Poster: PhotoLock: Autonomous Privacy-preserving Photo Sharing in Online Social Networks

Edwin Yang, Song Fang, Christan Grant, and Le Gruenwald
University of Oklahoma
{edwiny, songf, cgrant, ggruenwald}@ou.edu

***Abstract*—Posting photos on online social networking services (SNSs) has become popular nowadays, which allows a user to share individual photos with others. However, such photos may contain sensitive private information (i.e., a child's face), which should only be shared within certain groups like close friends or families, rather than by everyone who gets access to the page that includes the photo. To address such a multilevel security problem, we propose a novel attribute-based encryption scheme to secure photo sharing on SNSs. Specifically, we first utilize deep learning methods to identify and categorize "private" information, and for each category of sensitive information, we then encrypt it using a pre-defined subset of the photo uploader's private attributes (e.g., locations, occupations, and hobbies) which are not publicly available on SNSs. Each subset of attributes builds a specific access rule. As a result, only when a visitor's own attributes match the rule, the corresponding encrypted information can be then recovered. Not only does the proposed scheme not interrupt information sharing among legitimate users, but also it prevents unauthorized users from obtaining the private information.**

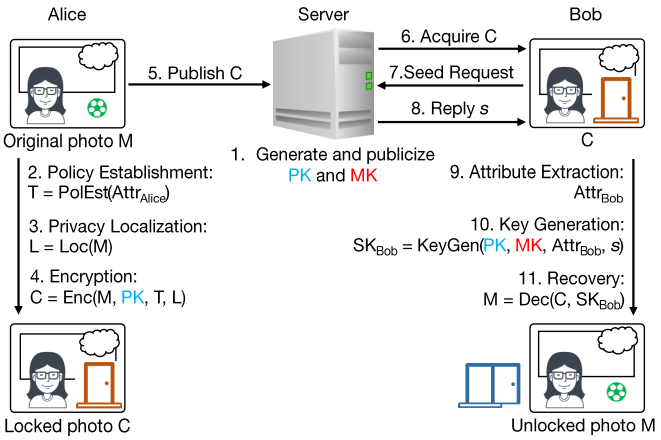


Fig. 1. *PhotoLock* scheme flow chart.

I. Introduction

With the growth and accessibility of technology and internet, the ease of posting and sharing photos on social networking services (SNSs) has increased exponentially. However, many users may unconsciously upload a photo that contains sensitive information in regard to themselves or someone else, such as baby faces, phone numbers, house locations and others. Certain private information should only be shared with certain groups of people. For example, some young parents may share pictures of their kids on SNSs and thus friends and family who do not live nearby can see them. Meanwhile, if such private information is disclosed to adversaries, it may cause serious consequences, including identify theft and digital kidnapping.

There are extensive research aiming to achieve privacy-preserving photo sharing on SNSs. Once sensitive information is detected on a photo before it is posted, existing work usually blur [4] or encrypt [5], [6], [3], [7] the sensitive part. Though both may stop adversaries from capturing the privacy when the employed blurring technique is irreversible and the key used for encryption is strong enough. Each has specific shortcomings: the former scheme also disables the intended receivers to view the content, causing the goal of sharing failed; the latter scheme requires the uploader to pre-share different keys with different groups who are allowed to view the corresponding private information, which is quite inflexible and causes a huge burden for the user. In this work, we propose a novel photo sharing scheme which enables the user automatically encrypts the detected private information leveraging attribute-based encryption (ABE) [1], [2], so that legitimate users can recover the private information with their attributes. The match level between attribute sets of the photo uploader and the viewer implements a natural fine-grained access control policy.

II. System Design

When a user uploads a photo, the proposed scheme automatically detects sensitive regions with a customized deep-learning based image analysis scheme. For each region, the user then adds different "locks" on them which are designed by leveraging the uploader's pre-defined attributes. The photo with locks will then be uploaded to the server. When another user visits the uploader's web page, aiming to view the picture, a key would be crafted based on the user's attributes in order to unlock the sensitive regions. Only the attributes of the user and the uploader match, the user is able to succeed in obtaining the private information.

