Facebook Users Have Become Much More Private: A Large-Scale Study

Ratan Dey
Polytechnic Institute of
New York University
Brooklyn, US
Email: ratan@cis.poly.edu

Zubin Jelveh
Polytechnic Institute of
New York University
Brooklyn, US
Email: zjelveh@gmail.com

Keith Ross
Polytechnic Institute of
New York University
Brooklyn, US
Email: ross@poly.edu

Abstract—We investigate whether Facebook users have become more private in recent years. Specifically, we examine if there have been any important trends in the information Facebook users reveal about themselves on their public profile pages since early 2010. To this end, we have crawled the public profile pages of 1.4 million New York City (NYC) Facebook users in March 2010 and again in June 2011.

We have found that NYC users in our sample have become dramatically more private during this period. For example, in March 2010 only 17.2% of users in our sample hid their friend lists, whereas in June 2011, just 15 months later, 52.6% of the users hid their friend lists. We explore privacy trends for several personal attributes including friend list, networks, relationship, high school name and graduation year, gender, and hometown. We find that privacy trends have become more pronounced for certain demographics. Finally, we attempt to determine the primary causes behind the dramatic decrease in the amount of information Facebook users reveal about themselves to the general public.

I. INTRODUCTION

Facebook was launched in February 2004. During its first few years, Facebook was largely restricted to university and high-school students, and not until October 2006 was it opened to all Internet users. Facebook has been growing dramatically since these early stages. In August 2008 Facebook reached 100 million users, and today it has over 800 million users. According to [1], the 35+ demographic is growing rapidly and represents around 30% of all Facebook users. The ethnic diversity of the U.S. Facebook users is now similar to the diversity of the overall population of the U.S. [2].

Facebook allows users to control and customize the personal information they make available to other users. For example, a Facebook user – let's call her Alice – can configure her account so that her friends can see her photos and interests, but the general public can only see her name and profile picture on her public profile page. In particular, Alice has the option of hiding her attributes such as friend list, age, gender, relationship status, and political affiliation in her public-profile page. Facebook users thus have a wide-range of options in deciding how public they will be to the Internet community at large.

In this paper, we investigate whether Facebook users have become more private in recent years. Specifically, we examine if there have been any major trends in the information Facebook users reveal about themselves on their public profile pages since early 2010. To this end, we have crawled the public profile pages of 1.4 million NYC Facebook users in March 2010 and again in June 2011. Although NYC users do not constitute a random sample of Facebook users, given the diverse demographics of the NYC population, trends among Facebook users in NYC are quite possibly indicative of general trends in the USA and perhaps in the world.

We have found that a large fraction of users in our sample have become dramatically more private during this period. For example, in March 2010 only 17.2% of users in our sample hid their friend list, whereas in June 2011, just 15 months later, 52.6% of the users hid their friend lists. In addition to friend list, we explore privacy trends for several personal attributes including networks, relationship, high school name and graduation year, gender, and hometown. We also explore whether privacy trends have become more pronounced for certain demographics. Finally, we attempt to determine the primary causes behind the dramatic decrease in the amount of information Facebook users reveal about themselves to the general public. As Facebook users become more private, not only will the interactions among members change significantly, but it will also become more difficult for third parties to collect and infer personal information about users, as described in the body of the paper.

The remainder of this paper is organized as follows. We present our data gathering mechanism and dataset properties in Section II. In Section III we study the friend-list attribute in detail, investigating trends in different demographics, including gender, age, and neighbourhood. We also investigate the primary causes behind the shifts in privacy for the friend list attribute. In Section IV we explore trends for several other attributes, including networks, relationship, high school name and graduation year, gender, and hometown. In Section V we review related work and in Section VI we summarize our conclusions.

