Face/Off: Preventing Privacy Leakage From Photos in Social Networks

Panagiotis Ilia[*]
FORTH, Greece
pilia@ics.forth.gr

Iasonas Polakis
Columbia University, USA
polakis@cs.columbia.edu

Elias Athanasopoulos
FORTH, Greece
elathan@ics.forth.gr

Federico Maggi
Politecnico di Milano, Italy
federico.maggi@polimi.it

Sotiris Ioannidis
FORTH, Greece
sotiris@ics.forth.gr

ABSTRACT

The capabilities of modern devices, coupled with the almost ubiquitous availability of Internet connectivity, have resulted in photos being shared online at an unprecedented scale. This is further amplified by the popularity of social networks and the immediacy they offer in content sharing. Existing access control mechanisms are too coarse-grained to handle cases of *conflicting interests* between the users associated with a photo; stories of embarrassing or inappropriate photos being widely accessible have become quite common.

In this paper, we propose to rethink access control when applied to photos, in a way that allows us to effectively prevent unwanted individuals from recognizing users in a photo. The core concept behind our approach is to change the granularity of access control from the level of the photo to that of a user's personally identifiable information (PII). In this work, we focus on the face as the PII. When another user attempts to access a photo, the system determines which faces the user does not have the permission to view, and presents the photo with the restricted faces blurred out. Our system takes advantage of the existing face recognition functionality of social networks, and can interoperate with the current photo-level access control mechanisms. We implement a proof-of-concept application for Facebook, and demonstrate that the performance overhead of our approach is minimal. We also conduct a user study to evaluate the privacy offered by our approach, and find that it effectively prevents users from identifying their contacts in 87.35% of the restricted photos. Finally, our study reveals the misconceptions about the privacy offered by existing mechanisms, and demonstrates that users are positive towards the adoption of an intuitive, straightforward access control mechanism that allows them to manage the visibility of their face in published photos.

Categories and Subject Descriptors

K.6.5 [**Management of Computing and Information Systems**]: Security and Protection

General Terms

Security, Privacy, Human Factors

Keywords

Access Control; Online Social Networks; Shared Photos; Photo Tagging;

1. INTRODUCTION

Online social networks (OSNs) have radically transformed the online behavior and activities of users. Unfortunately, such services have also introduced a number of privacy issues, which have caught the attention of both the research community and data protection agencies (e.g., [9, 31]). As the use of these services spans across multiple facets of daily life, users may face dire consequences when their personal and professional life can affect each other via OSNs. Many articles have reported incidents of users being fired because of sensitive photos which they considered to be private, while in actuality they were not (e.g., [7, 8]).

The implications of such privacy issues becomes alarming when considering the scale of adoption of these services. Apart from surpassing 1.49 billion monthly active users, with an average of 968 million daily users [3], Facebook has also become the most time-consuming online user activity [10], as well as the de-facto platform for sharing photos with over 350 million uploaded daily [6]. Accordingly, many companies regularly check up job applicants online during the hiring process. A recent study by Acquisti and Fong [16] revealed that they may also use what they find to discriminate against applicants. A Microsoft survey [15] found that 70% of recruiters in the US have rejected candidates due to information, including photos, they found online.

In certain cases, some users may not be concerned about privacy or may be unaware of the implications of their actions. Users may also not fully understand complex access control mechanisms, and disclose private information without hesitation, oblivious to the true visibility of the uploaded content. In an attempt to make users more aware of who can view their posts, Facebook recently altered the privacy

[*]Panagiotis Ilia is also with the University of Crete.

CCS'15, October 12–16, 2015, Denver, Colorado, USA.
Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM 978-1-4503-3832-5/15/10 ...$15.00.
DOI: http://dx.doi.org/10.1145/2810103.2813603.

selector for status updates, to explicitly describe the potential audience [5]. According to reports [13], Facebook is also building a system that will automatically identify embarrassing photos being uploaded (e.g., where the user is drunk), and warn the user if they are being shared publicly. However, while such an approach may prevent certain users from uploading photos (of themselves), it cannot prevent other users that may have "malicious" intent or lack better judgement. As users exhibit fundamentally different behaviors regarding how they disclose information [30], they may have different perceptions regarding what content is sensitive. Thus, in many cases the problem arises from content that is shared among multiple users (i.e., a photo that depicts several individuals). As such, these measures can only handle a subset of the actual problem.

In this work, we highlight that the essence of the problem is that existing mechanisms for defining access to photos in OSNs, cannot effectively handle cases where the interested parties have conflicting settings. First, the photo uploader is considered the owner of the photo and is granted full rights, whereas the people appearing in the photo are not considered co-owners and are not granted any rights. On top of this basic coarse-grained approach, OSN providers implement additional policies, some of which can significantly complicate issues. For example, the uploader can restrict the photo's visibility for the tagged users, and the maximum allowed visibility for them extends to their immediate contacts (i.e., a tagged user cannot set the visibility to include any users apart from his immediate social circle). Second, the photo uploader is not required to request the permission of the people present in a photo before publishing it, and may even ignore their requests to remove it. Furthermore, any users that are tagged affect the visibility of the photo, as the photo will be viewable by all their contacts (default privacy setting). Thus, even when the users tagged in the photo have restricted its visibility, if the uploader has not restricted access the photo will be publicly available, something which the remaining users will not even be aware of. In general, these situations can be characterized as cases of *conflicts of interest*, where the will of the content publisher goes against the will of the depicted users, or the privacy settings of one user override those of another. Note that even though the access control mechanisms may vary across OSNs, conflicts of interest are a general issue, as they arise from the *content* of the photos.

Previous work has proposed frameworks for integrating access control policies of collaborating parties [35], and mechanisms that allow the users to contribute to the specification of a collective policy [26,39]. However, such approaches only solve the problem partially, as they handle visibility at a photo-level granularity. In other words, current solutions are too coarse-grained for accommodating the privacy settings of all the associated users. In such cases, a user has to accept and follow the access control decision of the majority, even if his privacy concerns are not satisfied.

In this paper, we propose an approach that can effectively handle these conflicts by changing the granularity of the access control mechanism to that of the users' faces. This enables an OSN to express and enforce every user's privacy setting within an image; none of the users' settings are overridden no matter how restrictive or permissive they may be. In a nutshell, our approach employs face recognition to automatically identify the users depicted within a photo; subsequently, the visibility of each user's face is automatically restricted based on the privacy settings of the specific user and not the content publisher. The result of this phase is a "processed" photo that can be rendered selectively according to who is viewing it. Thus, when a photo is accessed, the system will automatically blur the faces of the users that have restricted access. We propose a simple technique to encode the pre-processed photos, so as to avoid the overhead of blurring them during the rendering phase.

We conduct a case study on over 4 million photos collected from 128 participants and their social circles, and explore the characteristics of their social graphs and their tagging behavior. We then quantify the privacy risks that users are exposed to, due to existing access control mechanisms.

To evaluate the feasibility of our approach being deployed at a large scale, we measure the overhead incurred by our proof-of-concept implementation. As popular OSNs already process photos with face recognition software, the overhead of our approach lies in retrieving the permissions of every user, enforcing access control, and processing the photo "on the fly". On average, our system requires only 0.05 seconds per photo, when running on a commodity machine.

To evaluate the effectiveness of our approach on preserving user privacy, we conduct an experiment with 34 participants. Each participant is shown a set of photos of their contacts, with the face of one user "hidden" in each photo, and is requested to identify those users. In 87.35% of the cases, the participants fail to identify their contacts, demonstrating that our approach can significantly improve user privacy.

We also interviewed 52 participants, to understand how users perceive existing access control mechanisms, and their opinion on the potential adoption of our approach by OSNs. Apart from the lack of understanding of existing access control settings due to their complexity, we find that most users are positive towards a simpler, yet, more privacy-preserving approach. After being informed about the conflicts of interest that arise in shared photos, 77% of them express positive opinions regarding the adoption of our approach, and 19.2% remain neutral. Only 3.8% are negative, and state a lack of concern for the privacy implications of content sharing.

