Technical Disclosure Commons

Defensive Publications Series

November 2022

Authenticated Attribution of Media Content Bound to Devices

Nic Watson

Chris Schneider

Follow this and additional works at: https://www.tdcommons.org/dpubs_series

Recommended Citation
Watson, Nic and Schneider, Chris, "Authenticated Attribution of Media Content Bound to Devices", Technical Disclosure Commons, (November 06, 2022)
https://www.tdcommons.org/dpubs_series/5486

This work is licensed under a Creative Commons Attribution 4.0 License.
This Article is brought to you for free and open access by Technical Disclosure Commons. It has been accepted for inclusion in Defensive Publications Series by an authorized administrator of Technical Disclosure Commons.

Authenticated Attribution of Media Content Bound to Devices

ABSTRACT

This disclosure describes techniques to authenticate that a media content item such as a photograph, a video, etc. is attributed to and originated by a particular user, based on the device on which the content was created. The content item is created on a device and associated with a digital signature using an attestation mechanism of the device such as trusted hardware or a sensor pattern that is unique to the device. The digital signature and content item are provided to a register or service for access by other users. When an accessing user views the content item, the content item is verified, via the digital signature, as having originated from the originating user's device. The accessing user's device displays the status of user attribution of the content item, indicating that the content item originates from a legitimate user and has not been forged, modified, or stolen. Described techniques are privacy-preserving and do not disclose the identity of the originating user.

KEYWORDS

- Content authentication
- Media attribution
- User attribution
- Media ownership
- Privacy
- Passive authentication
- Zero knowledge proof
- Trusted hardware
- DeepFake

BACKGROUND

User attribution of media content on the internet is an industry-wide problem. False attribution of media content to claimed originators represents an attack vector for malicious activity, e.g., to coerce individuals into particular actions (such as sending money) or represents lost revenue through the unauthorized reuse of the media content. Unauthorized modified and/or synthesized media content, such as DeepFakes, are also a rapidly growing problem. Users that download and display media content do not know, authoritatively, whether the content being displayed is attributed to a particular entity and has not been forged, stolen, or modified in an unauthorized manner.

DESCRIPTION

This disclosure describes techniques for authenticating and attributing origination and/or ownership of media content to users. The described techniques authenticate and attribute media to a specific device, and passively provides signals to users of media content in a way that lets them know, authoritatively, that the content being displayed is attributed to a particular entity. These techniques provide strong independent and external proof that the content item was generated on a particular device and by a particular user. This allows for attestation of provenance of that origination or ownership without additional burdens to the user and is provided via a privacy-preserving manner.

The described authoritative attribution of content provides value to content serving platforms and the described control mechanisms for content distribution provide value to content creators ("originating users"). The described techniques are also usable to detect unauthorized modified and/or synthesized media content such as DeepFakes.

The described techniques can be implemented on any suitable device or system, e.g., desktop or laptop computer, portable user device (e.g., a smartphone), server device(s), etc. The user is provided with options to enable or disable described techniques. The user can permit specific data or types of data (e.g., images, videos, etc.) to be processed and can deny processing of other data or types of data.

Creation of Authenticated Content Items

An example system that provides features described herein is shown in Fig. 1.

