Privacy Behaviour and Profile Configuration in Twitter

Taraneh Khazaei*
University of Western Ontario
tkhazae@uwo.ca

Lu Xiao*†
University of Western Ontario
lxiao24@uwo.ca

Robert Mercer*
University of Western Ontario
mercer@csd.uwo.ca

Atif Khan
InfoTrellis Inc.
atif@allsight.com

ABSTRACT

The dichotomy between users' privacy behaviours and their privacy attitudes is a widely observed phenomenon in online social media. Such a disparity can be mainly attributed to the users' lack of awareness about the default privacy setting in social networking websites, which is often open and permissive. This problem has led to a large number of publicly available accounts that may belong to privacy-concerned users. As an initial step toward addressing this issue, we examined whether profile attributes of Twitter users with varying privacy settings are configured differently. As a result of the analysis, a set of features is identified and used to predict user privacy settings. For our best classifier, we obtained an F-score of 0.71, which outperforms the baselines considerably. Hence, profile attributes proved valuable for our task and suggest the possibility of the automatic detection of public accounts intended to be private based on online social footprints.

Keywords

Social Privacy; Profile Attributes; Twitter Analysis

1. INTRODUCTION

The increasing levels of engagement in online social media have led to the accumulation of large social footprints left by millions of users on a daily basis. This massive source of information can deliver relevant information in the right context, leading to tremendous opportunities for both businesses and individuals. However, to effectively harness this treasure trove of data, it is imperative to address possible privacy complications. Such privacy issues are especially concerning due to the disparity between users' privacy behaviours and their attitudes in social media [6, 20], making their current privacy settings unreliable. As such, methods that can detect users' privacy preferences are desired so that data related to the privacy-concerned users can be discarded.

This privacy dichotomy may be due to users' misconceptions regarding the visibility of their data [20], the complex privacy specification interfaces [4], as well as their false, yet common, perception of the default setting as the recommended privacy policy [12]. In addition, even privacy-aware users may decide to choose public settings for the anticipated social gain, while they may not be willing to be profiled online by business and companies.

We argue that user *social footprints* [7] in social media environments can characterize their privacy preferences, offering an alternative and reliable source for the detection of privacy preferences. The social footprints are available in three types of social media data: users' profile attributes, their social context and ties, and their published content. In this study, we focus on the analysis of the profile attributes to explore their potential links to the user privacy preferences. In particular, we analyzed profile attributes of Twitter accounts to examine whether people with different levels of privacy setting configure these attributes differently.

Privacy configuration in Twitter is relatively simple and follows a binary specification. The Twitter users can follow the default *public* setting, which indicates that their tweets and follower/friend lists are accessible by the public. Alternatively, they can change the setting to *protected*, which makes their tweets and follower/friend lists accessible only by their approved followers. It is noteworthy that the users' profile attributes are visible to and accessible by the public and the Twitter API regardless of their privacy settings.

We analyzed a set of users' profile features and descriptions that are readily available from their Twitter accounts. We also developed and analyzed three additional features based on the existing profile attributes. Based on the analysis results, a feature set is developed and utilized in multiple classifiers to automatically detect the users with the protected privacy setting. Compared to the users' social network structure and their content-related features, their Twitter profiles contain very limited information. Despite this limitation, our classifier has obtained an F-score of 0.71, which improves a random and a naive baseline by over 20%. This finding can have implications for designing privacy-preserving personalization tools and indicates the value of profile attributes in the detection of privacy-concerned users.

The remainder of this paper is as follows: Section 2 reviews the earlier studies on privacy prediction in social media and the user attribute classification in Twitter. Section 3 describes our data collection process and the profile features

*Department of Computer Science
†Faculty of Information & Media Studies

WWW'16 Companion, April 11–15, 2016, Montréal, Québec, Canada.
ACM 978-1-4503-4144-8/16/04.
http://dx.doi.org/10.1145/2872518.2890088 .

in our dataset. Section 4 explains the analysis of the profile attributes for the users with the protected and default privacy settings. Section 5 presents the selected feature set and the evaluation of our classification. Finally, the paper concludes with our contributions and research plans.