Overall, the main contributions of this work are:

- We design an innovative fine-grained access control mechanism for photo-sharing services that enforces the visibility of each user's face based on their respective access control lists. Our approach effectively handles all the cases of *conflicts of interest* between the privacy settings of users.
- We build a proof-of-concept application that demonstrates the feasibility and applicability of our approach within the infrastructure of a real-world OSN. Our experiments show that performance overhead is small compared to existing processing of photos by OSNs, rendering the adoption of our approach suitable even at such a scale.
- Our first user study provides insights into the tagging behavior of users, and reveals the risk users face due to conflicting privacy settings on shared photos. Based on the collected data, we assess user tagging behavior, and quantify the risk presented in certain photo-sharing scenarios.
- A second user study demonstrates the effectiveness of our approach in hiding users' identities from their contacts. We also highlight the counter-intuitive approach of existing access control mechanisms, and the eagerness of users to adopt a mechanism that allows them to manage the visibility of their faces.

2. PHOTO-BASED PRIVACY LEAKAGE

Earlier work has reported that users are concerned about their privacy and tend to avoid publishing photos or any other private information publicly [43]. Furthermore, according to a survey by Besmer et al. [19], explicit requests by users for deletion of photos, or users un-tagging themselves, are complicated issues. These can lead to social tension and are a source of anxiety for users, who may abstain from such actions to ensure the stability of their social relations [42]. Also, the uploader may lack the incentives to actually fulfill the user's request and remove a photo. Thus, it is apparent that users are restricted by coarse-grained access control models regarding shared photos, and in many cases sacrifice their privacy in favor of not agitating social relationships. Moreover, the wide visibility of photos can also expose users to inference attacks [37], or be leveraged by attackers for bypassing account protection mechanisms [33, 34]. In the following, we present an example that illustrates users' privacy risk and we determine certain privacy leakage scenarios.

2.1 Privacy Leakage Example

To provide a visual illustration of the extent of the risk presented to users due to the existing access control mechanism, we present an example. We recreate a segment of the actual social graph collected in our user study (Section 3), and extend it by crawling publicly available data from Facebook. Specifically, we select four out of the 128 users that participated in our study that are connected (e.g., friends), and we re-create their social graph. We also use publicly available data regarding the users that were two hops away from them within the Facebook graph (i.e., friends of friends). This results in a social graph that contains 55,109 users. Note that, since certain users and friendships might not be publicly viewable and, thus, not collected by our crawler, these numbers are conservative estimates (i.e., lower bounds).

We consider an example case where a photo depicting the four users is uploaded, and calculate the privacy risk for one of those users (i.e., Bob) depending on who the uploader is. We quantify the risk as the number of people (i.e., nodes in the graph) that are not connected to Bob, but can access it in spite of Bob's settings. Recall that the uploader controls the photo's general visibility setting and also controls which of the tagged users' friends can view the photo. For simplicity we apply the default setting for each tagged users.

Figure 1a presents the ideal case, where Bob is tagged in a photo and only his 339 friends have access. In Figures 1c to 1e we illustrate which users can access the photo in different cases, and if they have been granted access by the user of interest (Bob), or by others. In these cases, the uploader allows users two hops away within the social graph to access the photo, i.e., the visibility setting is set to "friends of friends". For the remaining tagged users the setting is set to "friends only". As can be seen, depending on the position of the uploader in the social graph, the combined effect of (i) the coarse granularity of access control and (ii) multiple users appearing in the photo, the extent of privacy leakage covers up to 86.78% of the social graph (47,829 users).

This example highlights the extent of the problem, as the current mechanism allows users to access a photo that a user might want to restrict, even if the uploader does not set the privacy setting to "public". While these numbers will vary depending on the structure of each user's social graph, they are indicative of the risk-propagation effect.

2.2 Privacy Leakage Scenarios

Here we present certain scenarios that highlight the privacy implications that arise in everyday situations, due to the current access control mechanisms for managing the visibility of photos published in OSNs.

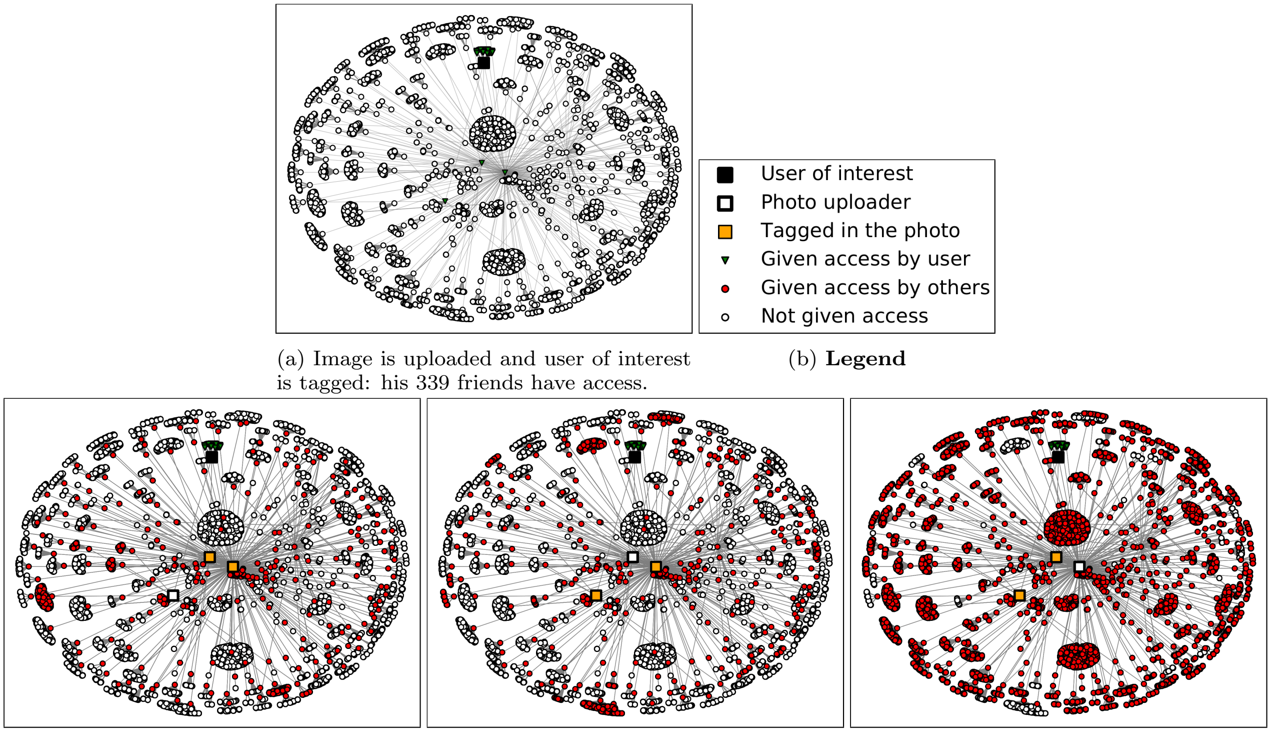
Scenario 1: The Malicious Tagger. Alice and Bob, who are coworkers, attend a party. During the event, Alice takes multiple photos, some of which depict Bob in an inebriated state. Despite that fact, Alice uploads the whole collection of photos and, subsequently, Bob is tagged in the embarrassing photos. In fear of other colleagues and supervisors seeing the photos, potentially creating negative implications, Bob sends Alice a request to remove the photos. Alice, however, does not remove them, and even though Bob un-tags himself, the photos are still viewable by colleagues.

Scenario 2: The Silent Uploader. The settings are similar to the previous scenario. Worse, in this case, Bob is never tagged in the photos and, thus, remains oblivious to the existence of the embarrassing photos. As such, even if Alice was willing to remove them upon request, the photos will be viewable by others until Bob becomes aware of their existence. A recent study [25] explored the extent to which users are aware of photos being shared by others that depict them or contain their tag. Results showed that users are not really aware of the extent of such content, and that there is a significant gap between users' expectations and reality.