II. DATA SETS

In Facebook, when Alice visits Bob's profile page, the information that is displayed to her depends on her relationship with Bob (for example, whether she is a friend or not) and on Bob's privacy settings. Roughly speaking, when Alice is a

Facebook friend of Bob, then she typically gets to see Bob's **full profile page**, which includes the list of Bob's friends as well as all of the information and photos that Bob puts into Facebook; if Bob is not a friend, Alice only gets to see a **public profile page**, which often includes no more than Bob's full name and his photo.

We were given IRB approval from our university to crawl Facebook to investigate privacy leakages and trends of privacy consciousness. We developed a multi-threaded crawler that visits Facebook user profile pages and stores the pages in a MySQL database. We performed three crawls. Before September 2009 Facebook allowed all members of the same regional network to see each others' full profile pages by default. In July 2009 we joined the NYC network and crawled the full profile pages for 1.67 million NYC users, which was the large majority of NYC Facebook users at that time. Facebook fully deprecated regional networks as of late September 2009 [3], [4]. A user's full profile is now, by default, only available to the user's friends.

In March 2010, we crawled these same 1.67 million users and collected the *public profile pages* for 1.47 million user IDs. We refer to this dataset as the ***March 2010 dataset***. We then launched another crawl in June 2011, revisiting the public profile pages of users in our March 2010 data set. For this third crawl, we were able to collect 1.41 million of the public profiles. We will call this data set the ***June 2011 data set***. We believe that the missing 200K users between the first and second crawl, and the missing 50K users between the second and third crawls either deactivated their accounts or were removed by Facebook in the intervening period. Our privacy analysis is based on the 1.41M users found in all three crawls.

In Table I, we show the *default* privacy settings that were in place at the times of our two crawls for a number of different attributes [5], [6], [7]. "All Facebook users" means visible to any one who has a Facebook account. From the table we see that the default settings were largely the same during the two crawls. There were only minor changes in Likes, Wall Posts, Hometown, and Current City. For these attributes, Facebook actually made the default settings less private, changing the default settings from "Friends of Friends" to "All Facebook users". As described in the subsequent section, we have found that many users have explicitly changed their default privacy settings for their public profile page.

Online Social Networks (OSN) recently began to introduce new privacy controls. These new controls are not only more user friendly, but also allow users to more precisely control the information they share with different sets of users. For example, as of August 23, 2011 Facebook provides in-line privacy control tools in profile pages [8]. The data and analyses presented in this paper, although very recent, reflects users' privacy behavior before this change.

Table II provides the big-picture view of information privacy trends for nine attributes: friend lists, age, high-school name and graduation year, network, relationship, gender, interested in, hometown, and current city. For each attribute, Table II reports the percentage of users for which the attribute is public to the entire world (that is, in their public profile pages) for each of the two periods. We see that for most of these categories, NYC Facebook users have indeed become more private (either by explicitly setting their privacy settings or by optionally not providing the information). In March 2010, for 12.3% of the users, all of these attributes were private; whereas in June 2011, for 33% of the users, all of these attributes were private. This is a large shift, especially if we consider that Facebok changed its default settings to disclose more information during this period. In the remainder of this section, we analyse many of these attributes in greater detail.

III. FRIEND LIST

In this section we investigate whether users are becoming more private with their friend lists and, if so, whether the trend is more pronounced for certain demographics. We also attempt to determine the primary potential causes behind the privacy trends.

Figure 1 shows the privacy trend for friend list. In this figure, the first column shows the percentage of users who are hiding (depicted as H) and are not-hiding (depicted as NH) in March 2010. The second column shows these percentages for June 2011. From Figure 1, we see a dramatic change has taken place, with users becoming much more private about who their friends are. In particular, a relatively small fraction went from hiding to not-hiding (3.6%), but a large fraction went from not-hiding to hiding (43.5%). As of June 2011, less than half the users made their friend lists public.

