A Semantic Framework for Content-based Access Controls

Sharon Paradesi[1]
paradesi@csail.mit.edu

Ilaria Liccardi[1,2]
ilaria@csail.mit.edu

Lalana Kagal[1]
lkagal@csail.mit.edu

Joseph Pato[1]
jpato@csail.mit.edu

[1]Computer Science and
Artificial Intelligence Laboratory
Massachusetts Institute of Technology
Cambridge, MA

[2]INRIA Saclay Île-de-France
Orsay, France

***Abstract*—Social networking sites provide role- or group-based access controls to help users specify their privacy settings. However, information posted on these sites is often intentionally or unintentionally leaked and has caused harm or distress to users. In this paper, we investigate possible improvements to existing implementations by introducing content-based access control policies using Linked Data. Users are able to specify the type of content in the form of tags or keywords in order to indicate which information they wish to protect from certain roles (for example employment), groups or individuals. Providing all possible keywords matching a specific topic may be too time consuming and prone to error for users. Hence using Linked Data we enrich the provided keywords by identifying other meaningful and related concepts. This paper presents the implementation and challenges of developing such a semantic framework. We have qualitatively evaluated this framework using 23 participants. Feedback from participants suggests that such a framework will help ease privacy concerns while posting and sharing social network content.**

***Index Terms*—privacy; access control; social network data;**

I. Introduction

Social networks are used by millions of people and have experienced an exponential growth in recent years [16]. While social media deals with sharing of content (such as status messages, pictures, and videos), not everyone is comfortable sharing all their information publicly. In fact, a recent Pew Internet report on social media sites [12] states that there is an increasing number of people who are setting more private access controls compared to the corresponding number in 2009. The study goes on to state that setting appropriate and effective privacy controls can sometimes be difficult to achieve. Keeping our social network persona private is not only dependent on what we decide to share and with whom (by restricting access using the available privacy settings), but also on what our friends share about us [11]. Further, studies conducted by Brandimarte et. al [4] highlight the paradox between access controls and privacy concerns. They observed that giving greater power to control the publication of information led to lowering of the participants' concerns over access and usage of that information.

Most of the current social networking sites provide users with the ability to describe roles or groups that may have access to their content (such as posts, photos, albums and so on). However, these controls can be considered *static* because users generally have to define new access controls for each data item created. We instead suggest *dynamic* access controls that augment the existing access controls by enabling users to specify tags or keywords to indicate the concepts in the content that they wish to protect. If a keyword matches certain content, that content should be suppressed during searches. We envision that these searches will be performed by organizations or individuals who are looking for information about a particular person for a specific intent. For example, it is becoming fairly common for employers to use Web search tools (such as SocialIntelligence[1], an FTC-approved commercial system) to gather specific information about a potential hire.

This paper presents a framework based on user generated policies to control access of social network data. Our framework constructs these policies based on the keywords provided by the users. However, entering long lists of keywords might be difficult for the users or they may fall short of expressing necessary keywords. Further, a syntactic match of the social network data to the keywords might not be adequate since we may miss many posts due to the variance in the syntax of the words (though the semantic meanings might be the same). Therefore we use Linked Data techniques to address the issue. Linked Data aims to link the resources on the Web and thereby form a graph of structured information [3]. We use these resources to enrich the provided keywords with meaningful and related concepts. Our framework utilizes the DBpedia Lookup service[2] to accept the input keywords from the users in the form of Linked Data terms (subject-predicate-object triples). These *semantically enhanced* keywords are then stored as policies. Before returning the social network posts in the search results, the policies are used to determine which posts to filter out based on individuals-, role-, or group-

[1]http://www.socialintel.com/
[2]https://github.com/dbpedia/lookup

978-0-7695-5137-1/13 $26.00 © 2013 IEEE
DOI 10.1109/SocialCom.2013.94

based access controls.

This paper makes the following three contributions:

1) A semantic framework that implements content-based access control policies for social networks using Linked Data. The framework makes use of the following two algorithms: (i) *SemanticEnhancement*, used to semantically enhance the keywords provided as input before creating a policy, and (ii) *DirectComparison*, used to apply the policy and filter the results of a search query.
2) An exploratory user study to understand the participants' perceptions and feedback about such a framework.
3) A discussion on the open challenges.