We utilize the technique of attribute-based encryption to generate such locks and keys. The encryption and decryption tasks are performed at client sides. To prevent an adversary from launching exhaustive attack (i.e., brute-force attack) to decrypt the locked regions, we involve the server in the secret key generation procedure, so that only limited number of decryption trials are accepted. Also, we explore to utilize

Network and Distributed Systems Security (NDSS) Symposium 2020
23-26 February 2020, San Diego, CA, USA
ISBN 1-891562-61-4
https://dx.doi.org/10.14722/ndss.2020.23xxx
www.ndss-symposium.org

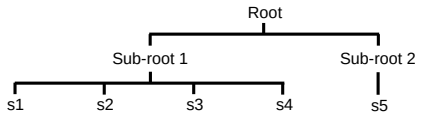


Fig. 2. Tree structure of a policy: `(s1 or s2 or s3 or s4) and s5`.

user-specific private attribute information to build a multi-level access control policy, with which, private information within an uploaded photo will be protected from unauthorized viewers, who do not pass the access control check.

In the following discussion, we illustrate the proposed photo sharing scheme using an example with one server and two legitimate users, named Alice and Bob. There are four major phases to securely share a photo between the two legitimate users, i.e., preparation, photo locking at the uploader side, key crafting at viewer side, and photo unlocking. Figure 1 demonstrates the overall structure of such an example.

A. Preparation

First of all, the server generates a public key PK and a master key MK. Suppose Alice has an original photo M, which is taken during a holiday trip with Bob, and she wants to share it with Bob. Before Alice uploads M to a SNS server (e.g., Facebook), she first determines different levels of access control policies using her private attribute sets (e.g., name, favorite sports, address). Let T_i represents such an access control policy, where i is the index of the determined security levels. T_i can be then denoted with a string including one or more conditional statements.

B. Photo Locking

Each access control policy T_i is parsed and stored with a tree structure. Figure 2 depicts an example of such a policy tree which is generated according to a user attributes based string, i.e., `(s1 or s2 or s3 or s4) and s5`, where s_j ($j \in \{1, 2, \cdots, 5\}$) indicates the j-th condition. During the photo locking phase, T_i is converted into a polynomial using the public key PK and the selected hash function. During the conversion of T_i, a corresponding sub-key k_i is derived from the polynomial. Meanwhile, private parts with different sensitivities in M are identified with a customized deep-learning based privacy localization algorithm. With the derived locations and sensitivities, different access control policies are utilized to encrypt corresponding sensitive parts. As a result, the user derives a polynomial part and an encrypted part from each of detected sensitive area in M. The user then generates and publishes the encrypted photo C, where the set of the polynomials are embedded.

C. Key Crafting

Suppose Bob and Alice are families, and he should be able to view the true content that Alice posts. When Bob visits Alice's web page which includes the photo, the server first displays the photo with private regions being encrypted (i.e., locked). In order to recover the original picture, Bob then generates a secret key SK_{Bob} using his own attributes $Attr_{Bob}$, along with the pubic PK, MK and the seed s obtained from the server. Note that for each decryption, s should be re-issued by the server, which is used for mitigating the brute-force attack that an illegal user may launch as the server can decline excessive seed requests from a same user within a short time.

D. Photo Unlocking

After Bob derives a secret key SK_{Bob}, he attempts to decrypt the locked photo C. In this phase, as Bob's private attributes $Attr_{Bob}$ satisfy the access control policy T_i specified by Alice (i.e., the access control policy for family members), Bob successfully obtains each sub-key k_i with the polynomials extracted from the photo C as well as his attributes $Attr_{Bob}$. Since the sensitive part in the photo C is encrypted with the corresponding sub-keys, Bob can then successfully reveal the original photo M.