There are many recent studies that exploit the friend list to uncover hidden information about users in OSNs. The paper [9] uses friend list to predict the sexual orientation of Facebook users. The paper [10] predicts the value of a user's attribute by using the most popular value among the user's friends. Similarly see [11], [12], [13], [14], [15], [16]. If a user X hides his friend list, but many of the X's friends do not hide their friend lists, then many of user X's friends can still be determined by intelligent crawling [17], [18]. But as more users choose to hide their friend lists in their public profile pages, it becomes increasingly more difficult (for good or for bad!) for third-parties to crawl Facebook, build a social graph, and infer hidden information about users.

A. Demographic Analysis

To take a closer look at the friend-list attribute, we classify all the users based on gender. For users who didn't publish their gender publicly, we accurately determine their gender using the techniques in [16]. Figures 2 and 3 show that women were more private than men in March 2010 and continue to be so in June 2011. We document a greater concern for privacy among female users as compared with an earlier Facebook study using a small data set [19].

We also examined the privacy behavior among Facebook users in the five boroughs of New York City (Manhattan, Queens, Brooklyn, Bronx, and Statan Island). Facebook users

TABLE I
FACEBOOK'S DEFAULT PRIVACY SETTINGS DURING MARCH 2010 AND JUNE 2011

Attribute Name	Visible Audience (March 2010)	Visible Audience (June 2011)
Name	All Facebook users	All Facebook users
Profile Picture	All Facebook users	All Facebook users
Gender	All Facebook users	All Facebook users
Networks	All Facebook users	All Facebook users
Contact Info	Friends	Friends
Friend List	All Facebook users	All Facebook users
Relationship	All Facebook users	All Facebook users
Interested In	All Facebook users	All Facebook users
High School Information	All Facebook users	All Facebook users
Birthday	Friends of Friends	Friends of Friends
Likes	Friends of Friends	All Facebook users
Wall Posts	Friends of Friends	All Facebook users
Hometown	Friends of Friends	All Facebook users
Current City	Friends of Friends	All Facebook users

TABLE II
BIG-PICTURE VIEW OF PRIVACY TRENDS

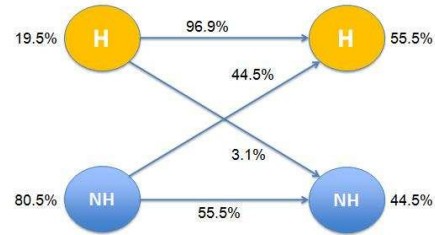| Attribute	March 2010	June 2011
% users with friend list public	82.7	47.4
% users with networks public	25.1	21.4
% users with relationship info public	11.3	4.9
% users with HS name and graduation year public	13.4	9.1
% users with gender public	58.9	52.8
% users with age public	1.5	1.4
% users with "interested in" public	7.7	6.4
% users with hometown public	10.4	24.0
% users with current city public	31.3	36.5
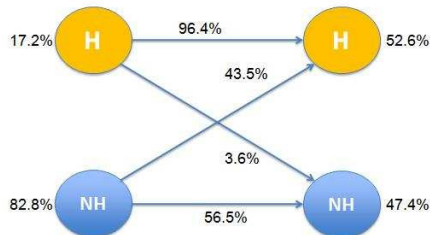
Fig. 1. Trend in hiding friend list from 3/2010 to 6/2011

Fig. 3. Trend in hiding friend list for females from 3/2010 to 6/2011
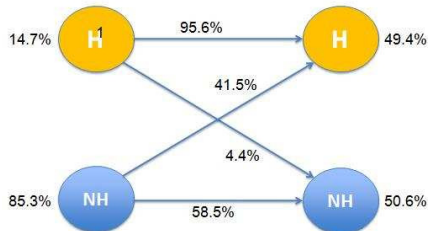