II. Related Research

Due to the lack of user awareness and proper privacy protection tools, a large amount of personal and sensitive information is being made accessible to authorities, strangers, recruiters and employers. Over one billion active users[3] share information on Facebook. It may be difficult for some users to create fine-tuned privacy policies, but there is a greater inherent difficulty in determining who the recipients of certain social media content should be. In other words, deciding which people or categories of people should have access to users' information is often cumbersome since it requires users to constantly manage their privacy settings or friend lists.

Every social network typically has some form of access controls mechanism to allow users to construct barriers around their data. The following works have investigated developing policies for such access control mechanisms. UPP (User Privacy Policy) [1] is an XML-based policy framework which incorporates the notion of access rights, reputation and the entities that can view data. Clifton *et. al.* in [7] propose a privacy framework for data sharing and integration by predicting the matches without revealing sensitive data and enabling querying across different data sources using semantic correspondences. *Privacy Preference Ontology (PPO)* is a fine-grained access control specification for Linked Data using the Web Access Control (WAC) vocabulary. Users would be able to specify their policies in PPO using a Privacy Preference Manager which can also enable the resolution of conflicting privacy preferences. Though role-based and group-based access controls are necessary privacy mechanisms, they may require constant maintenance and are not flexible. Content-based access controls, on the other hand provide users more flexible and dynamic protection since they only need to be created once for a particular concept or keyword.

Although social networks have in-built mechanisms for privacy of their users, they are inadequate to completely express the context in which the data subjects want their data to be viewed. Further, there are tools like SocialIntelligence which can be used by employers to search for information about an employee. If any data of the employee's social network profiles matches the employer's search criteria, SocialIntelligence notifies the employer of the same (possibly losing track of the context in which the data were specified). SocialIntelligence provides the same data gathering function as our framework for employers, but it does not have the notion of policies to protect the employees. An analogous commercial product to our framework is Reputation.com[4] which is a service aimed to replace malicious reviews with truthful, positive feedback. Though the specific implementation details differ, Reputation.com has the shared goal of giving users the ability to shape how others view information about them on the Web.

A recent Pew study [12] shows that the default access control settings are not sufficient to ensure privacy in social networks. Systems like Privacy Watch [2] and Secure Vault [14] enhance the current access control mechanisms to provide better privacy guarantees. Privacy Watch partially monitors the dissemination of information of the user and also proposes cryptographic techniques to ensure a trace that users can use for legal reasons in case of unauthorized access. Secure Vault addresses the concepts of data dislocation, fake information and encryption. Our framework does not provide false results to search agents because a fake social network post could possibly end up harming the reputation of the social network user compared to the user's real social network data.

Other frameworks that provide content-based access controls are based on either machine learning techniques [8] or rule-based approaches that model trust [6], [5]. While Content-Based Access Controls [6] uses a tag-based approach along with supervised classification techniques, rule based trust frameworks [6], [5] use created rules to reason over the specified data. However social network data is varied and unique hence it may not be so conducive to static techniques which require predefined rules or training data. Our framework has the ability to dynamically enhance users' generated input (keywords) with other related concepts.

Recent studies have analyzed the failure of the existing access controls prevalent on social networking sites [13] and the design conflicts between privacy and usability [16]. Combating "insider threat" by members in role-based access control lists who are not the intended audience of the posted content pose great concerns [10]. Our framework aims to address these concerns by providing flexible content-based access controls that are not restricted by the current role-based access control mechanisms.

III. A Semantic Framework for Content-based Access Controls

In this section, we describe how the searches performed by our framework differ from those currently feasible on the Web. We present the overall framework and explain the two algorithms used – SemanticEnhancement and DirectComparison. As described in our earlier work [15], when performing searches using content-based access controls, the framework interacts with the following users:

[3]http://newsroom.fb.com/Key-Facts

[4]http://reputation.com

- **Data subjects** are users who post information on social networks in form of status updates, posts, notes, or images.
- **Data Consumers** are users (such as insurance agents or employers) who search social network data in order to obtain information about specific data subjects with specific intents.

A. Overall Workflow

In popular social networks such as Facebook, Google+, and Twitter, users post information about themselves and specify people or categories of people with whom that information can be shared.

