 3]. In addition to that, the other industry sectors, i.e. advanced medical, new auto insurance, legal, being computerized has been in trend in the past years. The usage of digital images has been on the rise parallel to the growing digital data. As AI gains popularity, the growth of picture editing is happening at the same time [3,4]. Because Photoshop and GIMP are the two tools that offer the most advanced editing features and they are also free and user-friendly. The scope of the impact of the photomanipulation tool through the use of validation technologies is ever increasing. Through digital signatures, files are in a way checked whether they are original or not or the rest of the information included in the file is changed. And so, in a digitally driven world, digital certificates and documents are supposed to be kept safe and intact.

Image manipulation detecting and the most usual approaches for it are watermarking and CNN, or convolutional neural network. The term "watermarking" refers to the practice of embedding a watermark into an image. An array of methods, mostly invisible, transform domain, non-blind, blind, and semi-blind watermarking, have been suggested to this end. However, each of these methods has its own special challenges and disadvantages as well. Invisible watermarking could be an example of the fact that the signal overhead only increasing and the deterioration of the image quality may be the effect [5, 6]. Because the user not only needs the original data to extract the watermark but the additional storage area for the carrier signal, non-blind watermarking becomes impractical [5,6]. Despite the fact that blind watermarking does not require the original data, it is often costly and can be cumbersome to set up [5,6]. Additionally, semi-blind watermarking brings more complexity to the system as it requires more storage space for keys or additional data [5,6]. Furthermore,

spatial domain watermarking is also a bit problematic in terms of its security, as it is very simple to establish and comprehend but at the same time it is very vulnerable to many types of attacks [5,6]. While they bring in the benefit of error resilience, the drawbacks of transform domain techniques are their computational resource intensiveness and processing speed that make efficiency and scalability seriously problematic [5, 6]. Such weaknesses suggest the inevitability of incorporating new more capable and drastic measures for the protection of image digital content and/or its certification.

Convolutional Neural Networks (CNNs) are the prime technique used for classifying different kinds of anomalies in images and also widely used [8]. CNNs seeing that they use much more processing resources than the traditional classification methods [8]. At the same time, they need a considerable additional amount of processing power when compared with watermarking techniques [7]. CNNs are of the view of the opinion that in addition to the data that is fed to them they also operate at a specific level of processing power. One of the major problems in image editing is the absence of full datasets. The neural network model can only detect if tampering is based on the data, they have been exposed to their performance may be hampered in the absence of the original and the altered picture data available for training [7,8,9]. There is the possibility that the system may ignore the anomaly when an image tampering method is different than the training dataset and ultimately will provide false negatives. Besides that, the processing capability of the device that the CNNs run on also has a dramatic impact in the sense of their capacity to detect manipulation [7]. High-speed computers can contribute to the process of detection while single-board and low-level devices, which are usually employed in embedded technologies, lack the processing power needed [7]. These devices cannot access Neural Processing Units (NPUs), which are essential for CNNs. This problem becomes even more evident in real-time applications and environments with limited computer resources [7-9].

CNNs are usually very effective to detect image alteration, nevertheless, still most of the digital certificates include the watermarks or multi-signature QR codes that can be verified after some time in order to ensure the image is authentic [10]. The tradition of adding extra information to the image during the production process, such as with watermarks, so that it can be retrieved and confirmed after some time is a cornerstone of the conventional image authentication methods. However, these methods also have their drawbacks in that they necessitate modifying the source

image and therefore expose the embedded data to deletion or modification. Not all CNNs have become practical in real settings since most of their intricate setups and model training requirements, which are only possible to be worked on the latest and best equipped systems. Moreover, the computational resources devoted to CNN model training have also restricted their broad implementation [10,11].

In this study, we aim to combine JSON Web Tokens and Perceptual Hashing (P-hashing) to develop an innovative technique for image editing detection in the context of digital certificates. This brand-new method intends to let us affirm the authenticity of these images and also cut down use of resources for CNNs by a big margin. Our method is an appropriate solution to the problem of checking if tampering was done in the case of digital certificates without the complicated and high computational costs associated with classical CNN-based techniques. The security afforded by JWT and the robustness of the P-hashing method applied to image comparison allows for the simplicity and efficiency of the task. Consequently, this paper is a paradigm shift in the light of forensic image and provides businesses with a more scalable and less intimidating solution to digital asset integrity protection than the ones previously available.