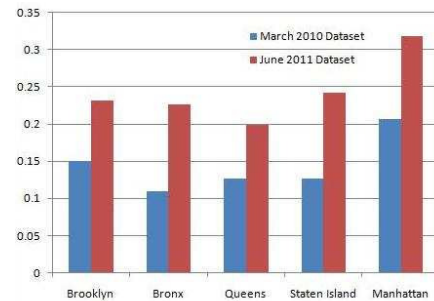Fig. 2. Trend in hiding friend list for males from 3/2010 to 6/2011

Fig. 4. Hiding friend list in different boroughs in NYC (Based on Current City)

can list their hometown or current city at a neighborhood level of granularity (such as Harlem, New York), at the borough level of granularity (such as Queens, New York), or at the city level granularity (New York, New York). We classified a user into a particular borough if he indicated he lives in that borough or in a neighborhood of that borough. Thus, for example, if a user indicated he lives in Harlem, we assigned him to the Manhattan borough. Users who indicated New York, New York as their current city were not classified into a borough. Figure 4 shows the percentage of users who hide their friend lists in each borough who are currently living in NYC.
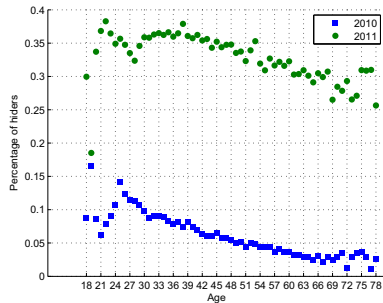
Fig. 5. Percentage of users who hide their friend list as a function of age.

Table III shows the average income and the percentage of foreign-born residents in each of the five boroughs [20]. We see that Facebook users from Manhattan – where the average income is the highest and percentage of foreign-born residents is relatively low – are the most privacy conscious. We see that Facebook users currently residing in the Queens – where the average income is mid-level and the percentage of foreign-born residents is the highest – are the least privacy conscious. Although further research is required here, we conjecture that, on average, the more money you make, the more private you are! We also conjecture that foreign-born residents, on average, are less privacy conscious than US-born residents. We also observe that there is a marked trend to become more private in all five boroughs.

We also investigate online privacy trends for different age groups. Figure 5 shows the percentage of users who hide their friend list as a function of age. For this experiment, we found 452K users who provided birth year in their full profile pages in the July 2009 dataset and increased the sample size with the users who provide birth years in their public profile pages either in March 2010 or June 2011 dataset. To increase the sample size further, we also estimate a user's birth year from his high-school graduation year, when high-school graduation year is publicly available but birth year is not available in all of the three datasets, as discussed in [17]. After including all these users in the sample, approximately 41% of the users in our data set are assigned a birth year. We see all age groups became more private (with respect to friend list) from March 2010 to June 2011. Focusing on 2011, we see that the younger and middle aged users are more likely to hide their friend list than older users (greater than 55 years). Among the younger users, the age group – $24-28$ years of age, is the least likely to hide. As of now, we do not have a good explanation for age-dependent behavior.

B. Analyzing factors behind hiding friend list

In this section we take a closer look at the potential reasons for the dramatic shift in the number of users who hid their friend list between our two crawls. Recall that during both the March 2010 and the June 2011 crawls, a user's friend list was by default public to all of the Facebook users. (It is also public by default at the time of this writing, November 2012.) Thus for this attribute, if a user went from NH to H, it was

clearly the user's conscious decision, and not due to changes in Facebook's default settings.

Figure 6 shows what we believe to be the two largest factors behind this change: a growing awareness of the risks associated with sharing personal information online, as well as Facebook-specific privacy issues. Using the Factiva news database, we retrieved the monthly number of news articles that included the terms 'facebook' and 'privacy' between January 2009 and September 2011. As can be seen, there is a persistent increase in the number of articles containing these words. By September 2011 this metric was 4.5 times greater than in January 2009 and 1.5 times greater than in March 2010. However, this last data-point masks the avalanche of attention which Facebook received in May 2010 as criticism grew over the company's December 2010 decision to make users' profiles public by default. In late May, Facebook responded [21] to the criticism and redesigned its privacy settings interface to reduce the number of clicks necessary to change many default settings, including whether to hide one's friends list. While it's likely that this greater awareness of privacy issues by the public and Facebook's privacy page redesign were major drivers in the dramatic increase we document in the number of users who hide their friends list, we are also interested in whether there was a social component to this shift.