Scenario 3: The Group Photographer. This is a very common case of privacy leakage due to conflicting interests. Alice uploads a group picture with Bob being one of the depicted friends. Although Bob is very wary of his privacy and has a strict privacy setting, with his photos being viewable only by his friends, Alice sets the photo to be viewable by all. Despite Bob having actively tried to ensure his privacy, the settings of another user overrules his settings, which results in a loss of privacy. This case is also reported by Yamada et al. [46]. A user study by Liu et al. [32] found that 18% of the users allow users two hops away (i.e., friends of friends) to view their photos, while 26% allow everyone.

Scenario 4: The Accidental Over-sharer. This is also a common case of privacy leakage, where users accidentally, or due to insufficient understanding of their privacy setting, end up sharing photos with a much larger audience than they intended. In [32], it is reported that 63% of the photos have privacy settings different from what intended, and almost always more open. Alarmingly, the privacy setting for 51% of those photos was set to public, allowing anyone to view them. Thus, overall, about one out of every three photos will be publicly viewable by accident. If Alice is the uploader, Bob's face may be accidentally viewable by anyone. This scenario can be attributed to the complexity of current mechanisms, and the uploader being responsible for setting the visibility options for the photo. We propose a simplified scheme where each user is responsible for its own face, and a user's privacy setting is enforced automatically.

Scenario 5: The Friendly Stranger. This case further exemplifies the ineffectiveness of current access control models. Consider that Alice uploads a photo of herself and Bob, and that both of them are cautious with their privacy settings and have opted for a strict setting where photos are only viewable by their friends. This offers a false sense of privacy because, while their interests seem to coincide, that is far from true. Unless Alice and Bob's social graphs perfectly overlap (i.e., identical sets of friends), both users will



(a) Image is uploaded and user of interest is tagged: his 339 friends have access.

(b) **Legend**

(c) The image is uploaded by the 2nd user. 2,871 red nodes (5.2%) have access.

(d) The image is uploaded by the 3rd user. 7,465 red nodes (13.54%) have access.

(e) The image is uploaded by the 4th user. 47,829 red nodes (86.78%) have access.

Figure 1: Risk for a "privacy-conscious" user tagged in a photo. In each case, a different user is considered the uploader (among the depicted users), allowing "friends of friends" to view the photo, while the remaining tagged users are set to "friends only".

be viewable by strangers; e.g., any of Alice's friends that Bob does not know will still be able to see him.

3. RISK ANALYSIS: USER STUDY

In this section we present the findings of our user study that explores the extent of conflicting user interests due to photos shared in social networks. As our focus is on the privacy risks they present to users, we study the characteristics of their social graph and their tagging behaviour.

IRB Approval. Before inviting users to participate in our user study, we issued an IRB protocol request to the review board of our institution, where we described our study and the type of data we would be gathering. After our request was approved, we invited users to participate.

Data and demographics. 128 users participated in our study by installing a Facebook application that collects information regarding the users, their social graph and their photos along with any tag information. The participants are from 14 different countries, with 71% of them belonging to the 20-29 age group and 17.9% to the 30-39 age group. Furthermore, not all the users disclose information regarding their gender, with 55% identifying as male and 16.4% as female. In summary, we analyse data for 4,064,445 photos that contain 4,621,064 tags.

The participants have an average of 344 friends, with a recent survey [11] reporting a similar value of 338. Moreover, about 7% of them have less than 100 friends, while 3% can be considered as *hub* users with more than 1,000 connections.

In Figure 2 we plot the cumulative distribution of the photos that are accessible from each user's profile, i.e., the photos uploaded by each user (or containing a tag of the user) and all the photos belonging to that user's friends (or containing their tags). We will refer to a user and all his/her immediate friends as a *clique*. We found that, on average, each clique has a collection of 31,753 photos belonging to a user and his friends, and 20% of the cliques have more than 44,700 photos. We also discovered that certain cliques of friends are prolific uploaders, with 4% having collections of over 100,000 photos. Based on the numbers stated in [32], we can infer that average users and their friends will accidentally allow almost 15,000 photos to be viewable by anyone, while for prolific uploaders that number will exceed 33,000.

In Figure 3 we plot the cumulative distribution of the total number of tags within the photo collection of each clique, and the number of tagged friends (i.e., unique userIDs). In the average case, a clique's photo collection contains 36,102 tags and has 250 tagged users. Furthermore, we find that 20% of the cliques have over 340 different tagged users in their photos, and have over 50,000 photos in their collection. In three cases, the clique has over 1,000 tagged UIDs. These numbers signify the risk of the aforementioned scenarios that arises from the current access control mechanism; within a clique of users, the ownership and visibility of thousands of photos (some being potentially embarrassing) is handled by multiple users that may have conflicting interests.

As described in the **silent uploader** scenario, users may never be tagged in the "embarrassing" photos and, therefore, never be alerted of their existence. To gain an estimation of this risk we conduct an experiment where we first manually inspect 2,000 randomly selected photos. Figure 4 shows the

Tags | **Faces**	**1**	**2**	**3**	**4**	**5**	**6+**
Photos (# of Faces)	15.2% (304)	32.5% (651)	17.9% (359)	10.7% (214)	8.3% (166)	15 .3% (306)
Photos (# of Tags)	87.6% (1753)	9.9% (199)	1.6% (33)	0.3% (7)	0.25% (5)	0.15% (3)

Table 1: Percentage (and number) of photos in our 2,000 photo dataset that contain a given number of tags or faces.

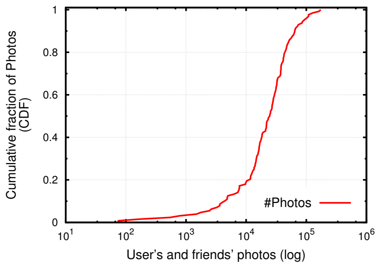


Figure 2: Cumulative distribution of uploaded photos depicting (or belonging to) users and their friends.

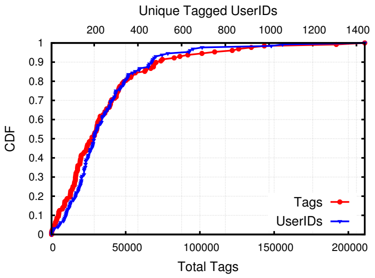


Figure 3: Cumulative distribution of total number of tags within a clique's photo collection, and the number of unique tagged UIDs.

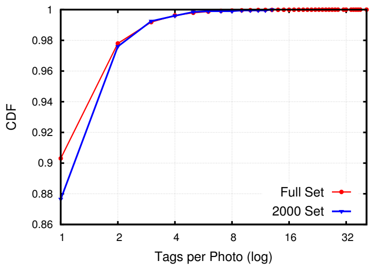


Figure 4: Cumulative distribution of number of tags per photo, for the complete dataset, and the set of 2,000 randomly chosen photos.

number of tags from the photos of our entire dataset, and of the 2,000 randomly chosen photos. As we can see, the randomly selected photos form a representative sample of our dataset, in terms of the number of tagged users per photo. Subsequently, we inspect these 2,000 photos and count the depicted faces that are discernible, both in the foreground and the background. We only take into consideration faces that could be identified by their friends, and overlook any non-identifiable faces (e.g., being too blurry, obstacles etc.). Table 1 presents the number of photos that contain identifiable faces. The photos depict a total of 7,244 faces (3.62 faces per photo) out of which 2,331 have been tagged (1.16 per photo). Only 15.2% of the photos depict one user, and about half of them depict two or three users. But, the vast majority (87.6%) contain only one tag. Thus, on average, *every photo depicts at least two users that have not been tagged and could be at risk due to the silent uploader scenario.*

According to the user study by Liu et al. [32], one out of four users has a public privacy setting for his photos. Thus, there is a high probability that photos depicting at least four people, will result in conflicting interests, as described in the **group photographer** scenario. In our dataset of 2,000 manually inspected photos, we found that 34.3% of them depicted at least four identifiable people.