2. Methodology

JSON Web Tokens (JWTs) are very similar to PHP cookies and are usually the main tool for authenticating clients to a server [12,13]. Technically, a JWT consists of three parts: the header, claims, and hash. There are timestamps reflected in the token along with the time it was issued, the time it will expire, the perceptual hash (P-hash), and other related aspects [14] that are prescribed in the claims section. Moreover, this claims data is included in the payload, which is then accepted with SHA512 encryption and then signed with the ECDSA (Elliptic Curve Digital Signing Algorithm). The hash or the third part of the JWT is in fact the same data that has been hashed and then encrypted using an encryption algorithm. Typically, this is to guarantee the integrity of claims; the claims are hashed and the JWT's hash part is decrypted using the decryption key, exposing a hash. It is only after both of the hashes have been compared, and found to be matching that the JWT is verified and confirmed. With this operation, the digital certificates' uptime, integrity, and authenticity are ensured [12,13]. Figure 1 shows the JWT structure.

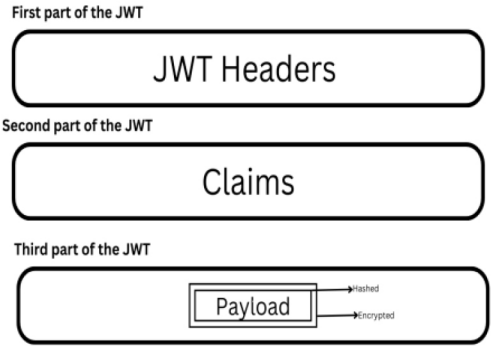

Figure 1

Structure of a JWT

Unlikely other hashing algorithms, which SHA512 is an example of, perceptual hashing (P-hash) is a special algorithm [14]. Perceptual hashing is designed to cope better with the particular difficulties occurring in certain data types, for example, images, which are not at all solved by traditional hash functions that, on the contrary, exhibit a very low degree of tolerance to the smallest data input changes, hence, they produce completely different hash values when the image is compressed. Formats like JPEG [14,15], that are based on lossy compression techniques, disable the use of the traditional hash algorithms for the images. The embedded compression mechanisms of these codecs may sometimes result in loss of some data. Such data loss might change the hash value significantly, even though the visual content of the image is almost unaltered. Thus, the introduction of changes that are not unique to the digital object may be incorrectly detected as tampering, although in reality they are just the remaining traces of the compression algorithm of the image [14, 15]. The high sensitivity of the algorithm may lead to vulnerabilities that come from the evolution of the image in the process of resaving and compression. In order that the integrity verification of the images and other similar data types be both more accurate and more reliable perceptual hashing creates hash values that are able to withstand benign transformations [14].

The SHA512 hashing technique is used to produce a safe and reliable hash, which is then incorporated into a JSON Web Token (JWT) once the image's P-hash has been determined. The Elliptic Curve Digital Signature Algorithm (ECDSA) using the secp521r1 curve is then used to encrypt this hash.

When contrasted with other encryption algorithms such as RSA [16] [25], elliptic curve cryptography (ECC) is faster. It is well demonstrated that ECC provides greater security for the same length of keys. To illustrate, an ECDSA with a 256-bit key is the same as an RSA algorithm with a 3072-bit key in terms of security [16, 19, 20]. 256 bits is a far cry easier and less resource-intensive than a 3072-bit key [16, 19, 20] to be spent Elliptic curve cryptography is a rather recent sphere of science. In the course of time, researchers have intensified their efforts to make the technology more efficient and secure, and they have also been doing their best to fix the security flaws [14]. Moreover, many experiments have come to the finding that ECC is superior to other hardware on both lower processing power and 32-bit platforms [17,18]. NIST-recommended secp521r1 curve is said to be the most efficient curve [24] and it is so because it utilizes Mersenne Prime and quicker residue multiplication modulo.[18]. Research contributions and developments of these techniques will continue, leading us to the conclusion that ECC will still be a strong and secure cryptographic technique in the face of ever-changing security threats [16].