In order to concentrate on active users, we restrict the sample to accounts with more than 10 friends in New York City and more than 10 friends total. Since we are interested in the decision to hide, we also restrict the sample to those users who didn't hide their friends list in the 2010 dataset. This leaves us a little less than $800,000$ users. We examine the change in the percentage of Alice's friends who have hidden their friends list as a predictor of whether Alice decides to do the same. We ran a logistic regression of the following form:

$$isHidden_{2011} = \beta_0 + \beta_1 \triangle \%Hidden_{10,11} + \beta\vec{x}$$

where $\triangle\%Hidden_{10,11}$ is the change in the percentage of Alice's friends who have hidden their friends list between 2010 and 2011 and $\vec{x}$ is a vector of covariates which capture changes in privacy settings along other dimensions. Our main parameter of interest, β_1, should be positive if an increase in the percentage of friends who are hiding their friends list is associated with Alice's decision to hide her friends list by the time of our second crawl in 2011. The covariates are observable attributes which we believe capture Alice's preference for privacy. These variables include dummies indicating whether, for a given year, a certain attribute such as gender, educational information, relationship status, etc. is listed on Alice's public Facebook profile. For example, one covariate in $\vec{x}$ is the change in whether Alice's gender was made public between 2010 and 2011. We subtract the 2010 value from the 2011 value so that a positive difference is associated with less concern over privacy. In general, we expect the coefficients on these variables to be negative.

Table IV shows the results from this model. The first column reports the baseline model with only our variable of interest,

TABLE III
PRIVACY TRENDS OF DIFFERENT BOROUGHS IN NEW YORK CITY

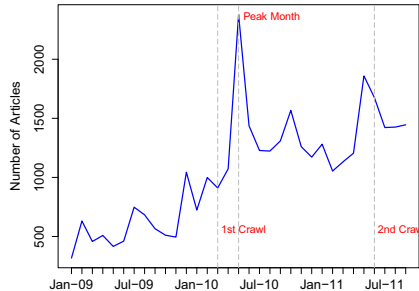| Borough Name	Average Income (Dollars)	Foreign-born Population (%)
Brooklyn	61, 917	37.3
Bronx	47, 276	32.0
Queens	69, 304	48.1
Staten Island	86, 162	21.0
Manhattan	126, 035	28.5

Fig. 6.   Number of news articles per month containing the terms 'facebook' and 'privacy'. Source: Factiva, All Sources data set

$\triangle\%Hidden_{10,11}$, as an independent variable. We find that a one standard deviation increase from the mean in this variable is associated with a 6.2% increase in the probability that a user hides her friends list. The second column reports the model with our set of covariates. Once these are included, we find that the relationship declines and a one standard deviation increase from the mean in the change in the percentage of a user's friends who are hiding being associated with a 2.7% increase in the probability that a user hides her friends list. If there is a network effect whereby users decide to change privacy settings based on the decisions of their friends, the evidence presented here suggests that this effect is likely to be small.

