CS563: Advanced Computer Security Research Proposal

Cassandra Jacobs and James Smith
csjacobs2@illinois.edu
smith152@illinois.edu

October 11, 2016

Abstract

With the huge growth in popularity of social media accounts and applications, privacy and security of users' personal information and the content they post publicly exposes them to a wide unknown audience which can potentially cause serious problems in their personal and professional lives. Lack of understanding how privacy settings work or carelessness (caused by human error) has led to severe repercussions and damages to thousands of people. We propose a solution through a Chrome extension for Facebook that will allow users to define sets of rules to categorize people in their Friends List into smaller groups in order to allocate personalized privacy settings. This extension would then analyze the contents of a post, prior to being made public, and measure the contents against a rules dictionary and determine if the post requires a security setting (based on the rules pre-defined by the user). The user can then decide if they want to apply the security/privacy filter or allow it to be posted without. Alternatively, the extension could automatically apply the security/privacy filter without prompting for confirmation if the user chooses to set it up accordingly. The goal of this extension is to lower the risk of human error and increase the privacy and security of the contents of the posts made by a user by monitoring and intervening (or reminding) the user to apply the appropriate filter if the contents are flagged by the extension.

1 Introduction

Social media applications (such as Facebook) allow users to connect with people from around the world, giving them the ability to maintain connections to friends and family regardless of distance, as well as meet new people and form new friendships and relationships. These types of connections were difficult, if not impossible, thirty or more years ago. As a result, more personal information about users are ending up online in order to make people easier to find. This can create privacy concerns and open up the potential for abuse and leaks in information. Larger networks of friends increases the risk of sensitive information being revealed as well as potential conflict between connections a user may have in common, yet are not connected to each other (thus are unlikely to be friends or associates with each other).

Through the use of customized privacy filters, users can choose the exact audience they wish to share their posts with in order to minimize who has access to personal data and information (such as sharing personal photographs with only close friends and family), or filter out others who may take offense over hot-button issues (e.g. sharing a news article in support of a particular political candidate). Privacy filters allow users greater control over their content and their intended audience, yet filters are not always easy to use or remember, introducing the possibility of user error.[1]

We propose a Chrome browser extension that, upon the user submitting the contents to be posted to Facebook, would analyze the contents searching for key words that have been defined in a rules dictionary, apply a weighted score based on the rules, and if the score is over a set threshold, the user is queried if they want to apply the defined privacy filter to the content before posting it. Alternatively, the user can set the privacy filter to automatically apply (without querying) if the threshold is exceeded. This would decrease the risk of inadvertently posting sensitive or controversial topics without at least first reminding the user of the contents and privacy filter.

The filters themselves would require personalization by the user, while the rules defined dictionary could be preset with default settings of the more popular controversial topics. The user can then add, remove and edit more rules, as well as set a priority level (which rules take precedence over others in the case of multiple rules matching) and the threshold level (what scoring is required to trigger the filter when a rule is matched).

2 Background

For many users, the primary use of a social media account is to share personal news and updates, as well as links to articles and websites they have an interest in or support. Most of these types of posts encourage feedback in the form of "likes" or "pluses" and comments that can be left by friends of the poster. This is where privacy risks can arise when a user posts without any privacy set at all (globally visible to anyone), leaving the contents of the post and the comments readable by anyone who visits

the social media page. The user or commenters may inadvertently reveal information about themselves or have this information picked up by a stalker or harasser. Small bits of information about a person may not be inherently harmful on their own but compiled together over time into a larger profile, may create a more complete picture that can be used to cross-reference and accurately identify a target.[2]

Users may also have a large cross-section of friends and family they stay connected with through social media, with a large gap of beliefs, culture and backgrounds. It may not be reasonable or wise to allow all connections on a friends network to have read/write access to a post that may be controversial, nor would it be reasonable to expect a user to refrain from posting on topics he/she is interested in solely to cater to a few people on their list that might take offense or cause a disruption. Many social media accounts allow the creation and use of *filters* which users can customize by adding (or removing) friends to and using these as custom privacy settings which can be applied on a per-post basis. This way, the user is no longer constrained to the types or subjects of posts he/she would like to share, and they can ensure the correct audience is exposed to the subject, minimizing potential privacy issues and offensive/disruptive content.

One of the inherent problems with social media though is the lack of emphasis placed on security and privacy, as one of the goals of these types of applications is sharing content and connecting people. Increased privacy and being less visible globally actually works against the connectivity model. With over 1.7 billion users on Facebook[3], high security, high privacy options should be standard, with opt-out being an active action the user has to set instead of having to opt-in and set their privacy settings on or to higher levels. Setting up filters aren't always clearly defined and easy to use and either default to the last one used or have to be set each time a post is made. This can introduce human error, as it relies on the user to be vigilant and cognizant each time they make a post to check the privacy settings and ensure they are correct at that particular moment.