2. RELATED WORK

Detecting privacy attributes in social media. To address users' privacy concerns, three main approaches have been reported in the literature: the use of privacy-enhancing principles for designing personalization systems [9, 21], the design of usable interfaces and visualizations that enable users to specify their privacy policies [11, 1, 5], and the computational methods that automatically predict users' privacy preferences [18, 4]. Regardless of its potential value, the prediction approach has received much less attention compared to the other two [8]. Additionally, the majority of the prediction models are structured on the users' social networks and their generated content, while only a few have utilized profile attributes and personal characteristics [8].

Minkus and Memon [13]conducted an online questionnaire study to examine the privacy settings of Facebook users and related the settings to their demographic and personality features. Their survey results were later used to build and deploy an online application, called MyPrivacy, that automatically recommends privacy settings. Specifically, MyPrivacy first asks multiple questions from users to determine the demographic and personality attributes and then uses a supervised machine learning algorithm to make recommendations based on the users' privacy settings. The evaluation of MyPrivacy showed that real Facebook users had positive subjective opinions toward the tool. However, this tool is semi-automated and requires direct input from the users.

Similarily, in [14] a supervised learning algorithm is proposed and is built on a large set of features to recommend privacy settings. These features include metadata elements regarding a shared item as well as users' demographic and profile features, such as the number of users' Facebook posts and their friends. It is worth noting that the algorithm is developed to recommend privacy settings for a particular shared item, such as individual Facebook posts, as opposed to predicting users' privacy preferences in general.

Dong et al. [3] proposed a privacy prediction model that takes into account social media behavioral analogs to psychological variables that are known to affect users' disclosure behavior. Some of the identified analogs are based the profile features of the users. For instance, user's trustworthiness is calculated based on the ratio of their followers to the total number of their social contacts.

Twitter user attribute detection. A variety of techniques and social data types have been utilized to detect Twitter users' latent attributes. For instance, [16] uses the profile fields, tweeting behaviour, tweet content as well as the network structure to understand political affiliation, ethnicity, as well as users' affinities to a specific business.

Another example is the work of Nguyen et al. [15], in which the connection between the language use and the age is studied in the Twitter context. The authors built a classifier based on the tweet unigrams and classified users according to their age category, life stage, and their age. In their study, Rao et al. [17] found distinctive variations in the language use of Twitter users across different gender, age, regional origin, and political orientation.

These studies on automatic detection of privacy preferences and Twitter user attributes provide valuable insights into the potentialities and limitations of the available features and techniques in the detection of users' latent attributes. Our purpose for using the profile attributes to detect users' general privacy preferences has received very little attention in the literature [3]. Although it is argued that the users' profile fields may not include enough good quality information for user classification purposes [16], our results suggest that they can be promising in the detection of social privacy.

3. DATA SOURCE

3.1 User Selection

We have built a directory of Twitter users by crawling a number of famous Twitter accounts and collecting their followers. Table 1 presents these Twitter accounts, the number of collected followers each account has, as well as the percentage of the number of protected followers to the number of collected followers of the account. Please note that in the table the numbers are rounded (e.g., the number of CNN followers is 12,246,514 and we rounded it to 12.2M), and the percentages are calculated based on the exact numbers and then rounded.

It is known that a large number of Twitter accounts are inactive thus are more likely not to follow any account [10]. In addition, such accounts are more likely to follow the default public privacy setting compared to the active accounts. Our underlying set of Twitter accounts follow at least one account (e.g., follow "Bill Gates" as shown in Table 1). Therefore, the percentage of protected accounts is anticipated to be higher than that of Twitter accounts in general. On average, 4.8% of Twitter users have protected accounts [10]. As can be seen in Table 1, the percentage of the protected accounts is similar or above the average for all our follower sets, which confirms our expectation.