To further explore how users are tagged, in Figure 5 we plot the number of tags for each userID in our collection. We have tags from 35,809 userIDs, and 30% of the users are being depicted in 72.4% of the tags. The majority of tags depict a small set of users that are tagged extensively, with the top 10% of users having an average of 594.9 tags and, when combined, amounting to 39.5% of the total tags. We do not have the information to conclude if this is due to these users not being concerned about privacy, or wrongfully "relaxed" privacy settings. The data, however, does suggest that certain users are more cautious about their privacy, as those from the least-tagged 10% have 3.41 tags on average.

Next, we focus on the risk that arises for users even when the uploader has strict privacy settings (i.e., photos are only visible to uploader's friends). In this experiment, we consider our participants as the "adversaries" of the **friendly stranger** scenario, and explore how many photos of strangers they are able to view, if the uploader had selected the "friends only" setting. Specifically, for each participant, we calculate the number of photos that have been uploaded by that user or his/her friends, and contain the tag of a user not in his/her friendlist (we refer to them as *strangers*). Figure 6 presents the results, with 92% of the participants having access to photos where strangers have been tagged. On average, these users can view 647 photos of 169 different users to which they are not connected, regardless of the privacy settings those users have set. One user can view 1,866 photos depicting 1,073 different strangers, while each of the top 10% users can view photos of at least 358 strangers. As such, even if the OSN opts for a more privacy-oriented approach, where the default setting for photos is "viewable by friends only", users' faces will remain viewable by many strangers.

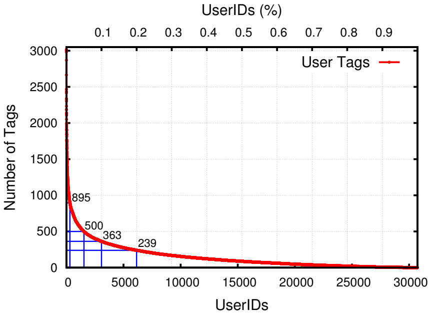


Figure 5: The number of tags contained in our dataset regarding every test subject and their friends, which follows a power law distribution. We sort users based on the number of times they have been tagged, and also depict the number of tags for the users at the 1%, 5%, 10% and 20% positions.

Overall, our study confirms concerns regarding the privacy risks that emerge from shared photos and threaten users, and demonstrates the necessity for a fine-grained access control mechanism, as the one we propose.

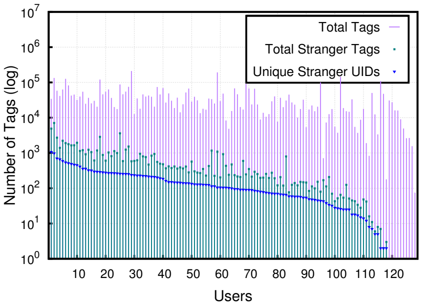


Figure 6: For each user study participant's clique, we plot the total number of tags, the tags that belong to users not associated with that participant (i.e., strangers), and how many (unique) UIDs belong to strangers.

4. ACCESS CONTROL MODEL

The key concept of our approach is to refine the object of the access control model, switching from photos (coarse-grained) to faces (fine-grained). As summarized in Figure 7, the *objects* in our access control model are the faces, the *subjects* are the users, whereas the photos are modeled as *object groups*. This allows us to define the concept of fine-grained, content-based, multi-owner control policy for photos.

The photo owner has a write-only right for publishing the photo. Read rights are enforced by the users whose faces are depicted in the photo. For example, in Figure 7, the user `U2` owns the photo `P2` (solid dot), which depicts `U1`, `U3`, and `U5`'s faces (empty dot, or solid dots on the diagonal axis).

This model could be implemented with a simple 2D sparse matrix, replacing any existing access control model, or as an extension, by adding an additional list containing the permission bits as necessary. By choosing to visualize it as a 3D matrix, we highlight that our model is an extension of the current model and does not interferes with it. As a matter of fact, this model can provide the exact functionality of the current one, simply by enabling the permission bits on all the objects. This model is implemented in the following.

4.1 System Design

Here we describe how our system resolves conflicting cases in requested photos. We design the system by assuming the availability of the existing functionalities of OSNs, namely face recognition (as in Facebook and Google+), image processing, and access control policy enforcement based on user preferences. Figure 8 provides an overview of the work-flow of our approach, which is further detailed in the following.

Step 1: Face recognition. We rely on face recognition to detect faces of known users, which become objects in the access control model. This process takes place once a user uploads a photo in the OSN. Each detected face is first compared to the classifiers of the uploader's contacts, as there is a high possibility that the depicted users will be friends with the uploader. Previous work [41] has also shown that social relationships can be used to further improve face recognition results. Detected faces that do not match any of the uploader's contacts, will subsequently be compared to the contacts of the other depicted users. Depending on the computational resources available, this step can be extended to include an arbitrarily larger portion of users.

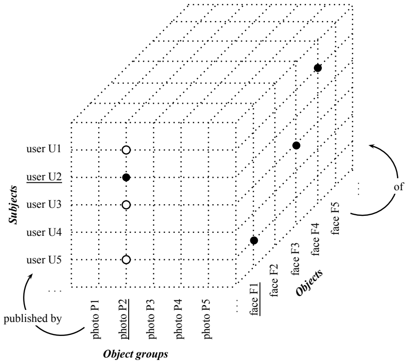


Figure 7: Visualization of our revised access control model.

Auto tagging and suggestion. Auto-suggestions of the identified faces are displayed to the user to verify the subjects' identity, if necessary. Moreover, we request the uploader to tag any unidentified users. The auto-suggestion mechanism is already implemented in modern OSNs.

Step 2: Template generation. The depicted (recognized) users are notified about the photo and everyone sets its own permissions. If a default setting has been already set by a user, the system can enforce it automatically and allow adjustments on a per-photo basis. Then, a template of the processed photo is generated.

User notification. Every user identified in the photo is automatically notified that a photo with his/her face has been uploaded. Users will be asked to verify the validity of the face (if its actually him/her) and set the access control for the specific photo. Until the depicted user has processed the specific face, even if tagged by other users, the face will remain hidden and no tag will appear. The mechanism for allowing a tag is already implemented by Facebook, in the form of *tag review* [4], and users have grown accustomed to such requests. However, differently from our approach, the user's selection is reflected solely on the visibility of the entire photo within the user's albums.

The output of this phase is a *template photo*, which is composed by the uploaded photo and a set of F layers, where F is the number of faces recognized. Each layer represents a face f appearing in the original photo p and has the size of the patch corresponding to the associated face area. Each tuple $\langle p, f \rangle$ in the template is processed: the face patch f is pixelized/blurred, or set to a solid color.

Step 3: Template rendering. When a photo is to be viewed by a subject, we select the corresponding row in the access control matrix (see Figure 7). This allows us to determine, in constant time, the faces (objects), $f_1, f_2, \ldots$, that the subject is allowed to view (read) according to each face's owner's privacy setting for that photo, p. Based on this information, we create a photo "on the fly" and serve it to the user. Thanks to the template photo, this can be performed efficiently, by simply superimposing the required layers $\langle p, f_i \rangle$ on the original photo.