The National Institute of Standards and Technology (NIST) has introduced the secp521r1 curve, which is a prime curve with a 521-bit prime number that provides strong cryptographic security [16 17, 18]. In fact, this is the first revision of the specification, as is indicated by the "r1" designation. The secp521r1 is among the most vulnerable curves expressing one among the most reliable alternatives demonstrating security criteria [16, 17]. The encrypted P-hash is now part of the JWT, which is combined with the image's information, such as EXIF data [21]. So, this, in fact, is an [26] element of the security whereby a person can later communicate to the server by looking at the p-hash of the image the embedded p-hash with or to the original hash the image's integrity. Figure 2 shows the process outline.

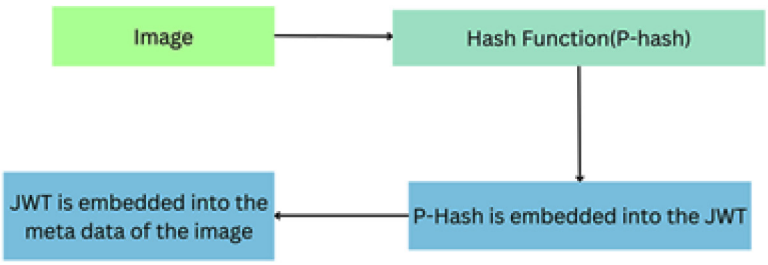

Figure 2

Outline of the process

The same procedure is applied to the photo to check if its tampered, the JWT after this is compared with the cached JWT of the image in the EXIF data. If they are the same, the photo has not been tampered with. In the event JWT comparing returns a mismatch, the image is deemed altered.

3. Challenges and Limitations

Still, aside from the implicit advantages of speed and precision in the detection of tampered pictures, this method also has some obvious drawbacks. One of the key shortcomings of this method is that it does not allow for localization. As a result, one cannot determine the exact area of a picture that has been manipulated. Moreover, this method still has room for improvement in the processing of EXIF data [22]. The procedure of verifying integrity is the weak point that may be taken advantage of by a skillful adversary who modifies or removes the EXIF data, even if the ordinary user may not so do. In addition, there are certain programs that tend to delete the EXIF data [22,23], thereby inadvertedly removing the embedded JWT, resulting in the picture manipulation's identification errors. Though, it is able to use the hash of the forgone file on the manipulated picture to confirm the integrity of the original image, the method requires access to the original photo so that it can be verified. Thus, it is only suitable for some use-case scenarios.

4. Future Directions

Much of the new AI technology now comes with a specific feature of tamper detection along with other AI capabilities for filtering materials and detecting fake images. The use of this tactic would give us an excellent means to improve tampering detection as it develops. Moreover, high levels of technology have been experiencing in the domains of picture fingerprinting and recognition with the aid of several types of EXIF data and hash-based approaches. Research has shown that, with some more iterations, the technique might be a proven tool to the ongoing effort to assure the integrity of the image and its validity.

5. Results

The testing was done on a dataset of four thousand images, the program was written in python and tests were performed on a laptop with intel i5-1155G7 and with 16GB of RAM, the entire calculation of the hashes

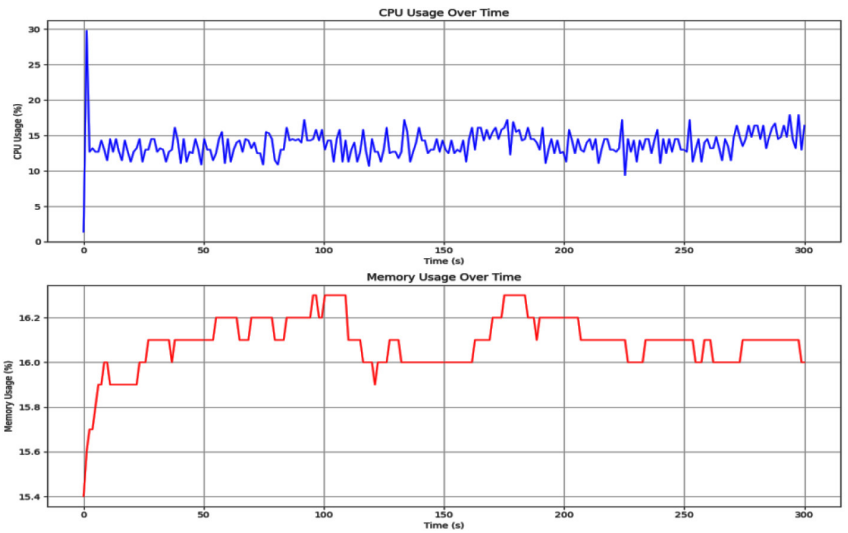

Figure 3

CPU and Memory Uses

and the JWT along with them in the EXIF data took about 5 minutes with an average CPU usage of about 14% and an average RAM usage of about 15%. The following graph shows the CPU and RAM data. Figure 3 shows the CPU and memory uses.