As shown in the table, the percentage of the protected accounts for "CNN Breaking News" is 11%, considerably higher than the other follower sets and the average percentage in Twitter. One possibility is that CNN tweets may cover more topics that attract private users in Twitter, such as privacy-related news, compared to the other accounts in the table. This might also be true for other Twitter news accounts as we expect the news accounts to cover a wider variety of the topics that are related to and have a potential impact on the users' lives and societies. Further analysis of the percentage of the protected Twitter accounts that follow other Twitter news accounts and the analysis of the news content are expected to offer more insights on this issue.

As mentioned in the introduction section, the users' privacy settings may not reflect their actual privacy preferences [6, 20]. In other words, Twitter users who have the default public privacy setting may, in fact, be private. Therefore, these users' profile attributes reflect the private users' configurations. To study the possible differences in how protected and public profiles are configured, we need a set of

Table 1: A set of popular accounts in Twitter and the statistics of their collected follower sets.

Account	#Followers	#Protected	%Protected
Facebook	4.8M	261K	5%
CNN Breaking News	12.2M	1.5M	11%
Youtube	14.8M	788K	5%
Bill Gates	5M	374K	7%
Obama	23.8M	52M	7%
Katy Perry	75.3M	52M	7%

that belong to key individuals and brands are marked as *verified*. To focus on the general public, we removed these *verified* accounts from the set. Finally, we filtered the set to include only those accounts whose language is set to English. By applying these three criteria, we were able to select roughly 850K protected accounts from the original 1.5M accounts. Our public accounts also dropped from 12.2M to almost 10M. To have a relatively balanced set of accounts, we randomly pulled 1M public accounts from this set.

3.2 Profile Features

Each Twitter account is associated with a set of profile attributes. A set of profile features is configured by the account holder, often when the account is created, and is intended to represent who the user is in the network. Examples of such profile attributes include the username, the profile image, and the location information. Another set of attributes, which are also specified by the user, is related to the settings of their account. For instance, they can specify whether or not they want their tweets to be geo-tagged by setting the value of the geo-enabled attribute. Other examples of such attributes include their preferred interface language and whether or not their account should be withheld from certain countries. Finally, a set of contextual attributes is specified by Twitter. For instance, the time of the account creation, the number of tweets published by each user, and the number of followers/friends.

A subset of the available profile attributes is deemed to be relevant for our purpose and selected in our analysis. This list is shown below, along with a brief description of the attributes. Please note that the descriptions are adapted from Twitter API specification[1].

- Name: The name of the user.
- Username: The alias that users identify themselves with.
- Description: A piece of text users provide to describe their account.
- URL: A URL provided by the user in association with their profile.
- Location: The user-defined location for this account's profile.
- Geo-enabled: When true, indicates that the user has enabled the possibility of geotagging their Tweets.
- Default Image: When true, indicates that the user has not uploaded their picture and a default avatar is used instead.
- Default Profile: When true, indicates that the user has not altered the theme or background of their user profile.
- Favorite Count: The number of tweets this user has favorited in the account's lifetime
- Tweet Count: The number of tweets issued by the user.
- Follower Count: The number of followers this account currently has.
- Friend Count: The number of users this account is following.
- List Count: The number of public lists that this user is a member of.

[1]https://dev.twitter.com/overview/api/users

4. ANALYSIS OF PROFILE ATTRIBUTES

In our analysis, the geo-enabled attribute, default profile, and default image are all binary attributes and are studied as binary variables. Similarly, the numeric attributes of the favorite count, the tweet count, the follower/friend count, as well as the list count are analyzed as is.

Based on the declared name in the Twitter account, we created a binary and a numeric attribute: we matched the account name against a directory of English names to check whether any part of their declared name is indeed a person's name in the dictionary. We also counted the number of parts in the account name that are available in the dictionary. For example, an account name that has only the first name matched has the value of 1, whereas an account name that has both the first and the last name appearing in the dictionary has the value of 2. For the Twitter account's username, we checked to see if it contains the declared name of the user. For description, URL, and location attributes, we simply checked whether the corresponding piece of information is provided by the user.