Generally, any user who is able get access to Alice's web page including the photo may send a seed request for recovering locked regions. Suppose an adversary Trudy is a strange to Alice, aiming to gather private information about Alice for malicious goals. Similar with Bob, Trudy is also able to derives her corresponding secret key SK_{Trudy} with her own attributes $Attr_{Trudy}$. However, as $Attr_{Trudy}$ may not pass any access control policies determined by Alice, she would thus fail to recover the original photo M.

III. Conclusion

We have verified the feasibility of a novel autonomous privacy-preserving photo sharing scheme on SNSs leveraging attribute-based encryption, which enables a user to upload a photo without worrying that the private sensitive information on photo will be disclosed to unauthorized users. Continuing this work, we will further improve our scheme by considering more security levels of different groups as well as more user-related attributes for providing finer-grain access control.

References

[1] R. Baden, A. Bender, N. Spring, B. Bhattacharjee, and D. Starin, "Persona: an online social network with user-defined privacy," in *Proceedings of the ACM SIGCOMM 2009 conference on Data communication*, 2009, pp. 135–146.

[2] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-policy attribute-based encryption," in *2007 IEEE symposium on security and privacy (SP'07)*. IEEE, 2007, pp. 321–334.

[3] J. He, B. Liu, D. Kong, X. Bao, N. Wang, H. Jin, and G. Kesidis, "Puppies: Transformation-supported personalized privacy preserving partial image sharing," in *2016 46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*. IEEE, 2016, pp. 359–370.

[4] F. Li, Z. Sun, A. Li, B. Niu, H. Li, and G. Cao, "Hideme: Privacy-preserving photo sharing on social networks," in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*. IEEE, 2019, pp. 154–162.

[5] Ra, Moo-Ryong, R. Govindan, and A. Ortega, "P3: Toward privacy-preserving photo sharing," in *Presented as part of the 10th USENIX Symposium on Networked Systems Design and Implementation (NSDI 13)*, 2013, pp. 515–528.

[6] M. Tierney, I. Spiro, C. Bregler, and L. Subramanian, "Cryptagram: Photo privacy for online social media," in *Proceedings of the first ACM conference on Online social networks*. ACM, 2013, pp. 75–88.

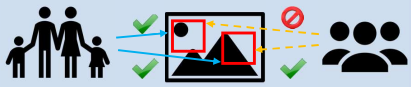
[7] L. Yuan, P. Korshunov, and T. Ebrahimi, "Privacy-preserving photo sharing based on a secure jpeg," in *2015 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2015, pp. 185–190.

Poster: PhotoLock: Autonomous Privacy-preserving Photo Sharing in Online Social Networks

Edwin Yang, Song Fang, Christan Grant and Le Gruenwald
University of Oklahoma
{edwiny, songf, cgrant, ggruenwald}@ou.edu

Motivation

- Photo sharing over social networking services becomes increasingly popular.
- An uploaded photo may contain multiple categories of sensitive information, each requiring different protection mechanisms.
- Existing privacy-preserving photo sharing techniques (e.g., encryption [1], blurring [2]) do not distinguish different levels of information sensitivity in a photo, and thus can not support multi-level access control.

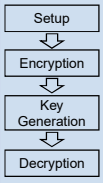


Multi-level access control for different groups.

Objectives

- Develop a deep-learning based privacy localization method to localize and categorize private information in a photo.
- Apply a customized attribute-based encryption scheme to automatically generate public keys and corresponding secret keys.
- Achieve multi-level access control by automatically establishing access control policies based on the photo-uploader's private information.

Attribute-Based Encryption Scheme

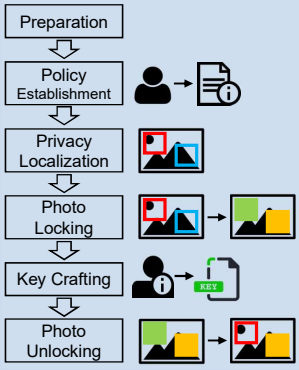


1. **Setup** (Server side): Generate a public key and a master key.
2. **Encryption** (Uploader side): Encrypt plain-text with the public key and the uploader's attributes.
3. **Key Generation** (Viewer side): Derive a secret key using the master key and the viewer's attributes.
4. **Decryption** (Viewer side): Decrypt cipher-text using the secret key.