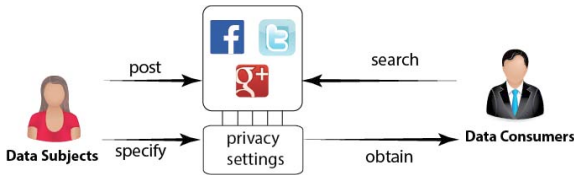

Fig. 1: *Workflow of performing search based on privacy settings in the current social networks.*

Data subjects can configure the privacy settings on their posts so that those posts are public and hence viewable by everyone. However, if they choose to disclose certain information only to selected recipients, different social networks provide different ways to achieve this goal. Facebook allows its users to (i) approve and add "friends", and (ii) create custom friend lists that can be used to set the audience for their information. Google+ has a similar approach to Facebook using a "circle" as the means to collect recipients. A user can either set the post to be public or disclose it to a particular set of circles when publishing the post. Twitter employs a simpler model, people can follow and thereby view the tweets of any public user without needing prior approval. However, they need the approval of a user with a private profile to be a "follower" and thus view that user's tweets.

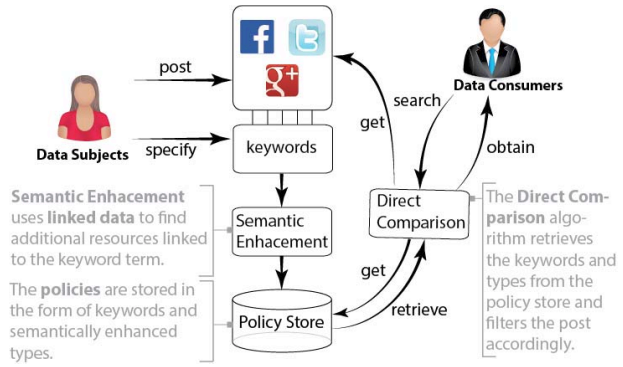

Fig. 2: *Workflow of performing search using content-based access controls through our framework. The data consumer can only obtain results that reflect the data subjects' disclosure specifications.*

When a data consumer looks for information about data subjects using the current search capabilities available in these social networks, they are shown the resulting posts based on the privacy settings placed by the data subjects on their respective profiles. This workflow is shown in Figure 1.

Our framework augments existing privacy settings in the social networking sites by using policies created by data subjects. The overall workflow (shown in Figure 2) using the framework is as follows:

1) Data subjects post information directly on the social networking sites.
2) Data subjects specify keyword(s) as input.
3) The keywords are then used by the *SemanticEnhancement* algorithm to gather types of Linked Data terms that are conceptually related to these keywords. The resulting type, along with the input keywords, constitute the policies of the data subjects.
4) The policies are then stored in a centralized *Policy Store*.
5) Data consumers conduct a search using the framework for information about a specific data subject's posts.
6) The framework gets the posts related to the search query from the data subject's social network profile. These posts are accessible to the data consumer according to the access controls specified by the data subjects.
7) The framework consults the policies that were specified by the data subject and stored in the policy store.
8) Based on the policies, the framework displays only the information that the data subject allows to be shown.
9) Finally, data consumers obtain the results of their search query.

B. Semantic Enhancement of Policies

The SemanticEnhancement algorithm uses Linked Data techniques to gather terms related conceptually to the keywords entered by the data subjects during the creation of policies. This is necessary to enable data subjects to enter a few keywords instead of a long list of words. Without this ability, the process of entering keywords would be cumbersome and prone to errors. The pseudocode of the SemanticEnhancement algorithm is provided in Figure 3.

A data subject types the input keywords into our framework which then looks up the Linked Data term equivalent of the keyword using the DBpedia Lookup service. It then presents the data subject with a list of DBpedia resource labels who can select one of the suggested terms. The data subject can then choose to enter more keywords and thereby more terms into the form. The terms entered by the data subject are sent as input to the SemanticEnhancement algorithm. For each resource in this list, the SemanticEnhancement algorithm first identifies other resources linked using the *"sameAs"* relation on DBpedia. These identified resources are then looked up on publicly available endpoints to determine if additional linked resources can be found. The algorithm then fetches

(using automatically constructed SPARQL queries) the types of the resources identified at the endpoints. This list of newly-identified types is shown to the data subjects who can then select one or more types.

The resources associated with those types are then looked up at the endpoints and the types of those identified resources are then collected and shown to the data subjects. The list of types selected using this process is stored along with the input keywords as a *policy* in the policy store.