Finally, we used a linguistic analysis tool to study the account's profile descriptions to understand how the users of different privacy settings describe themselves in Twitter. The analysis results of the surface-based profile features are provided in Section 4.1, while Section 4.2 explains the results of the linguistic analysis of the profile descriptions.

4.1 Surface-based Profile Features

Table 2 presents the selected binary features, along with the percentage of the protected and public accounts for which these binary attributes hold. Although the Chi-Square test results suggest statistical significance for all the features, the effect size values suggest that only three features have practically different values in the public versus the protected accounts: *has location*, *is geo-enabled*, and *is default profile*. We calculated the effect size using Cramer's V and followed the convention to interpret the value [2]. A Cramer's V needs to be at least .1 to show a practically significant effect in reality. As shown in the table, a larger percentage of protected accounts has enabled their geo-tagging feature and has provided information for the location attribute. Besides, more protected accounts have changed their default profile settings compared to the public accounts.

Table 3 provides an average value of our numeric features in the two types of accounts. We calculated the effect size using Cohen's d, and followed the convention to interpret the value [2]. Specifically in our study context, a feature's

Table 2: Analysis of binary profile attributes of protected and public accounts.

Binary Attributes	%Protected	%Public	Effect Size
Has Name	71.17	68.86	0.02
Username Has Name	3.01	3.45	0.01
Has Description	56.79	51.69	0.05
Has URL	15.01	16.78	0.02
Has Location	64.70	49.05	0.15
Is Geo-enabled	39.78	25.59	0.15
Is Default Profile	33.26	71.17	0.38
Is Default Image	6.43	8.80	0.04

Table 3: Analysis of numeric profile attributes of protected and public accounts.

Numeric Attributes	Protected	Public	Effect Size
Favirote Count	189.32	115.43	0.09
Tweet Count	1389.16	384.55	0.29
Follower Count	80.71	166.78	0.03
Friend Count	255.78	242.76	0.03
List Count	1.01	0.93	0.0006
Name Count	1.10	1.07	0.04

Cohen's d value needs to be at least .2 to be considered as a practically useful feature that distinguishes the protected and public accounts. Although the t-test results suggest statistical significance for all the features, the effect size values suggest that only the Tweet count feature has a practically different value in the public versus the protected accounts. The results show that on average, protected accounts tweet more often and this feature's effect is close to medium (d = .29) (see Table 3). The protected account seems to have a larger number of favorite tweets although the effect is still quite small (d = .09).

In general, the results are interesting and contrary to what we expected before the analysis. For example, we anticipated that because the protected accounts represent a more private or more privacy aware population, they would be less likely to enable the location tracking feature or change the default profile theme, or even tweet often. These findings, however, indicate otherwise.

4.2 Profile Descriptions: A Closer Look

As explained earlier, the Twitter users can provide up to 160 characters in the description field. In our set of the CNN followers, there are almost 500K of the protected accounts and roughly 500K of the public accounts that have descriptions. We used Language Inquiry and Word Count (LIWC 2015) to analyze the language categories in these descriptions. The LIWC program processes each text file word by word and compares them against a pre-built dictionary to detect the LIWC category that the word belongs to. After processing all the words in the text, LIWC calculates and outputs the percentage of each LIWC category. Before conducting the linguistic analysis by LIWC, we applied the following pre-processing steps on the descriptions:

- removed HTML characters
- replaced apostrophe elisions (e.g., I'm -> I am).
- replaced URLs with the word "url"

Table 4: LIWC categories and their corresponding percentage for protected and public descriptions.

LIWC Category	Protected	Public	Effect Size
Function Words	37.51	33.68	0.15
Affect	8.68	7.50	0.10
Social Processes	11.08	10.75	0.02
Cognitive Processes	7.86	6.71	0.09
Drives and Needs	11.01	11.38	0.02
Relativity	10.03	10.23	0.0006

- replaced emoticons with their corresponding meanings (e.g., :) -> smile)
- removed punctuation marks
- replaced user handlers with the word "mention"

The LIWC dictionary is structured in a hierarchical format, wherein each category may encompass several subcategories. Details about these categories can be found in the LIWC website [2]. Since the users' profile descriptions are usually very short (commonly between 8-10 words), the percentages provided by LIWC are often very small for the majority of the categories. Therefore, we only focused on the higher-level categories that are at the top of the LIWC hierarchy.