Future Work

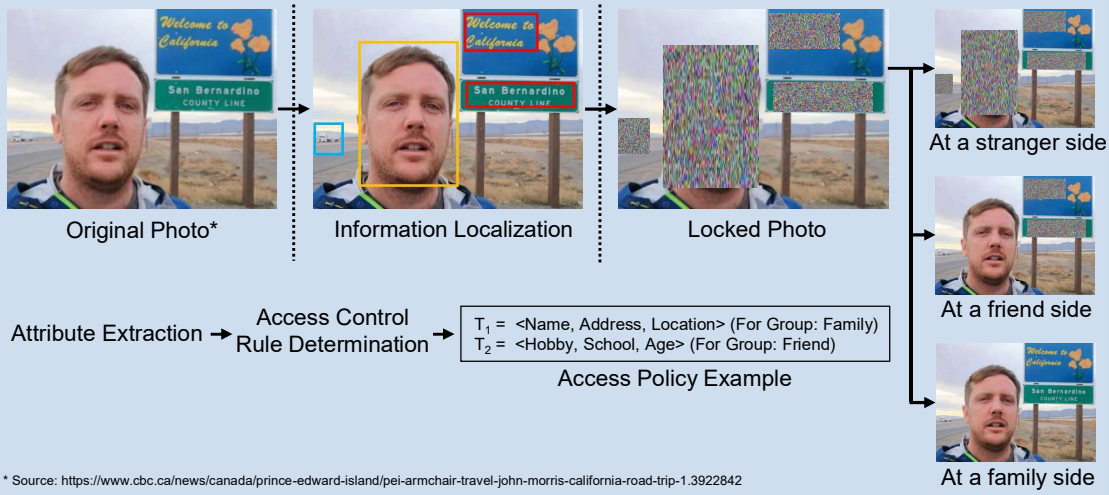
- Investigate commonly shared private information between different human relationships (e.g., families, close friends).
- Develop a deep-learning based privacy localization algorithm.
- Implement a fully automated access control policy generation scheme.
- Perform extensive user studies to evaluate the performance of the proposed technique.

System Design



1. **Preparation:** Initializes a public key and a master key on server side.
2. **Policy Establishment:** Extract access control policies from the uploader's personal information.
3. **Privacy Classification:** Localize and categorize sensitive data from an uploaded photo.
4. **Photo Locking:** Encrypt the identified sensitive regions in the photo with the established access control policies.
5. **Key Crafting:** Generate the secret key leveraging the viewer's personal information.
6. **Photo Unlocking:** Decrypt each encrypted region of the photo using the secret key.

Demonstration



* Source: https://www.cbc.ca/news/canada/prince-edward-island/pei-armchair-travel-john-morris-california-road-trip-1.3922842

References

[1] M. Tierney, I. Spiro, C. Bregler, and L. Subramanian, "Cryptagram: Photo privacy for online social media," in *Proceedings of the first ACM conference on Online social networks*. ACM, 2013, pp. 75–88.

[2] F. Li, Z. Sun, A. Li, B. Niu, H. Li, and G. Cao, "HideMe: Privacy-preserving photo sharing on social networks," in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*. IEEE, 2019, pp. 154–162.

[3] R. Baden, A. Bender, N. Spring, B. Bhattacharjee, and D. Starin, "Persona: an online social network with user-defined privacy," in *Proceedings of the ACM SIGCOMM 2009 conference on Data communication*, 2009, pp. 135–146.

[4] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-policy attribute-based encryption," in *2007 IEEE symposium on security and privacy (SP'07)*. IEEE, 2007, pp. 321–334.