6. Conclusion

This study demonstrates a unique approach to pinpointing manipulated images. The method, which is more reliable and efficient than the common methods (such as convolutional neural networks (CNNs) and watermarking), is proposed. This technicality involves JSON Web Tokens (JWTs) that will act as a guarantee that the hash value will not change and that will thus improve the integrity verification process. The hash function is calculated by the system while at the same time the hash is generated using some components, for example, a constitutional hashing function, which was originally created for recording information on compressed photographic images and is very resistant to their decompression methods. The study elaborates on the said technologies' usage on low-resource devices to provide fast and accurate photo tampering detection. A total of 4,000 photographs require 400MB of Random Access Memory (RAM) for the perceptual hash (P-hash)

calculation. In about five minutes, the whole process was completed using a laptop with Intel i5-155G7 processor and 16GB of RAM which shows the method's efficacy in resource-constrained environments.

References

[1] W. Sun, J. Zhou, Y. Li, M. Cheung, and J. She, "Robust High-Capacity Watermarking over Online Social Network Shared Images," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 3, pp. 1208–1221, Mar. (2021), doi: 10.1109/TCSVT.2020.2998476.

[2] Z. Zhang and B. B. Gupta, "Social media security and trustworthiness: Overview and new direction," *Future Generation Computer Systems*, vol. 86, pp. 914–925, Sep. (2018), doi: 10.1016/j.future.2016.10.007.

[3] T. Hassan, A. Asaad, D. Ali, and S. Jassim, "Artificial Image Tampering Distorts Spatial Distribution of Texture Landmarks and Quality Characteristics," arXiv preprint arXiv:2208.02710 (2022).

[4] A. Ilavendhan and N. Vignesh, "A Comprehensive Analysis of Anti-Image Manipulation Techniques: Advancements and Challenges," 2023 2nd International Conference on Automation, Computing and Renewable Systems (ICACRS). *IEEE* (2023).

[5] D. K. Mahto and A. K. Singh, "A survey of color image watermarking: State-of-the-art and research directions," *Computers and Electrical Engineering*, vol. 93, Jul. (2021), doi: 10.1016/j.compeleceng.2021.107255.

[6] C. Kumar, A. K. Singh, and P. Kumar, "A recent survey on image watermarking techniques and its application in e-governance," *Multimedia Tools and Applications*, vol. 77, no. 3, pp. 3597–3622, Feb. (2018), doi: 10.1007/s11042-017-5222-8.

[7] S. Singhania, A. N. Arju, and R. Singh, "Image Tampering Detection Using Convolutional Neural Network," *International Journal of Synthetic Emotions*, vol. 10, no. 1, pp. 54–63, Sep. (2019), doi: 10.4018/ijse.2019010103.

[8] O. A. Montesinos López, A. Montesinos López, and J. Crossa, "Convolutional Neural Networks," in *Multivariate Statistical Machine Learning Methods for Genomic Prediction*, Cham: Springer International Publishing, pp. 533–577 (2022). doi: 10.1007/978-3-030-89010-0_13.

[9] F. Ghasemi, A. Mehridehnavi, A. Pérez-Garrido, and H. Pérez-Sánchez, "Neural network and deep-learning algorithms used in QSAR studies: merits and drawbacks," *Drug Discovery Today*, vol.

23, no. 10, Elsevier Ltd, pp. 1784–1790, Oct. 01 (2018). doi: 10.1016/j.drudis.2018.06.016.

[10] H. A. Ahmed and J. W. Jang, "Document certificate authentication system using digitally signed QR code tag," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Jan. (2018). doi: 10.1145/3164541.3164586.