Table 4 provides these categories as well as their corresponding percentages for the protected and public accounts. Here, we dropped those LIWC categories that had less than 5% of matching words in the entire corpus of descriptions. In addition, LIWC outputs a set of summary dimensions along with the percentage of their matching words. Table 5 provides the summary variables deemed relevant and their corresponding percentages for the two sets of accounts. A t-test is performed for these categories, along with the effect size measured by Cohen's d. All the categories have statistically significant different values between the protected and the public accounts, but these differences are small based on the Cohen's d (see Table 4). It is still interesting to note that the protected account has a larger percentage of the *function words* and *a ect words*, which being similar to our findings regarding the surface-based attributes is in contrast to our prior expectation.

In addition to the LIWC main categories, an analysis of the summary dimensions shows that protected accounts contain a smaller number of lengthy words (i.e., words with six or more letters). They use fewer words representing *analytical thinking* and *clout*. However, they have a higher percentage of words that bear *emotional tone* and *authenticity*. The differences are statistically significant based on the t-test results, but are not practically significant from the Cohen's d value (see Table 5).

5. USER CLASSIFICATION

We utilized the binary and numeric features introduced in Section 4.1, along with the LIWC features discussed in Section 4.2, in multiple classifiers to identify protected accounts. The classifications are conducted by the machine learning toolkit Weka [3] and the results are evaluated using stratified 10-fold cross validation.

[2]http://liwc.wpengine.com/

[3]http://sourceforge.net/projects/weka/

Table 5: LIWC summary variables and their corresponding values for protected and public descriptions.

Summary Dimension	Protected	Public	Effect Size
Six Letter Words	22.73	26.69	0.14
Analytical Thinking	75.64	84.04	0.15
Clout	66.83	72.99	0.10
Emotional Tone	98.58	96.55	0.05
Authentic	28.89	21.24	0.13

For the classification, we modified some of the LIWC features. We changed the LIWC categories that matched less than 25% of the entire corpus to binary attributes (see Table 4 for category percentages), such as *a ect*, *social processes*, and *drives and needs*. For instance, if a description contains any word that is categorized as an *a ect word* in LIWC, the corresponding feature is set to 1; otherwise, it is set to 0. The remaining LIWC attributes are kept as is.

We also added the presence of four keywords as supplementary features. Throughout our keyword analysis of the descriptions, the two keywords of *follow* and *business* were found to be commonly present in the public descriptions compared to the protected ones. In addition, the public descriptions seem to *mention* other accounts more often than their protected counterparts. On the other hand, the word *smile* is frequently used in the protected descriptions, which can either be the use of the word directly or a smiley face replaced with the word *smile* in the cleaning phase.

Since our classification is conducted on the users who have a description, the feature that checks the presence of the description is of no value here and so is removed. In addition, Twitter users with protected accounts need to approve their followers, while any user can instantly follow the public accounts. Therefore, the differences in the follower counts may not necessarily stem from differences in privacy attitude; hence, the follower count is also removed from our feature set. In summary, our classifiers are built on a total of 26 profile features extracted from each account, consisting of 11 surface-based attributes, 11 attributes extracted by LIWC, as well as four keyword-based features.

Table 6 provides our evaluation results of multiple classifiers. Our best classification results are obtained using *Classi cationViaRegression* with a performance of 0.71. Since there is no comparable study in the literature, we used a random classifier as our baseline. Based on the exact numbers of protected and public accounts in our underlying set, 48% of the set is composed of protected accounts, while 52% are users with public accounts. Therefore, a random classifier will label protected instances with an F-score of 0.48. Our feature set outperforms this baseline across all algorithms, and improves the results by 23% in the best case.