IV. OTHER ATTRIBUTES

In this section we examine the privacy trends for some other personal attributes. In Facebook, users can join networks, such as the network consisting of all the employees of a company or the network consisting of all the students in a university. Figure 7 examines the network attribute, that is, whether or not users make the networks to which they belong public. During both time periods, the default settings were to make networks public (see Table I). For this attribute, we interpret a user as "hiding" if he either explicitly changed the default settings or if he doesn't belong to a network. Observe that 20.5% of non-hiding users in 2010 converted to hiding in 2011, whereas only 0.6% of hiding users converting to not-hiding. This trend also has important consequences for inferring characteristics about users. For example, if a user Bob does not make his university network public, but has many friends who publicly indicate attending the Foobar Univesity network, then we can infer that user Bob most likely also attends Foobar University. Because there is a pronounced trend to hide network information, inference of schools and workplaces will become more difficult.
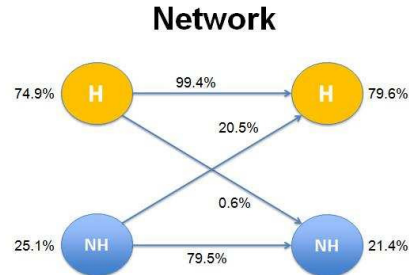

Fig. 7.   Trend in hiding network information from 3/2010 to 6/2011
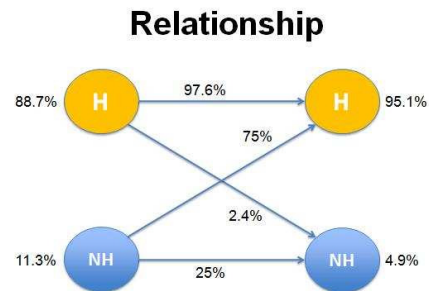

Fig. 8.   Trend in hiding relationship information from 3/2010 to 6/2011

Figure 8 shows the privacy trend for the relationship status attribute. From the figure we can see that 75% users who publicly disclosed their relationship status in March 2010 chose to hide this information in June 2011. As of 2011, only 5% of users make their relationship status public to the general public. Figure 9 shows the trend in hiding high-school graduation information (high school and year of graduation). From the figure, we can observe that a significant percentage (62.1%) of users who disclosed their high-school information in March 2010 chose to hide this information in June 2011. Because a person's graduation year can be used to predict a user's age with high accuracy [17], this trend will make it more difficult to accurately estimate Facebook users' ages in the future.

Figure 10 shows the trend for the gender attribute. The results for this attribute are less intuitive than for the previous attributes analyzed in this paper. Although on the whole NYC Facebook users have become more private about their gender (going from 41.1% in March 2010 to to 48.2% in June 2011), surprisingly a significant fraction (24.3%) of users went from hiding to not hiding. Perhaps many of these users have concluded their gender can easily be inferred from their first name [16] or from their photo, and therefore hiding gender information does not truly provide additional privacy.

For each of the attributes discussed above, users are more private in June 2011 than they were in March 2010. But for

TABLE IV
RESULTS FROM LOGISTIC REGRESSION MODEL

	Model 1		Model 2	
	Estimate	S.E.	Estimate	S.E.
(Intercept)	-1.1081**	(0.0075)	-1.4634**	(0.0204)
$\triangle\%Hidden_{10,11}$	2.3519**	(0.0227)	1.3132**	(0.0303)
$\triangle Sex\_Listed_{10,11}$	.		-2.5617**	(0.0097)
$\triangle Relationship\_Listed_{10,11}$	.		-0.4237**	(0.0113)
$\triangle Interests\_Listed_{10,11}$	.		0.0348*	(0.014)
$\triangle Hometown\_Listed_{10,11}$	.		-0.9299**	(0.0079)
$\triangle CurrentCity\_Listed_{10,11}$	.		-1.1945**	(0.0058)
$\triangle Network\_Listed_{10,11}$	.		-1.7035**	(0.0194)
$\triangle BirthYear\_Listed_{10,11}$	.		0.904**	(0.0109)
$\triangle Education\_Listed_{10,11}$	.		-0.9772**	(0.0099)
$log(Total\_Friends\_NYC)$	.		0.0506**	(0.0038)
N	790,081		790,081	