Algorithm III.1: SEMANTICENHANCEMENT(R)

```
for each r ∈ R
  do { S = (SAMEAS(r, "http://dbpedia.org/sparql"),
       MAPPING(TYPE(r)))
for each r ∈ R
  do { for each (s, e) ∈ S
       do {T = T ∪ TYPE(s, e)
       //till no more selections
       show T to user
       for each t ∈ T selected by user
       do { {s1} = SELECT(s, e)
            for each s' ∈ {s1}
            do { if s' ∉ R
                 do { then { R = R ∪ s'
                             T = T ∪ TYPE(s', e) } } } }
return (T)
```

Fig. 3: *SemanticEnhancement(R): The algorithm traverses linked data at the endpoint (e) using the list of initial resources (R). This procedure returns a list of types (T) that the user can select from.*

Example: Upon entering "Diabetes" and selecting the corresponding Linked Data term, the framework would show the following types identified by the endpoint (for example, from linkedlifedata.com[5]) to the user: *Disease or Syndrome*, *Pathologic Function*, *drugs*, *Biologic Function*.

Our framework differs from the automated Linked Data traversal approach [9] by allowing users to control the traversal process. This process is repeated as long as the data subjects continue to select the newly-identified types shown to them until there are no new types found in the endpoint. This flexibility for the user to be in control of how the Linked Data graphs are traversed is not available in the automated process [9]. The benefits of involving the data subjects in the traversal process is that they are in control of which types of concepts are included in their policy. This involvement would lead them to more fully appreciate how their policy is enhanced. Further, unnecessary branches (corresponding to the types not selected by the data subject) are not traversed, leading to more efficient traversal. The downside to this approach is that branches leading to potential useful concepts might not be traversed if the data subject does not select the corresponding types. Thus, the framework might not have access to related information because of the cognitive load on the data subject.

[5]http://linkedlifedata.com/sparql

C. Direct Comparison and Filtering of Posts

The DirectComparison algorithm uses the semantically-enhanced policies when performing the searches initiated by data consumers. It takes three inputs: (i) a list of the data subject's posts (P) that match the search query and that are accessible to the data consumer through the existing access controls, (ii) a list of the semantically-enhanced keywords (R) obtained from the data subject's policy, and (iii) the endpoint (e) corresponding to the location where additional resources can be found. The pseudocode of the DirectComparison algorithm is provided in Figure 4.

Algorithm III.2: DIRECTCOMPARISON(P, R, e)

```
for each p ∈ P
  do { for each r ∈ R
       do { for each n ∈ p //n = noun phrase
            do { if SELECT(n, r, e)
                 then { remove p from P
                        break //next post
return (P)
```

Fig. 4: *DirectComparison(P, R, e): The algorithm filters a post if any of the noun phrases in the post matches the content of the resources mentioned in the policy. This procedure returns a list of residual, unfiltered posts.*

When the search query is applied to the collection of the data subject's posts, a subset of posts that match the query is initially identified. This subset of matching posts is sent as input to the DirectComparison algorithm along with the semantically-enhanced keywords in the policy and the endpoint used by the SemanticEnhancement algorithm. The DirectComparison algorithm identifies the noun phrases in the input posts and compares those phrases to the content of the resources mentioned in the policy. If the comparison results in a match, the corresponding post is removed from the list. The residual list of unfiltered posts is then returned to the data consumer.

Example: Suppose that the data consumer executes a search query for the term "medicine" within a specific data subject's profile. The data subject explicitly mentioned "Diabetes" while creating the policy. The data consumer will see all the posts that contain "medicine" with the exception of the posts containing "Diabetes". Suppose that the data subject selected the type *drugs* during the SemanticEnhancement process. The framework would also filter out the posts mentioning medications used to treat to diabetes (for example, Metformin). Using Linked Data techniques, the data subject did not have to specify Metformin (or any other medication) when creating the policy.

IV. EXPLORATORY USER STUDY

We conducted an exploratory user study to understand the perceptions people will have using such a framework. In order to understand potentially sensitive topics, we conducted an

experiment using Amazon's Mechanical Turk[6] to understand potentially "sensitive" topics. Based on the feedback from the study we selected three topics – "diabetes", "heart attack", and "football".

A. Methodology

Participants: We recruited 23 participants (ten female and thirteen male) for this study. Twelve users were between 20 and 25 years of age, six were between 26 and 30, three were between 31 and 35, and, two were between 50 and 70. Eighteen participants were graduate students, one was a software consultant, one was a stay-at-home parent, one was a professor, one was an undergrad and one was a research scientist. All but one participant declared that they actively used social networking sites.

Procedure: The participants were asked to rate 15 posts using a 5-point Likert scale (1 = very sensitive; 2 = moderately sensitive; 3 = neutral; 4 = not sensitive and 5 = indicating uncertainty). The posts were a mix of public Facebook status and wall messages covering two medical topics (diabetes and heart attack) and one sport topic (football). The proportion of posts with respect to the chosen topics was – 7 (diabetes), 5 (heart attack), and 3 (football).