User lists. Each user has a personalized set of lists and populates them with certain contacts. Every list may represent a group of friends with a common characteristic (e.g., coworkers, family, close friends). These lists are used for assigning permissions to groups of contacts for our fine-grained access control mechanism. The user can create a new list at

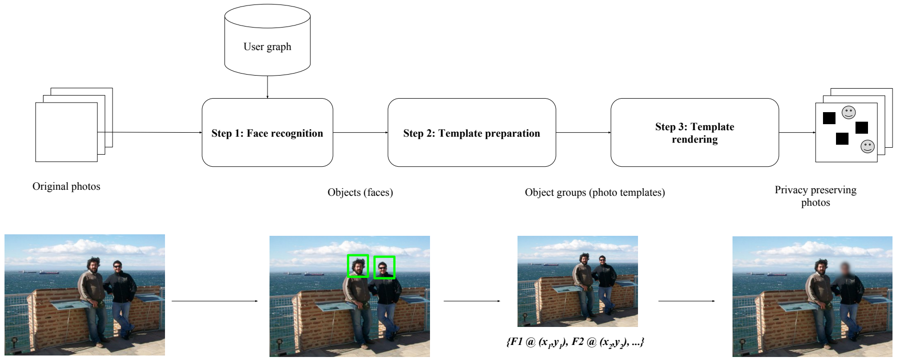

User graph
Original photos
Step 1: Face recognition
Step 2: Template preparation
Step 3: Template rendering
Objects (faces)
Object groups (photo templates)
Privacy preserving photos
{F1 @ (x1,y1), F2 @ (x2,y2), ...}

Figure 8: Overview of our approach. In **Step 1** we identify the depicted faces and associate them to the users' identities. In **Step 2** we create a template of multiple layers, each containing a single hidden face from the original photo. In **Step 3** we create a "processed" photo by superimposing the template layers on the original photo, according to users' permissions.

```
Algorithm 4.1: ViewPhotos($U, V$)

$\mathcal{P} \leftarrow$ ListOfPhotos($U$)
$\mathcal{F} \leftarrow$ Tag_Present($U, P$)
$\mathcal{N} \leftarrow$ Tag_Not_Present($U, P$)
comment: $\{P_1, ..., P_i\} = \{F_1, ..., F_j\} \cup \{N_1, ..., N_k\}$,
        $where$ i = j + k
for each  x $\in \mathcal{N}$
  do { $photo \leftarrow$ FaceOff(x, $V$)
       Show($photo$)
for each  x $\in \mathcal{F}$
  do { $access\_flag \leftarrow$ TagAccess(x, $U, V$)
       if $access\_flag = 1$
         then { $photo \leftarrow$ FaceOff(x, $V$)
                Show($photo$)
```

Figure 9: Pseudo-code of photo selection and access control enforcement mechanism.

any time or remove an existing one. Access permission is not irrevocable or permanent, as the user can modify his friend-lists by adding new friends or removing some of the existing ones, to revoke their permissions. Lists can be managed (create/modify) during the permission assignment phase, as the user may wish to create a new list for the specific photo (e.g., friends that attended event X). Note that the custom friend-list functionality is already provided by most OSNs.

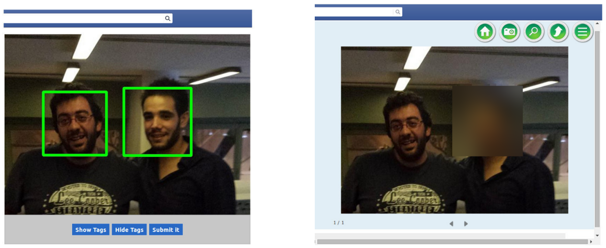
Access control. Our goal is to provide an efficient face-level, fine-grained access control mechanism that smoothly operates on top of the traditional photo-level mechanisms. Thus, the existing photo-level access mechanism used to populate the photo albums a user is attempting to view, remains as is. After the set of photos is identified, our face-level granularity access mechanism is employed, for determining which depicted faces can be viewed and which must be hidden from the user. Thus, if our model is adopted by an OSN it can extend the existing mechanisms.

The procedure of selecting the photos of user U that will be shown to the viewer V is presented in Figure 9. Using the existing photo-level mechanism, we create the set of photos P that the viewer is allowed to access. This set can be broken down to two subsets, F where U's face is present, and N where the U is absent. For every photo in N, we check the permissions for every individual user depicted and hide any faces, that should not be viewable. For photos in F, we only present photos where the viewer has the permission to view U's tag, and once again, we check the permissions of every individual face.

The reason for using our fined-grained access control mechanism in conjunction with the existing mechanism can be highlighted with the following scenario, as it demonstrates how we achieve stronger privacy in certain cases. Consider the case where Alice is trying to view Bob's photos. For a specific photo where Bob is depicted along with Marjorie, who is also a friend of Alice, Bob has set a privacy setting that prohibits Alice from viewing her face. However, Marjorie has a less restrictive setting. If Alice was able to view the photo, where Bob's face would be blurred, she would be able to argue that the hidden face most likely belongs to Bob, as she is accessing Bob's photo album. One could state that this precaution may be redundant because Alice can view the specific photo through Marjorie's albums. However, in an extreme case where Bob, Alice and Marjorie have the exact set of users as online friends, Alice could reveal that Bob's face is hidden, by accessing the photo albums of all of her friends. Since the photo will be presented only in Bob's and Marjorie's albums, she can infer without a doubt that Bob is depicted in the photo. While this example may present a very extreme case, even in normal cases Alice is inclined to consider Bob as the most plausible candidate. Thus, we choose to hide such photos from Alice, so when viewing the photo through Marjorie, any other user is equally possible to be hidden beneath the blurred section.

5. IMPLEMENTATION DETAILS

In this section we describe the proof-of-concept implementation of our mechanism. We built our prototype as a third-party Facebook app that is hosted on our web server, which is also used for storing the uploaded photos, user information and users' permission matrices. We store all users' data locally, as our intention is not to provide another tool for altering the current access control mechanisms, but to demonstrate the functionality of our approach and to verify that it can be easily integrated into existing OSNs. The fact that we were able to implement our approach as an external application, without any modification in the backend, indicates the unobtrusiveness of our mechanism.



(a) The photo is uploaded, faces are detected and marked for tagging.

(b) User requests access. Photos are blurred selectively, according to depicted users' permissions.

Figure 10: Screenshot of our proof-of-concept application.

Installation. When the Face/Off application is installed by the users, it requests permissions for reading and managing users' friend-lists. These enable us to allow the user to create custom friend-lists within the application. When the user loads the app, it is authorized through the Facebook authentication mechanism and the application's database is updated with the user's current friend-lists. This allows us to easily keep track of newly created lists, of users that have been un-friended or simply removed from some of the lists (a simple way to revoke permissions).

Initial photo review. The application first determines if any new photos that contain a tag of the user have been uploaded. In such a case, thumbnails of these photos are presented to the user, who is able to load each photo for inspecting the tag and for choosing which of his friend-lists are permitted to access it. It should be noted that the face of the user remains hidden to others as long as the permissions have not been set, similarly to the case where the viewer has not been granted access.

Face Detection. When a new photo is uploaded our application performs face detection, and the detected faces are marked, as shown in Figure 10a. The main omittance of our proof-of-concept implementation is that we do not perform face recognition but rely on the uploader to tag the photo. Similarly, Facebook prompts the user to assign names to the detected faces upon each uploaded photo. We decided to only implement face detection but not recognition as that would have required us to collect the Facebook photos of the user and all of his friends to achieve accurate face recognition results. However, Facebook has acquired `face.com` and according to a comparative study [28], the face recognition algorithm of `face.com` was the most accurate and effective tested. Moreover, in [44] the authors state that they achieve a 97.5% identification rate.

The server generates and stores a unique photoID for the uploaded photo and a faceID for each one of the faces. For the generation of IDs the server uses the userIDs of the uploader and each one of the tagged user, the server's internal time and a one-way hash function. After that, the server starts processing the image by cropping and blurring the depicted faces. This functionality does not affect user experience as all the processing is performed in the background.

Photo rendering. When access to a photo is requested, we fetch all the information of this photo and its tags, and determine which faces can be revealed and which should remain hidden, by checking the users' friend-lists. Then, we generate a processed image "*on the fly*", by superimposing the blurred layers of the template on top of the photo, and we populate it into the user's album, as shown in Figure 10b.

In our prototype we implement the functionality of the fine-grained access control mechanism, but do not replicate the existing photo-level mechanism. We follow a simplified approach by considering that all the photos can be accessed by the *friends* of the uploader and the *friends* of each tagged user. However, our implementation takes into consideration the case where the uploader's face should remain hidden, as described in Section 4.1, and does not populate these photos in the uploader's photo album.