Our hope in a Chrome browser extension that can analyze the contents and flag the content for potential privacy leaks and controversial topics before the content is posted to Facebook is that it would lower the risk of human error and forgetfulness and allow users to post more freely without having to worry about whether they've set the proper privacy/security permissions or not.

3 Related Work

Previous work in the area of privacy within Facebook was explored by Ghazinour[2], who developed a machine learning tool, YourPrivacyProtector, in order to suggest privacy settings for users. The authors harvested data from Facebook profiles, primarily user profile information and privacy settings from photo albums, and based on the data compiled, used it to learn about the privacy settings already in place, and made recommendations on what other profiles should do based on similar data. This study focused primarily on the user's personal profile information.

The ability to select better, or more relevant, privacy settings, is explored by Naini[1]. In this paper, they approach a similar idea we are proposing, a method for predicting the privacy setting for a post at creation time, based on a set of features within the post itself. By allowing automatic privacy-setting assistance, they hypothesize a potential decrease in leaked information to the wrong audience and increased security and privacy for a user.

Ideally, we would like to build on the research developed by Naini et al, and attempt to develop the application that could predict and apply the recommended privacy setting for the user.

4 Proposed Approach

To achieve our goal in helping to protect the scope of a user's posted content on Facebook we propose a two pronged approach. The first prong of this approach is a Facebook application that will intelligently control the audience of a user's post based off of security guidelines set in place by the user. When a user downloads our chrome extension for the first time he/she will be prompted to perform an initial setup. The setup will consist of the user creating different mappings or rather rules that link a given Friend List (e.g. Family) to a given category and action (e.g. politics, do not share). The categories will be predefined to begin with but the application will allow for custom categories to be created by the user later on. The application will be using the Graph API provided by Facebook to grab the valid set of Friend Lists for the user[4].

Once a user has established an initial set of rules they will then be presented with a dialog box. That dialog box will allow them to type in the contents of a post. It will also display the set of Friend Lists that the post will be visible to, a button to validate that list, and a button to post the content. When the user clicks 'post' the content will be published out to the shown Friend Lists. That list will be generated by our algorithm upon the user clicking the validate button. The algorithm to generate that list consists of two main parts: a weight function that assigns categories and their given weights to the given content and a decision function that takes

the assigned categories and the user's rules to spit out a recommend set of Friend Lists. Here the categories for a post will be assigned using a dictionary that will be preloaded with words that are mapped to a given category.

The second prong is a set of guidelines that will outline best practices for the user to take when setting up security zones in Facebook through the use of Friend Lists. This will take the form of a set of rules to be given to the user that will be used to assess the risk of a given friend based off of the information disclosed on the friend's profile[2]. Facebook no longer allows this information to be collected through the APIs; therefore this piece of our research has become a more manual process for the user. The guidelines will also include suggestions for different over-arching social privacy zones that are clearly defined and will get the user up and running quickly. The Friend Lists defined here will then be used by the application described above when a user is requesting advice for the specific Friend Lists a post should be sent to.

5 Research Plan

The overall goal of this research is to establish a mechanized way to intelligently protect the dissemination of a user's Facebook posts to a specific group of friends that is visible and transparent to the user in a simple and easy to understand fashion. Through the use of our application the user will be alleviated from having to tediously choose what friends a given post will be visible to and can instead follow the recommendations provided to them by the app. The application will provide a clean and easy to read interface that gives full transparency to the user on who his/her posts will be visible to. This functionality will help keep the posted content of a user confined to parties that have been verified as trustworthy for the given category of the post's content.

The second goal of this research is to provide a clear recommendation for users to use when setting up Friend Lists in Facebook and to provide them with the tools needed to use these lists as personal privacy security zones. The goal here is to educate users on how to be more conscientious of their social information and how to prevent privacy leaks into the social web. This will accomplished through the startup rules that will be laid out in the guidelines for creating Friend Lists (security zones) and how to best use these lists when posting content.

The two goals above have been presented separately but in all actuality they are closely related. The efficacy of the primary goal is very much dependent on the successful setup of meaningful and accurate Friend Lists. That is why making the manual part of our research process easy to implement fort the user is important. We would like to automate this process but due to restrictions on the information that is publicly available from the Facebook Graph APIs this is not feasible at the current time. The guidelines that we will lay out could later be considered for developing an automated protocol for bucketing users into security zones.

References

[1] Kaweh Djafari Naini et al. "Analyzing and predicting privacy settings in the social Web". In: *International Conference on User Modeling, Adaptation, and Personalization*. Springer. 2015, pp. 104–117.

[2] Kambiz Ghazinour, Stan Matwin, and Marina Sokolova. "YOURPRIVACYPROTECTOR, A recommender system for privacy settings in social networks". In: *arXiv preprint arXiv:1602.01937* (2016).

[3] Statista. *Number of monthly active Facebook users worldwide as of 2nd quarter 2016 (in millions).* `https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide`. [Online; accessed 10-October-2016]. 2016.

[4] Jesse Weaver and Paul Tarjan. "Facebook linked data via the graph API". In: *Semantic Web* 4.3 (2013), pp. 245–250.