[11] S. Xie, S. Zhong, R. Chen, and H.-Z. Tan, "Two-stage textured-patterns embedded QR codes for printed matter authentication," *Research Square* (2021).

[12] "JSON Web Token (JWT) Profile for OAuth 2.0 Client Authentication and Authorization Grants," Accessed: Aug. 25 (2024). [Online]. Available: https://doi.org/10.17487/RFC7523.

[13] P. Sharma, L. Raja, and S. Panwar, "Technological Use in Predicting Customer Churn for Long Term Sustainability and Viability Using Machine Learning Models," *Multidisciplinary Approaches for Sustainable Development*, CRC Press, pp. 224–228 (2024).

[14] H. Farid, "An Overview of Perceptual Hashing," *Journal of Online Trust and Safety*, vol. 1, no. 1, Oct. (2021), doi: 10.54501/jots.v1i1.24.

[15] T. Gillespie, "Content moderation, AI, and the question of scale," *Big Data and Society*, vol. 7, no. 2, SAGE Publications Ltd, Jul. 01 (2020), doi: 10.1177/2053951720943234.

[16] M. Al-Zubaidie, Z. Zhang, and J. Zhang, "Efficient and secure ECDSA algorithm and its applications: A survey," arXiv preprint arXiv:1902.10313 (2019).

[17] S. Ali and M. Cenk, "Faster residue multiplication modulo 521-bit mersenne prime and an application to ECC," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 65, no. 8, pp. 2477–2490, Aug. (2018), doi: 10.1109/TCSI.2018.2791285.

[18] A. Kumar, P. Dadheech, V. Singh, R. C. Poonia, and L. Raja, "An improved quantum key distribution protocol for verification," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 22, no. 4, pp. 491–498 (2019).

[19] M. Hübner, M. Gokhale, and R. Cumplido, "2015 International Conference on ReConFigurable Computing and FPGAs (ReConFig15): Cancun, Mexico, December 8–10, 2014," IEEE (2015).

[20] H. Cohen, A. Miyaji, and T. Ono, "Efficient elliptic curve exponentiation using mixed coordinates," *Advances in Cryptology—ASIAC-*

RYPT'98: International Conference on the Theory and Application of Cryptology and Information Security, Beijing, China, October 18–22, 1998, Springer Berlin Heidelberg (1998).

[21] S. M. N. Khan, J. M. Ghazali, L. Q. Zakaria, S. N. Ahmad, and K. A. Elias, "Various Image Classification Using Certain Exchangeable Image File Format (EXIF) Metadata of Images," *Malaysian Journal of Information and Communication Technology (MyJICT)*, pp. 1–12 (2018).

[22] B. Toevs, "Processing of metadata on multimedia using exiftool: a programming approach in python," in *Proceedings - 2015 Annual Global Online Conference on Information and Computer Technology, GOCICT 2015*, Institute of Electrical and Electronics Engineers Inc., pp. 26–30, Aug. (2016). doi: 10.1109/GOCICT.2015.14.

[23] M. Olvecky and M. Host'ovecky, "Digital image forensics using EXIF data of digital evidence," *ICETA 2021 - 19th IEEE International Conference on Emerging eLearning Technologies and Applications, Proceedings*, pp. 282–286 (2021), doi: 10.1109/ICETA54173.2021.9726649.

[24] K. R. Prasad, "Big Data Sentiment Analysis Using Distributed Computing Approach," in Bapi, R., Rao, K., Prasad, M. (eds), *First International Conference on Artificial Intelligence and Cognitive Computing*, Advances in Intelligent Systems and Computing, vol 815, Springer, Singapore, pp. 1–8 (2019). https://doi.org/10.1007/978-981-13-1580-0_66.

[25] K. Naveen Kumar, C. Krishna Mohan, and L. Reddy Cenkeramaddi, "The Impact of Adversarial Attacks on Federated Learning: A Survey," Accepted in *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2023), DOI: 10.1109/TPAMI.2023.3322785.

[26] C. N. Santhosh Kumar, A. Kiran, V. Kumar, S. Somnath, and S. Suneel, "Social media spammer identification using deep learning," in *MATEC Web of Conferences*, vol. 392, p. 01164, EDP Sciences (2024).

Received November, 2024