6. EVALUATION

In this section we evaluate several aspects of our approach. First, we measure the overhead introduced by our system. Next, we conduct a user study to evaluate the effectiveness of our approach in preserving the privacy of users. Finally, we explore the willingness of users to adopt our fine-grained access control mechanism for protecting their privacy.

6.1 Performance Evaluation

Regarding the performance overhead imposed by our approach, one must take into account that several OSNs already have the infrastructure available for performing real-time face recognition on uploaded photos. This functionality has already been implemented by Facebook and Google+ for supporting their tagging suggestion mechanism. Here, we measure the processing overhead of our mechanism; we do not measure the time required for the face detection process, as we focus on the overhead incurred by actions that are not already performed by the service. All experiments were conducted on a commodity desktop machine.

Overhead: First, we measure the overhead presented by the photo preparation phase, which takes place after a photo has been uploaded and faces have been identified. This includes cropping detected faces and creating a blurred layer of each face. We select 100 random photos from our user study and process them. This phase takes 0.0023 seconds on average per tag, and is performed before the photo is added to the uploader's albums. This overhead is negligible, especially when considering that OSNs already perform transformations to uploaded photos (e.g., resizing).

Figure 11 presents the results from the experiment regarding the access control enforcement and photo transformation. Again, we upload 100 photos to our app, and tag one of the faces. We then access the photos from an account that does not have permission to view that tag, and measure the total time required for creating the processed photo "on the fly". This includes retrieving the access control lists for the photo, selecting the faces to be blurred, overlaying the blurred sections, and saving the transformed image. Overlaying the blurred layer for a single tag requires merely 0.001 on average (0.022 seconds in the worst case) which is negligible. The time required for the complete process ranges from 0.012 to 0.109 seconds, with an average value of 0.052.

Thus, the main overhead of our approach is loading the photo and the template layers from the filesystem, retrieving the access permissions from the database for each depicted user, and deciding which faces the accessing user should view. This process is dependent on the number of people depicted in the photo, the permissions of each user and their number of friends. In our experiments, we selected a user with 452 friends, which is higher than the average of 344.

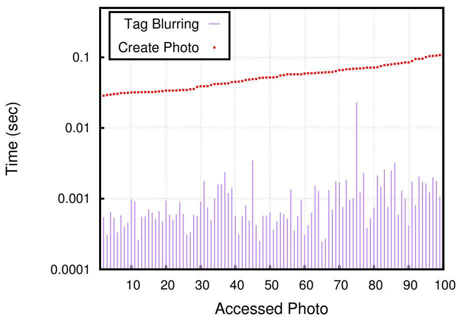


Figure 11: The total time required for serving a photo, which includes reading the access permission and blurring a face.

According to [44], the complete processing of an image for face identification, conducted by Facebook, lasts 0.33 seconds when executed on a single core machine. Thus, on average, our fine-grained access control will incur at most a 15.6% increase of the duration of the photo processing already conducted (if other processing is done, the overhead will be even less). Moreover, these values will be much lower when executed on high-end servers found in the data centers of major web services. Also, our proof-of-concept implementation can be optimized, which will further reduce the overhead. Overall, we believe that this small overhead is justified by the privacy gain the users will benefit from.

Scalability. In an attempt to further explore the performance and the scalability of our approach, we select another set of 100 random photos that contain at least three depicted faces. At first, we upload all the photos, tag a single face in each photo and access them from multiple accounts that are not allowed to view the face. We repeat this process 2 more times, by uploading again the same photos and tagging two and three of the depicted faces respectively. The three tagged users have 452, 1173 and 442 friends. Each extra tag increased the processing time by 0.002 seconds.

From the last experiment, we can conclude that our mechanism is scalable, as the number of tags in a photo and the number of the tagged users friends has a very small impact on the performance of the system. It can be observed, that the bulk of processing time is spent on fetching the photo from the filesystem, and not on retrieving the access lists or computing the permissions. While our experiments are not an extensive measurement of the overhead of our approach under all possible scenarios, they are indicative of the small overhead imposed by our access control mechanism.

6.2 Privacy Evaluation

To evaluate the effectiveness of our approach in preventing the identification of depicted users, we invited the participants of the risk analysis study (Section 3) to take part in an experiment where we would apply our approach to photos of their friends. The 34 users that participated were shown a set of randomly selected photos of their contacts, with one friend "hidden" in each photo, and were requested to identify the hidden friend. In cases where they supplied a guess for the hidden user, they were also required to provide feedback regarding the visual clues that influenced their guessing. To reflect actual use cases, all photos depicted multiple people.

Ideally, this experiment would be conducted by deploying our proof-of-concept application at full scale and asking the participants to identify their restricted friends within each accessed photo. This would allow us to ensure the "freshness" of the photos, and avoid using photos that the participants have previously seen. However, this experimental setup requires the participants' friends to also install the application and upload new photos, which poses many practical difficulties. If only a small number of the user's friends installs the application, the pool of users to "hide" will be limited, and results could be heavily biased.

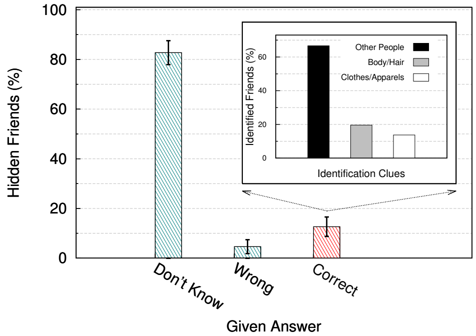


Figure 12: Identification of hidden contacts (95% confidence interval). For correct answers, we break down the visual clues that led to the identification.

Thus, we opt for an alternative experimental setup; we use photos collected during the risk analysis study. To obtain an accurate evaluation of the privacy offered by our approach, we do not consider photos where the user feedback stated that they remembered seeing them before. First, we randomly select a set of photos that depict at least one of the participant's friends. Apart from containing the tag of a friend, we also ensure that they have not been uploaded by our participants, nor do they contain their tag. Moreover, we manually verify the correctness of tag placement, which will result in the hidden area. Then, our mechanism blurs out the friend's face in each photo, and presents the photo challenge to the participants.

The results of our experiment are shown in Figure 12. We prepared and presented a total of 476 challenges, out of which 448 had not been seen before by the participants, according to their feedback. We manually verified answers to avoid erroneous evaluation due to spelling mistakes. Users stated that they could not identify their friends, and did not suggest a name, for 82.7% of photos they were shown. On average, users correctly identified the hidden user in 12.6% of their challenges, and gave a wrong answer for 4.6%.

As can be seen, the dominating clue for correctly guessing the identity of a restricted user was the existence of other people within the photo known by the participant. The non-restricted people in the photo allowed users to correctly infer the hidden user in 66.7% of the cases. In 19.6% of the identified challenges, the body or hair led to identification, while clothes were helpful in 13.6%. Thus, while other people are the dominating reason for inferring the identity of the user, other visual clues that can be potentially removed, have significant contribution. We discuss how we plan to extend our approach for mitigating this effect in Section 7.

These numbers offer a upper bound as, in practice, users may be presented with multiple hidden faces in a photo, which will make identification harder. Furthermore, the par-

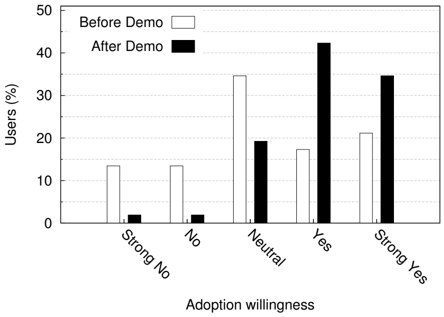


Figure 13: Willingness of users to adopt a mechanism that blur out faces in uploaded photos. Users' opinion before and after a short demonstration of our proof-of-concept app.

ticipants knew that the hidden users were friends of theirs. In an actual deployment, not all the hidden users will be contacts of theirs, which will increase uncertainty and may result in even less identified users. Overall, while the number of participants is relatively small, our results are promising as they indicate the effectiveness of our approach in hiding the identity of users from their contacts.