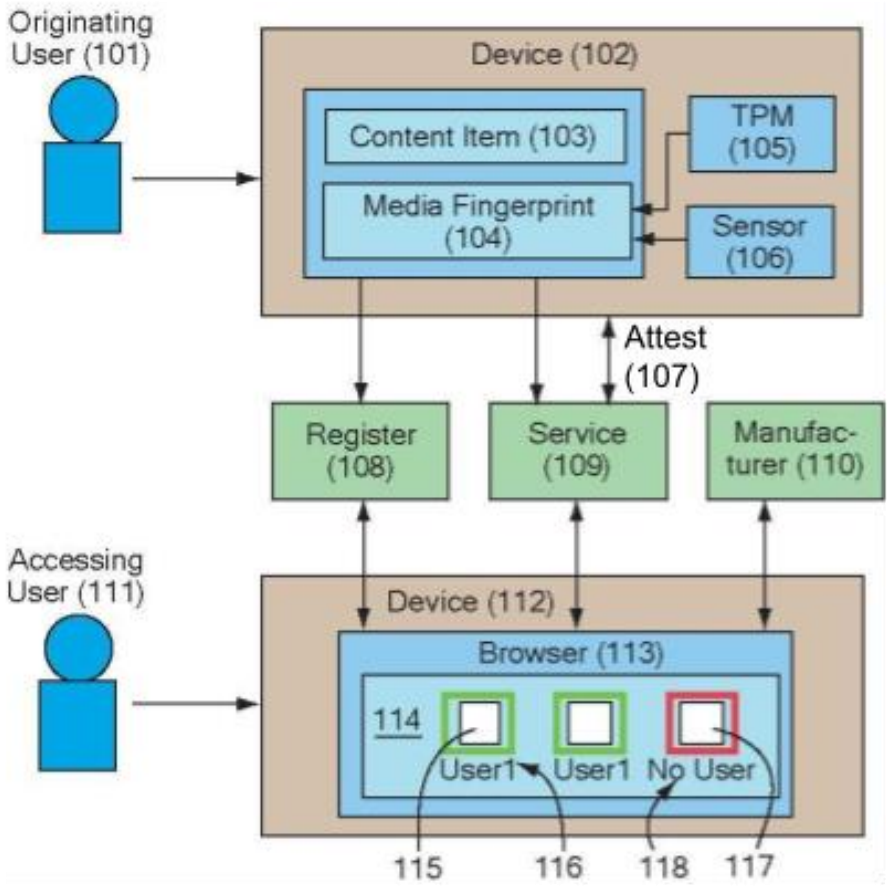

Fig. 1: Example system for authoritatively attributing media content to user device

As shown in Fig. 1, an originating user (101) operates a device (102) to create a media content item (103), such as an image (e.g., photo), video, or other content item. The device can be a camera, smartphone, computer, etc. A media fingerprint (104) is generated for the content item. The media fingerprint is bound to the specific device.

The media fingerprint includes metadata, such as a location of creation, name (e.g., username or account, obtained with user permission), camera settings when capturing an image, etc. The media fingerprint also includes a hash of the content item, and a digital signature based on the hash of the content item. The hash can be obtained by applying a suitable mathematical function to the content item and is unique to the media item.

In some cases, the digital signature can be a cryptographic digital signature that provides assertive provenance to indicate the source of the content item. In some cases, the digital signature can be provided by using a local attestation mechanism such as a Trusted Platform Module (TPM) (105), which may be hardware that is on the device and provides security-related functions using RSA (Rivest-Shamir-Adleman), ECDSA (Elliptic Curve Digital Signature Algorithm), or other cryptographic techniques. For example, the digital signature can be determined based on the hash of the content item and a private key stored in the TPM.

In some cases, e.g., if the device does not include a TPM, noise detected by a device sensor (106) can be used to provide the digital signature, as inferred provenance. For example, an optical CMOS sensor (complementary metal-oxide-semiconductor sensor) on the device that captures images may have random noise included in the captured images due to its manufacturing process. The noise is unique to that CMOS sensor and thus unique to the device. This noise is inherent in a content item that is a captured image and can be used to identify the device and as a form of digital signature.

Disseminating Content Items for Public Access

The content item and media fingerprint (e.g., both included in a content item file) can be disseminated in any of multiple ways. Referring to Fig. 1, in a first example, the originating user sends the content item and media fingerprint to a centralized authority that is a trusted verification service, referred to as a register (108). At the register, the content item is ready for dissemination across the internet as attested content. The register can be accessed by any accessing user's device to download the content item from the register. The privacy of the originating user is retained since the register does not reveal the identity of the originating user to devices accessing the register.

In a second example, the originating user uploads their signed content item to an internet service (109) that provides the content item for access to other users. For example, the service can be a social media service or other type of service/ website that provides a user account or other user identity under which the originating user can post content items.

When uploading the content item file to the service, the originating user attests (107) the device to a vendor, which can be the service or can be a third-party vendor that can provide verification to the service. In an example attestation process, the originating user attests to the vendor that he or she owns the device, the vendor sends a barcode or other coded representation to the originating user, the originating user displays the barcode on a different device, takes a photo of the bar code with the device, and sends the photo from the device to the vendor, and the vendor accepts that the originating user owns the device. The service then associates that device (and the associated digital signature) with the originating user's account on the service that is being used by the originating user in the attestation process.

In some cases, the originating user can choose to maintain his or her privacy in association with the content item, e.g., hide any association of the user's identity and account with the content item when it is provided for access by the service. If this privacy option is chosen, a privacy cryptographic protocol such as Private Set Membership (PSM) or other privacy protocol or technique, e.g., zero knowledge proofs such as Succinct Non-Interactive Arguments of Knowledge (SNARK), Bulletproof, etc. can be used to maintain privacy of the user's identity. In an example, the service hashes and stores the device signature and user account identification in a PSM table (accumulator) that is fully encrypted. When a request for verification is received, only a hash of the digital signature is received in the request, the hash is looked up in that table, and the service only informs the requestor whether a matching hash is present or not, without revealing to the requestor the stored hash or other information in the table.