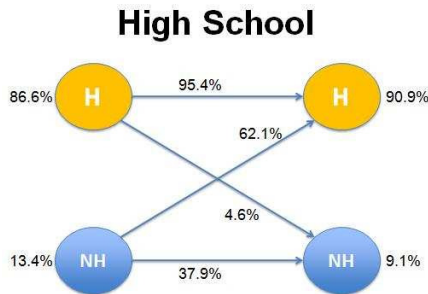* $p \leq 0.05$, ** $p \leq 0.01$

High School

Fig. 9.   Trend in hiding high school graduation information from 3/2010 to 6/2011
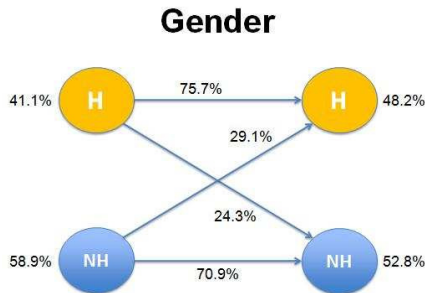
Gender

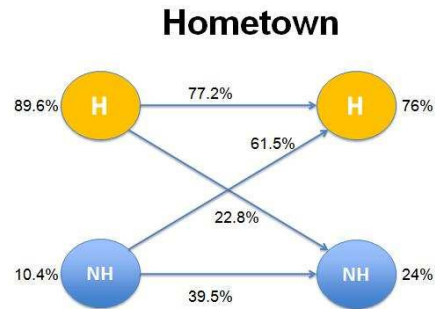Fig. 10.   Trend in hiding gender information from 3/2010 to 6/2011

Hometown

Fig. 11.   Trend in hiding hometown information from 3/2010 to 6/2011
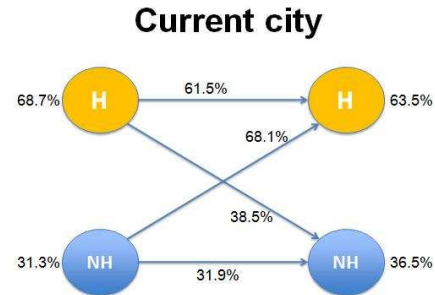
Current city

Fig. 12.   Trend in hiding current city information from 3/2010 to 6/2011

some attributes people are actually collectively less private than they were. Figures 11 and 12 shows the trend for the hometown and current city attributes. For both categories, there is a significant fraction of users "crossing over" from hiding to not-hiding and from not-hiding to hiding. Here we can attribute much of the cross over from hiding to not hiding to Facebook's change in default privacy settings (see Table I). Indeed many users who were hiding by default in March 2010 became non-hiding by default in March 2011.

V. RELATED WORK

In 2005 Gross and Acquisti [22] studied patterns of information revelation in OSNs and their privacy implications. Their study evaluated Facebook privacy disclosure and privacy settings for 4,000 Carnegie Mellon University (CMU) students. In 2006 Acquisti and Gross [23] further surveyed 294 Facebook users at CMU and compared the survey results to the information retrieved from Facebook. They also examined underlying demographic or behavioral differences between members and non-members and analyzed the impact of privacy concerns on member behavior.

In 2007 Strater and Lipford [24] surveyed 18 Facebook users, all students from University North Carolina Charlotte. They performed a formative study of the privacy concerns and strategies of Facebook users. In 2009 Stutzman et. al. [25] surveyed 122 Facebook student users from University of North Carolina. They explored how privacy settings and privacy policy consumption affect the relationship between privacy settings and disclosure behaviors. In 2007 Lewis et. al. [19] analyzed the factors that are predictive of a student having a private versus public profile. Their experiment is based on $1,740$ students at Harvard University. They proposed four hypotheses and concluded that privacy behavior is an upshot of both social influences and personal incentives.

Recently Liu et. al. [26] measured the disparity between desired and actual privacy settings, quantifying the magnitude

of the privacy-management problem. Their analysis is based on survey of 200 Facebook users who were recruited via Amazon Mechanical Turk. They explore neither privacy trends nor information disclosure for different demographics.