The participants and our framework (using the SemanticEnhancement and DirectComparison algorithms that were described in section III) rated all 15 posts independently. The participants were then shown the posts for which their rating differed from the framework's rating. For posts that were rated as sensitive by the framework but not by the user, a justification was shown to the participant. The participants were given an option to change their ratings and were asked to provide feedback about their response to the rating given to the post by the framework. At the end of each survey we conducted a semi-structured interview for 15-30 minutes in order to obtain the user's feedback on the potential usefulness of such a tool, possible improvements that they would like to see, and general comments and suggestions.

B. Results

Survey: Regarding sharing their own information or viewing friends' information on social networking sites, 13 participants explained that such a framework could be useful in understanding the possible damaging effects of exposing their own post or their friends.

> [Participant 22]: *"If I weren't aware of the implications of the drug, the information could be damaging to my credit, for example to a potential employer."*

> [Participant 4]: *"If talking about diseases of other people, he/she might get offended. It might not be polite to leak the privacy of other people."*

Our framework provided users with the ability to understand the justifications for why a post was considered "sensitive". 14 participants highlighted the importance of this understanding. For example, we received the following quote:

> [Participant 5]: *"The URL(s) kind of gave me "YIELD" signs and helped me think twice about the sensitivity of mentioning the word "cholesterol.""*

Overall, the participants highlighted the usefulness of using such a tool to understand the possible implications of what they might post on social networking sites.

Semi-structured Interview: The overall feedback underlined the need to be cautious in order to preserve the privacy of their information on social networking sites. Having a system that flags posts as "sensitive" would allow users to become aware about what type of information they disclose about themselves or their friends. In particular, two participants mentioned that our framework could be used as a "cautionary tool" or a "moral compass". It can be used to alert users to think twice before posting certain type of data (for example, medical information). Further, it can provide feedback and help users understand the consequences of sharing that information in different contexts. All 23 participants indicated the usefulness of such a tool. However they expressed wanting to make the final decision indicating wether the post is sensitive or not. Specifically,

> [Participant 3]: *"assuming it is not given the final say, ... being more sensitive is more valuable ... because it is very hard to unsay something"*

> [Participant 9]: *"For me, I would rather it be over cautious than under cautious"*

> [Participant 13]: *"sometimes [an over cautious system may] get annoying ... I'd rather get less information than more information"*

> [Participant 22]: *"helpful when typing something really quickly ... good to give a second warning... caveat is that maybe it is annoying over time. In my mind, I should know what is sensitive to me better than the system. Unless it is something where the system thinks it is useful to be more public."*

V. Open Challenges

Our framework provides a way to incorporate content-based access controls when performing searches for information about data subjects. However, during the design, implementation, and evaluation of our framework, we came across the following open challenges.

A. Incentives

Data consumers already have other search mechanisms available online. For example, a data consumer could easily (i) perform a search using Google, Bing, or other search engines, or (ii) use a commercial product like SocialIntelligence to find out information about a particular data subject. The incentives

[6]https://www.mturk.com/

that would convince a data consumer to adopt such policy-aware searches are not readily apparent. Along the same lines, although such a framework can be directly implemented in social networking sites, none of the leading social networking sites currently provide their users with the ability to create such content-based access controls. Thus, incentivizing the social networking sites to adopt this technology is another open challenge.

B. Data Context

It would not be wise for a data consumer to come to strong conclusions about a data subject solely based on the results returned by any search tool. This is because there could have been additional clarifying posts in the data subject's profile or other contextual information but not been returned in the search results. This may happen if those additional posts (i) did not match the search criteria, (ii) had different privacy settings, or (iii) were not returned due to technical issues with the API or the search tool implementation.

C. Shareability

On one hand, there are instances where one may want to explicitly share their information. For instance, job seekers and employees seem to prefer to extensively share their work experiences and professional accomplishments on LinkedIn. On the other hand, the naive response to privacy concerns about leakage of information on social networking sites is to hide everything (from search agents). In our semi-structured interviews, the majority of the participants indicated that they would prefer a tool that was over cautious. Balancing these two views on sharing of information requires further research.

VI. Conclusion

Most social networking sites have implementations to control access to their users' data through predefined settings. However, those users do not have means to specify content-based controls for their information. In this paper, we presented a semantic framework to augment the existing access control implementations on social networking sites with content-based access control policies created with the help of Linked Data techniques. We described the following differences between the privacy settings offered by the current social networking sites and our framework: how to specify the access controls, how searches can be performed, and how results are filtered out in these two settings.