In addition, we used a naive baseline to compare the results. This baseline decides to label users based on their geo-enabled feature. This rule is established based on how the Twitter *setting* configuration page interface is arranged. In this page, the section designed to change the geo-enabled attribute is placed right at the top of the one designed to modify the privacy setting. Therefore, this baseline naively assumes that the users who have changed their default geo-enabled field from *false* to *true* are aware of the default privacy setting and thus have changed their privacy setting to the *protected* mode as well. The naive baseline reaches an F-score of 50.77, which is roughly 21% worse that our best algortihm. These results suggest that even by only relying on the profile attributes, one may be able to automatically detect users' privacy preferences in social media. This finding is encouraging. We call for the analysis of other attributes of Twitter users and their potential relationships to their privacy behaviour.

Table 6: Evaluation of classification results

Algorithm	Precision	Recall	F-score
Naive Bayes	0.66	0.67	0.66
Regression	0.71	0.70	0.71
Logistic	0.69	0.70	0.69
J48	0.68	0.66	0.67
KNN	0.67	0.59	0.63

6. DISCUSSION

Characterizing user privacy preferences in social media is a difficult and challenging task that requires a careful examination of various aspects of the users' social footprints. One class of such footprints can be found in how users shape and build their account profiles. In this study, we identified possible connections between users' profile attributes in Twitter and their privacy settings.

In particular, we found that protected accounts enable the geo-tracking feature more often compared to the public ones. As well, they tend to provide their location information and change their default profile theme. Besides, they tweet much more frequently. These differences, along with the common presence of emotion bearing and affect words in protected profile descriptions, can be associated with extraverted personality. Based on this interpretation, our finding is in contrast to earlier research on privacy in traditional settings, stating that introverts tend to be more privacy-concerned and are more likely to feel invaded when asked to reveal private information [19].

A possible speculation is that since users with protected accounts are aware that their tweets and accounts are private, they feel secure in this environment and are willing to voluntarily reveal more information about themselves and participate more actively in the network. On the other hand, users who are consciously following the public setting are utilizing a different strategy to protect their privacy, which is not including their location information, using a default theme, tweeting less, or using fewer function and affect words in their descriptions. If this is the case, then the users who are engaging in social media more actively (e.g., making changes to their profile attributes, sharing personal information, and revealing emotional states), tend to feel secure in the environment. Therefore, they are more likely to feel invaded by targeted advertising and marketing messages.

As discussed earlier, users' privacy behaviours may not necessarily match their privacy attitudes. Despite our efforts to choose a set of accounts with a minimal number of such false positives, our underlying set can still include public accounts that were meant to be protected. In spite of this issue, we obtained an F-score of 0.71, indicating the value and importance of profile attributes in the detection of privacy behaviour. Based on this finding, unsupervised or semi-

supervised techniques can be developed to effectively identify the public accounts that belong to privacy-concerned people, taking into account user profile attributes.

7. CONCLUSIONS

Ongoing creation of large social footprints offers immense potentials for both business and the individual. However, as users' privacy concerns are often not well-translated into their privacy settings, their data may be unintentionally visible to the public and thus should not be used for user profiling purposes. Therefore, it is desirable to characterize users' privacy attitudes, allowing companies to make informed decisions whether to discard or use the publicly available data for business intelligence purposes.

In this study, we explored the benefits of using Twitter profile attributes to infer privacy settings. By building a feature set based on these attributes, we obtained an F-score of 0.71 for the detection of privacy-concerned accounts. The classifiers in our experiments consistently outperformed both a random and a naive baseline and proved to be of value for our task. To the best of our knowledge, this is the first study that attempts to find differences in how people with different privacy settings manage their profile attributes.

Our analysis of the profile features and our classifiers are based on CNN followers' accounts. To generalize these findings, we will conduct a similar analysis and classification process with other Twitter accounts. We also plan to explore other available data sources. For example, we may be able to offer more informed labeling of the users' privacy levels based on their network structures. Furthermore, the content of user tweets is expected to be of great potential toward user classification since natural language has been shown to be a reflection of internal states. We will also investigate the generalizability of our approach by analyzing similar feature sets across different social platforms.

8. ACKNOWLEDGMENTS

This project is supported by the Mitacs Accelerate program in Canada. We also thank our industry partner InfoTrellis for its support.