Access of Content Items by Other Users

Referring again to Fig. 1, a different user ("accessing user") (111) can access the service via a device (112) and download the content item for display on the device. For example, the accessing user can use a browser (113) or other application program that retrieves the content item file (including content item and media fingerprint) and displays the content item on a display screen (114), e.g., as an image, icon, etc. The media fingerprint for the content item is parsed by the browser to verify the attribution of the content item based on instructions (e.g., as part of the browser or of a browser plugin) that enables the described verification. This parsing and verification process occurs in the background; the accessing user does not need to provide input or be aware of the verification process. The browser verifies the content item through a third-party entity (e.g., register 108, service 109, or manufacturer 110).

In some cases, if a media fingerprint is provided with the content item and the originating user sent the content item to a register (108) as described above, the browser attempts to verify the signature by sending a query to the same register, where the query includes the digital signature and hash from the media fingerprint. If the register has a matching hash associated with a matching signature, the register returns an answer to the browser that indicates that the content item is verified as being attributed to the originating user. If the hash is not present or metadata is missing in the media fingerprint (and the digital signature is present), then the content item can be hashed by the accessing user's device and sent to the register with the signature.

In some cases, e.g., if the content item was not provided to a register by the originating user, the verification process of the browser can include querying the manufacturer (110) of the originating user's device, e.g., over the internet, by sending the content item and its media signature to the manufacturer. The manufacturer can verify that the content item originated from (e.g., was created on) a particular device that has the device signature associated with the content item. For example, the manufacturer can verify, based on its data records, that the digital signature originated from a device having a particular public key from a TPM on that device, or from a device having a particular noise pattern from the CMOS sensor. The manufacturer provides a reply to the browser that indicates whether the content item is verified, e.g., whether or not the digital signature originated on the originating user's device.

In some cases, the query for verification can be sent by the browser to the service (109) providing access to the content item. The browser can send the media fingerprint of the content item to the service. If the service has previously verified the content item, such as via the attestation of the originating user to the service as described above, the service replies that the content item is verified for the originating user.

If the originating user requested privacy as described above, then the browser sends only the digital signature to the service in the request for verification, with no metadata. In this way, neither the verifying entity nor the accessing user get to know anything about the originating user's identity. When the signature associated with the content item is verified as described above, the service associates the content item with an identity (e.g., user account) of the originating user. If a media fingerprint is not provided with the content item, the browser can indicate to the accessing user that the content item is not verified and that it does not have an attributable originating user.

The browser displays the status of the verification, without requiring input from the user. A content item is displayed in the browser (or other application program) with a visual indicator, e.g., a visual overlay, that indicates the verified origination of the content item and whether the content item is authentic, e.g., has not been reattributed or modified since that origination. For example, a verified content item (115) is displayed with a visual indicator (116) that indicates that the originating user originated the content item. In an example, if permitted by the originating user, the originating user's account name can be displayed in or associated with the visual indicator.

If the originating user requested privacy in association with the content item, the visual indicator can indicate that the content item originates with an (unnamed) user and has not been modified since that origination. For example, the source of verification can be indicated by the visual indicator, e.g., the register or service that verified the origination of the content item. The user's identity (e.g., account) is not displayed in association with the content item.

In an example, the accessing user accesses the originating user's photo website that displays images of the user's photos. A visual indicator is displayed with each image and

indicates that the originating user originated these photos and that they have not been modified or stolen.

If the content item signature is not verified, or a digital signature is not provided with the content item, the content item (117) is displayed with a visual indicator (118) that indicates that there is no verified user attribution or origination for the content item.

Users are provided with options to grant permissions to and/or to disable described features entirely. The various features of the system are implemented only with user permission to access user information that serves as input to the system (e.g., user images or other content items, user context information, camera input, user's preferences, etc.). Users may be provided with controls allowing the user to make an election as to both if and when systems, programs or features described herein may enable collection of user information, and if the user is sent content or communications from a server. Certain techniques are not implemented if users deny permission. In addition, certain data may be treated in one or more ways before it is stored or used, so that personally identifiable information is removed. For example, a user's identity may be treated so that no personally identifiable information can be determined for the user. Thus, the user may have control over what information is collected about the user, how that information is used, and what information is provided to the user.

CONCLUSION

This disclosure describes techniques to authenticate that a media content item such as a photograph, a video, etc. is attributed to and originated by a particular user, based on the device on which the content was created. The content item is created on a device and associated with a digital signature using an attestation mechanism of the device such as trusted hardware or a sensor pattern that is unique to the device. The digital signature and content item are provided to

a register or service for access by other users. When an accessing user views the content item, the content item is verified, via the digital signature, as having originated from the originating user's device. The accessing user's device displays the status of user attribution of the content item, indicating that the content item originates from a legitimate user and has not been forged, modified, or stolen. Described techniques are privacy-preserving and do not disclose the identity of the originating user.

REFERENCES

1. Lee, Shinhaeng. "Content authentication system and method." U.S. Patent Application 11/259,000, filed April 26, 2007.