The two algorithms used in the framework – *SemanticEnhancement* and *DirectComparison* were then presented. SemanticEnhancement is used to enhance the keywords provided by the data subjects by finding conceptually-related Linked Data terms. The semantically-enhanced policies are then stored in the Policy Store of the framework. DirectComparison is used to perform the searches conducted by data consumers in a policy-aware manner.

We then presented our exploratory user study which investigated the participants' perceptions of our semantic framework. All participants highlighted the usefulness of such a tool in order to understand the possible damaging effects of their posts containing "sensitive" information. In particular, they highlighted the fact that such a framework would have the ability to dynamically alert them about "sensitive" posts before posting them on social networking sites. Finally, we discussed the major open challenges facing the implementation of such a framework, in particular: *Incentives*, *Data Context*, and *Shareability*.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant Number 1228687. Ilaria Liccardi was supported by the European Commission Marie Curie International Outgoing Fellowship grant 2011-301567 *Social Privacy*.

Our thanks to Oshani Seneviratne, Fuming Shih, Hal Abelson and other members of the Decentralized Information Group at CSAIL for comments and suggestions on the research.

References

[1] E. Aimeur, S. Gambs, and A. Ho, "Upp: user privacy policy for social networking sites," in *Fourth International Conference on Internet and Web Applications and Services, 2009. ICIW'09*. IEEE, 2009, pp. 267–272.

[2] E. Aimeur., S. Gambs, and A. Ho, "Towards a privacy-enhanced social networking site," in *International Conference on Availability, Reliability, and Security (ARES) 2010*. IEEE, 2010, pp. 172–179.

[3] C. Bizer, T. Heath, and T. Berners-Lee, "Linked data-the story so far," *International Journal on Semantic Web and Information Systems (IJSWIS)*, vol. 5, no. 3, pp. 1–22, 2009.

[4] L. Brandimarte, A. Acquisti, and G. Loewenstein, "Misplaced confidences: Privacy and the control paradox," *Social Psychological and Personality Science*, 2012.

[5] B. Carminati, E. Ferrari, R. Heatherly, M. Kantarcioglu, and B. Thuraisingham, "Semantic web-based social network access control," *Computers Security*, vol. 30, no. 2-3, pp. 108–115, 2011.

[6] B. Carminati, E. Ferrari, and A. Perego, "Enforcing access control in web-based social networks," *ACM Transactions on Information and System Security (TISSEC)*, vol. 13, no. 1, p. 6, 2009.

[7] C. Clifton, M. Kantarcioğlu, A. Doan, G. Schadow, J. Vaidya, A. Elmagarmid, and D. Suciu, "Privacy-preserving data integration and sharing," in *Proceedings of DMKD 2004.*, 2004.

[8] M. A. Hart, "Content-based Access Control. PhD Thesis. State University of New York at Stony Brook," 2006.

[9] O. Hartig and J. Freytag, "Foundations of traversal based query execution over linked data," in *Proceedings of the 23rd ACM conference on Hypertext and social media*. ACM, 2012, pp. 43–52.

[10] M. Johnson, S. Egelman, and S. M. Bellovin, "Facebook and privacy: it's complicated," in *Proceedings of the Eighth Symposium on Usable Privacy and Security*. ACM, 2012, p. 9.

[11] A. Lampinen, V. Lehtinen, A. Lehmuskallio, and S. Tamminen, "We're in it together: interpersonal management of disclosure in social network services," in *Proceedings of the 2011 annual conference on Human factors in computing systems*. ACM, 2011, pp. 3217–3226.

[12] M. Madden, "Privacy management on social media sites," *Pew Internet Report*, pp. 1–20, 2012.

[13] M. Madejski, M. L. Johnson, and S. M. Bellovin, "The failure of online social network privacy settings," 2011.

[14] S. Malik and A. Sardana, "Secure vault: A privacy preserving reliable architecture for secure social networking," in *7th International Conference on Information Assurance and Security (IAS), 2011*. IEEE, 2011, pp. 116–121.

[15] S. Paradesi, O. Seneviratne, and L. Kagal, "Policy aware social miner," in *IEEE Symposium on Security and Privacy Workshops (SPW), 2012*. IEEE, 2012, pp. 53–59.

[16] C. Zhang, J. Sun, X. Zhu, and Y. Fang, "Privacy and security for online social networks: challenges and opportunities," *Network, IEEE*, vol. 24, no. 4, pp. 13–18, 2010.