9. REFERENCES

[1] M. Anwar, P. W. Fong, X.-D. Yang, and H. Hamilton. Visualizing privacy implications of access control policies in social network systems. In *Data Privacy Management and Autonomous Spontaneous Security*. Springer Berlin Heidelberg, 2010.

[2] J. Cohen. *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Earlbaum Associates, 1988.

[3] C. Dong, H. Jin, and B. Knijnenburg. Predicting privacy behavior on online social networks. In *Proceedings of the AAAI Conference on Web and Social Media*, 2015.

[4] L. Fang and K. LeFevre. Privacy wizards for social networking sites. In *Proceedings of the International Conference on World Wide Web*, pages 351–360, 2010.

[5] B. Gao and B. Berendt. Circles, posts and privacy in egocentric social networks: An exploratory visualization approach. In *Proceedings of the Conference on Advances in Social Networks Analysis and Mining*, pages 792–796, 2013.

[6] R. Gross and A. Acquisti. Information revelation and privacy in online social networks. In *Proceedings of the ACM Workshop on Privacy in the Electronic Society*, pages 71–80, 2005.

[7] D. Irani, S. Webb, K. Li, and C. Pu. Large online social footprints–An emerging threat. In *Proceedings of the Conference on Computational Science and Engineering*, volume 3, pages 271–276, 2009.

[8] T. Khazaei, L. Xiao, R. Mercer, and A. Khan. Detecting privacy preferences from online social footprint: A literature Review. In *Proceedings of the iConference*, 2016.

[9] A. Kobsa. Privacy-enhanced personalization. *Communications of ACM*, 50(8):24–33, 2007.

[10] Y. Liu, C. Kliman-Silver, and A. Mislove. The tweets they are a-changin': Evolution of Twitter users and behavior. In *Proceedings of the AAAI Conference on Weblogs and Social Media*, 2014.

[11] A. Mazzia, K. LeFevre, and E. Adar. The PViz comprehension tool for social network privacy settings. In *Proceedings of the Symposium on Usable Privacy and Security*, pages 13:1–13:12, 2012.

[12] C. R. McKenzie, M. J. Liersch, and S. R. Finkelstein. Recommendations implicit in policy defaults. *Psychological Science*, 17(5):414–420, 2006.

[13] T. Minkus and N. Memon. Leveraging Personalization To Facilitate Privacy. *ArXiv e-prints*, 2014.

[14] K. Naini Djafari, I. Altingovde, R. Kawase, E. Herder, and C. Niederee. Analyzing and predicting privacy settings in the social Web. In *User Modeling, Adaptation and Personalization*. Springer International Publishing, 2015.

[15] D. Nguyen, R. Gravel, D. Trieschnigg, and T. Meder. "how old do you think i am?" a study of language and age in twitter, 2013.

[16] M. Pennacchiotti and A. Popescu. A machine-learning approach to twitter user classification. In *Proceedings of the AAAI Conference on Weblogs and Social Media*, 2011.

[17] D. Rao, D. Yarowsky, A. Shreevats, and M. Gupta. Classifying latent user attributes in twitter. In *Proceedings of the Workshop on Search and Mining User-generated Contents*, pages 37–44, 2010.

[18] A. Squicciarini, S. Karumanchi, D. Lin, and N. DeSisto. Identifying hidden social circles for advanced privacy configuration. *Computers & Security*, 41:40 – 51, 2014.

[19] D. L. Stone. Relationship between introversion/extraversion, values regarding control over information, and perceptions of invasion of privacy. *Perceptual and Motor Skills*, 62:371–376, 1986.

[20] K. Strater and H. R. Lipford. Strategies and struggles with privacy in an online social networking community. In *Proceedings of the British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction*, pages 111–119, 2008.

[21] E. Toch, Y. Wang, and L. Cranor. Personalization and privacy: A survey of privacy risks and remedies in personalization-based systems. *User Modeling and User-Adapted Interaction*, 22(1-2):203–220, 2012.