6.3 Adoption Willingness

A crucial factor in determining the merits of our approach, is the attitude of users towards the potential adoption of our system by popular services. To explore that aspect, we conducted a user study for identifying the willingness of users to adopt and use a face-level fine-grained access control mechanism. To obtain a more objective understanding of users' opinion, we opted for a set of new subjects that had not participated in any of our previous experiments and were unaware of our approach. This offered an unbiased view of how people will react to such a mechanism being deployed. A total of 52 users participated, with 65.4% being male and 34.6% female, all in the age range of 18-36.

First, we presented a photo processed by our mechanism that contained some hidden faces, and asked users if they would like such a mechanism to be implemented by photo-sharing social networking services. After the users' response, we presented the privacy implications that arise from conflicts of interest and briefly demonstrated our proof-of-concept application. Users were allowed to interact with it. Then, we asked them if they wanted OSNs to adopt such a mechanism, selecting from answers modelled after the Likert scale.

The results are shown in Figure 13. Initially almost 27% of the participants were against the adoption of such a mechanism, while about 35% reported a neutral opinion. In most cases, users responded negatively due to a false belief that current access control mechanisms are effective. The remaining negative answers were from users that were not interested in privacy implications created by widely accessible photos. The users that had selected a neutral stance, recognized the need that other users may have for preserving their privacy, but did not have a strong motivation in using such a mechanism. However, these users were also not aware of the true visibility of their photos. On the other hand, 38.4% of the participants immediately shaped a positive opinion of OSNs adopting a fine-grained access control mechanism.

Interestingly, there was a remarkable shift in user opinions after introducing the problem of conflicting interests, and demonstrating our application. Only 3.8% of the participants maintained a negative opinion, and 19.2% remained neutral. Almost 77% of the users wanted such a mechanism to be adopted. We observed that most of the initially negative and neutral participants care about their privacy, but were not aware of the current access control mechanisms and the visibility of their data. Moreover, several of the initially negative users, having stated that they do not care about privacy, became neutral and accepted the necessity of such a mechanism, as they recognized the privacy needs of others.

Finally, we asked users to assess the usability of our approach, in a 5-point rating scale. 86.5% of the users rated our mechanism as usable and very usable (4 and 5 points). 11.5% and 1.9% of the users rated the mechanism with 3 and 2 points respectively, due to the lack of an option for assigning the same permissive lists to multiple photos, at once. This, however, does not impact the usability of our approach, as this concerns our proof-of-concept implementation, and not the core access control mechanism, and can be easily addressed in the future.

7. LIMITATIONS AND FUTURE WORK

User Removal. A recent user study [33] demonstrated that users are effective at recognizing their friends even in photos where their face is not clearly visible. However, in the study, users were significantly aided as they had to select from a list of 6 possible friends. In our study, participants were able to only guess the identity of 12.6% of the users. Thus, while our current approach offers a significant step towards a more privacy-preserving sharing of content within OSNs, we plan to explore methods to further improve effectiveness. Specifically, we plan to explore the feasibility of completely removing the depicted user from the presented photo. A large body of work has demonstrated effective techniques for automatically removing objects from images and reconstructing the affected region (e.g. [17, 22]) with performance suitable for processing big data [47]. Thus, after the user's body/pose is identified [38], the photo can be processed to completely remove him/her.

Collateral Inference. Even with our mechanism in place, a user's identity might be inferred from information found in the photo's comments. As such, further exploration is required for determining the extensibility of our mechanism to also handle comments associated with a photo.

Identification accuracy. The effectiveness of our approach relies, to an extent, on the accuracy of the face identification software employed by the social network. To prevent malicious user behavior, such as uploaders not tagging users (to prevent the users from hiding their face), or falsely tagging faces, our system has to employ highly accurate software for the identification of the depicted users. According to Taigman et al. [44] Facebook's method reaches an accuracy of 97.35%, rendering it suitable for our approach. In cases where a face cannot be identified, users may be asked to provide a suggestion and the system can accept answers only if there is consensus among several users.

Non-members. A case where our approach cannot protect a user's privacy, is when a photo depicts a user who does not have an account in the social network. If such an event occurs, various approaches can be applied, such as following a strict permission where all such faces are hidden, or a more lenient setting where the photo uploader is considered the owner and applies the privacy setting.

Legislation. European data protection agencies have pressured Facebook into removing the tag-suggestion mechanism due to privacy concerns over face recognition software processing uploaded photos without users' consent [1, 12]. This resulted in the tag-suggestion mechanism being temporarily disabled and the deletion of biometric data collected, for users located in European countries [2, 14]. To that end, many consider that face recognition software will have limited adoption. Fortunately, there is active research towards privacy-preserving face recognition [24,36] and, therefore, we envision that this very effective technology will be adopted by such services. Nevertheless, this paper is orthogonal to privacy concerns and legislation issues related to face recognition. In actuality, our approach takes advantage of automated face recognition for *enhancing user privacy.*

8. RELATED WORK

In [18] Besmer et al. studied the behavior of users regarding photo sharing applications, and identified the reasons users choose to tag or un-tag a photo. During their study they demonstrated a simple prototype that obfuscates faces, in an attempt to initiate a discussion about user privacy and photo ownership. Their findings highlighted user concerns in regards to the visibility of images and the lack of effective access control mechanisms. Their results argue that users are interested in shaping their identity in order to manage impressions and avoid exposing situations they are not comfortable with. In a follow-up [19], they presented a "negotiation" tool that allows each tagged user to send an out-of-band request to the photo uploader, for requesting the photo to become non accessible by particular users. However, it remains entirely up to the uploader to accept or reject user's request. Even though users can contribute in access control by sending a request, this does not solve conflicts of interest.

Multiple works [20, 21, 29, 40] follow the rule-based access control approach. In [29] users are allowed to annotate their photos with semantically meaningful tags and to specify access control rules based on these tags. The work presented in [40] uses previously uploaded photos, and their access control rules, for classifying each new photo by its content and for predicting an access control rule that will be acceptable by the uploader. The advantage of this approach is that the prediction is adaptive to the behavior of the user. However, all these approaches create a complex set of rules and also consider access control at the photo level.

Al Bouna et al. presented a system for preserving privacy regarding multimedia objects [21], which can be specifically used for photos [20]. They have designed a security model and built a security rule specification toolkit that uses the SWRL language for specifying content-based access control rules. Their prototype has the ability to hide faces among others, but it does not distinguish access control from the conflict resolving mechanism. Importantly, this approach does not allow each depicted individual to set his/her own rules, but only the uploader. When two or more rules are conflicting, a security administrator is required to set priority values on the execution of the rules. This, of course, is not feasible at the large scale of an OSN.

In [45] Thomas et al. highlighted the lack of a multi-party access control mechanisms for shared content that is uploaded by other users in OSNs. They studied the conflicting privacy settings between friends and how these settings can reveal sensitive information that was intended to be private. But, their proposed approach is very strict and far from usable, as objects are revealed only to the mutual friends of the related users. Also, [26, 27, 39] proposed multi-party mechanisms for allowing collaboration between the users regarding the specification of the access control policy. However, even if collaboration is allowed, the access control is enforced at photo level, which cannot effectively accommodate the privacy preferences of all the depicted users.

Cutillo et al. [23] presented a demanding cryptography-based face obfuscation mechanism for a specific decentralized OSN, namely, the Safebook. This mechanism is far from applicable within the environment of existing OSNs, as it leverages the multi-hop routing protocol of the specific OSN. On the other hand, our approach is designed for easy integration with existing social networks, relying on technological capabilities widely available to such services.

9. CONCLUSIONS

In this work we tackled the problem of conflicting interests that arise from photos being shared in social networks. The problem stems from the current design of OSNs, as users associated with a shared photo have limited control over its visibility, and their privacy settings usually are overridden by those of other users. As such, we identified the different scenarios where conflicts of interests can occur, and we conducted a case study in order to quantify the privacy risks presented. We collected a large number of photos, along with their tags, for assessing users' tagging behavior, and for determining the true visibility of shared photos.

We designed a fine-grained access control mechanism that allows depicted users to define the exposure of their own face, by setting their preferred permissions. When a photo is requested, our mechanism determines which faces should be hidden and which should be revealed based on the requesting user, and presents a "processed" version of the photo. Our mechanism can be implemented on top of the existing access control mechanisms and smoothly interoperate with them, as demonstrated by our proof-of-concept implementation. The proposed approach is scalable, as it imposes only a small processing overhead. Finally, we conducted a user study to evaluate the effectiveness of our approach, and found that hiding users' faces is an effective measure for enabling privacy in shared photos. Our study also revealed the misconceptions users have regarding existing access control mechanisms, and showed that users are positive towards the adoption of a face-level access control mechanism.

Acknowledgements

We thank the anonymous reviewers for their valuable comments. This work was supported by the FP7 Marie-Curie ITN iSocial funded by the EC under grant agreement no 316808, by the NSF under Grant CNS-13-18415, and by the MIUR under the FIRB2013 FACE grant. Any opinions, fundings, conclusions, or recommendations expressed herein are those of the authors, and do not necessarily reflect those of the US Government or the NSF.

10. REFERENCES

[1] Data Protection Commissioner - Facebook Ireland Audit. [accessed Aug-2015].

[2] Data Protection Commissioner - Facebook Ireland Re-Audit. [accessed Aug-2015].

[3] Facebook - Stats. [accessed Aug-2015].

[4] Facebook - Tag Review. [accessed Aug-2015].

[5] Facebook Privacy Selector. [accessed Aug-2015].

[6] Bussiness Insider - Facebook Users Are Uploading 350 Million New Photos Each Day. [accessed Aug-2015].

[7] Business Insider - A High School Coach Was Fired For Facebook Photo. [accessed Aug-2015].

[8] CBS news - Did the Internet Kill Privacy? [accessed Aug-2015].

[9] Germany Sues Facebook For Violating Users' Privacy. [accessed Aug-2015].

[10] Social, Digital Video Drive Further Growth in Time Spent Online. [accessed Aug-2015].

[11] Pew Research Center - Facebook Survey. [accessed Aug-2015].

[12] Telegraph - Facebook defends using profile pictures for facial recognition. [accessed Aug-2015].

[13] Wired - Facebook Envisions AI That Keeps You From Uploading Embarrassing Pics. [accessed Aug-2015].

[14] Wired - Facebook complies with EU data protection law. [accessed Aug-2015].

[15] Microsoft - Online Reputation in a Connected World, 2009.

[16] A. Acquisti and C. M. Fong. An experiment in hiring discrimination via online social networks. 2013.

[17] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester. Image inpainting. In *SIGGRAPH '00*.

[18] A. Besmer and H. R. Lipford. Privacy perceptions of photo sharing in facebook. SOUPS '08.

[19] A. Besmer and H. R. Lipford. Moving beyond untagging: Photo privacy in a tagged world. In *Proceedings of CHI '10*, 2010.

[20] B. A. Bouna, R. Chbeir, A. Gabillon, and P. Capolsini. A flexible image-based access control model for social networks. In *Security and Privacy Preserving in Social Networks*. Springer, 2013.

[21] B. A. Bouna, R. Chbeir, A. Gabillon, et al. The image protector-a flexible security rule specification toolkit. In *SECRYPT*, 2011.

[22] A. Criminisi, P. Pérez, and K. Toyama. Region filling and object removal by exemplar-based image inpainting. *Transactions on Image Processing*, 13(9).

[23] L. A. Cutillo, R. Molva, and M. Önen. Privacy preserving picture sharing: Enforcing usage control in distributed on-line social networks. In *SNS '12*, 2012.

[24] Z. Erkin, M. Franz, J. Guajardo, S. Katzenbeisser, I. Lagendijk, and T. Toft. Privacy-preserving face recognition. In *PETS*, 2009.

[25] B. Henne, M. Linke, and M. Smith. A study on the unawareness of shared photos in social network services. In *Web 2.0 Security Privacy (W2SP)*, 2014.

[26] H. Hu, G.-J. Ahn, and J. Jorgensen. Detecting and resolving privacy conflicts for collaborative data sharing in online social networks. In *ACSAC '11*.

[27] H. Hu, G.-J. Ahn, and J. Jorgensen. Enabling collaborative data sharing in google+. In *GLOBECOM'12*, 2012.

[28] G. B. Huang and E. Learned-Miller. Labeled faces in the wild: Updates and new reporting procedures. Technical Report UM-CS-2014-003, UMass Amherst.

[29] P. Klemperer, Y. Liang, M. Mazurek, M. Sleeper, B. Ur, L. Bauer, L. F. Cranor, N. Gupta, and M. Reiter. Tag, you can see it!: Using tags for access control in photo sharing. In *CHI '12*.

[30] B. P. Knijnenburg, A. Kobsa, and H. Jin. Dimensionality of information disclosure behavior. *IJHCS*, 71(12):1144 – 1162, 2013.

[31] B. Krishnamurthy and C. E. Wills. Characterizing privacy in online social networks. In *WOSN '08*.

[32] Y. Liu, K. P. Gummadi, B. Krishnamurthy, and A. Mislove. Analyzing facebook privacy settings: User expectations vs. reality. In *IMC '11*.

[33] I. Polakis, P. Ilia, F. Maggi, M. Lancini, G. Kontaxis, S. Zanero, S. Ioannidis, and A. D. Keromytis. Faces in the distorting mirror: Revisiting photo-based social authentication. CCS'14.

[34] I. Polakis, M. Lancini, G. Kontaxis, F. Maggi, S. Ioannidis, A. Keromytis, and S. Zanero. All your face are belong to us: Breaking facebook's social authentication. In *ACSAC '12*, 2012.

[35] P. Rao, D. Lin, E. Bertino, N. Li, and J. Lobo. Fine-grained integration of access control policies. *Computers & Security*, 30(2-3):91–107, 2011.

[36] A.-R. Sadeghi, T. Schneider, and I. Wehrenberg. Efficient privacy-preserving face recognition. ICISC'09.

[37] Y. Shoshitaishvili, C. Kruegel, and G. Vigna. Portrait of a privacy invasion: Detecting relationships through large-scale photo analysis. In *PETS*, 2015.

[38] J. Shotton, T. Sharp, A. Kipman, A. Fitzgibbon, M. Finocchio, A. Blake, M. Cook, and R. Moore. Real-time human pose recognition in parts from single depth images. *Commun. ACM*, 56(1), Jan. 2013.

[39] A. C. Squicciarini, M. Shehab, and F. Paci. Collective privacy management in social networks. WWW '09.

[40] A. C. Squicciarini, S. Sundareswaran, D. Lin, and J. Wede. A3P: Adaptive policy prediction for shared images over popular content sharing sites. HT '11.

[41] Z. Stone, T. Zickler, and T. Darrell. Autotagging facebook: Social network context improves photo annotation. In *CVPRW '08*.

[42] M. M. Strano and J. Wattai Queen. Covering your face on facebook. *Journal of Media Psychology: Theories, Methods, and Applications*, 24(4), 2012.

[43] K. Strater and H. R. Lipford. Strategies and struggles with privacy in an online social networking community. In *BCS HCI '08*.

[44] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf. DeepFace: Closing the Gap to Human-Level Performance in Face Verification. In *CVPR '14*.

[45] K. Thomas, C. Grier, and D. M. Nicol. Unfriendly: Multi-party privacy risks in social networks. In *Proceedings of PETS' 10*, 2010.

[46] A. Yamada, T. H.-J. Kim, and A. Perrig. Exploiting privacy policy conflicts in online social networks. Technical report, CMU, 2012.

[47] J. Yang, K. Hua, Y. Wang, W. Wang, H. Wang, and J. Shen. Automatic objects removal for scene completion. In *INFOCOM Workshop on Security and Privacy in Big Data '14*.