With the exception of Liu et. al., all of the above surveys took place in 2005-2007 with a relatively small number of users (under 2,000) in a narrow demographic (US undergraduates). Since 2007, Facebook has experienced dramatic changes in both the size and diversity of its user base. Our study took place in 2010 and 2011, and involved a large and highly diverse user base of 1.41 million users from New York City. Because of the diversity of the user base, we were able to provide new insights into user privacy concerns for many demographics based on age, gender, and neighborhoods. We also identified a dramatic trend from 2010 to 2011, namely, many NYC Facebook users are disclosing less information in public profiles. Finally, we provide evidence that media attention and Facebook's privacy page redesign are the most important potential factors behind the trend.

VI. CONCLUSION

We have found that NYC users have become dramatically more private. In March 2010 only 17.2% of the users in our sample hid their friend list, whereas in June 2011, just 15 months later, 52.6% of them hid their friend lists. In March 2010, for 12.3% of the users, all of the attributes in Table II were private, whereas in June 2011, for 33% of the users, all of these attributes were private.

We have found that women tend to be more private than men, and that young and middle aged people tend to be more private than older users. We have found that people living in the wealthier boroughs and in boroughs with more US-born users tend to be more privacy conscious. We have found that people's decisions to be private are not significantly influenced by their friends' decisions. We provide evidence that media attention and Facebook's privacy page redesign are the most important potential factors behind the trend.

REFERENCES

[1] "Facebook statistics, stats facts for 2011," January 9, 2011, available at: http://www.digitalbuzzblog.com/facebook-statistics-stats-facts-2011/.
[2] "How diverse is facebook?" December 16, 2009, available at: http://www.facebook.com/note.php?note_id=205925658858.
[3] "Facebook to fully deprecate regional networks by september 30," available at: http://www.insidefacebook.com/2009/08/05/facebook-to-fully-deprecate-regional-networks-by-september-30.
[4] "Developer blog: July 2009 platform news," available at: http://developers.facebook.com/blog/post/285.
[5] "The evolution of privacy on facebook," May 2010, available at: http://www.mattmckeon.com/facebook-privacy/.
[6] "What's it like to create a new facebook account?" May 10, 2010, available at: http://rhosgobel.blogspot.com/2010/05/whats-it-like-to-create-new-facebook.html.
[7] "New tools to control your experience," December 9, 2009, available at: http://blog.facebook.com/blog.php?post=196629387130.
[8] "Making it easier to share with who you want," August 23, 2011, available at: http://www.facebook.com/blog.php?post=10150251867797131.
[9] B. F. M. Carter Jernigan, "Gaydar: Facebook friendships expose sexual orientation," First Monday, vol. 14, no. 10, 2009.
[10] J. Becker and H. Chen, "Measuring privacy risk in online social networks," in W2SP, 2009.
[11] A. Mislove, B. Viswanath, K. P. Gummadi, and P. Druschel, "You are who you know: Inferring user profiles in online social networks," in WSDM, 2010.
[12] E. Zheleva and L. Getoor, "To join or not to join: the illusion of privacy in social networks with mixed public and private user profiles," in WWW, 2009.
[13] R. Heatherly, M. Kantarcioglu, B. Thuraisingham, and J. Lindamood, "Preventing Private Information Inference Attacks on Social Networks," University of Texas at Dallas, Tech. Rep. UTDCS-03-09, 2009.
[14] W. Xu, X. Zhou, and L. Li, "Inferring Privacy Information via Social Relations," in 24th International Conference on Data Engineering Workshop, 2008, pp. 154–165.
[15] J. He, W. W. Chu, and Z. Liu, "Inferring privacy information from social networks," in ISI, 2006, pp. 154–165.
[16] C. Tang, K. W. Ross, N. Saxena, and R. Chen, "What's in a name: A study of names, gender inference, and gender behavior in facebook," in DASFAA Workshops, 2011, pp. 344–356.
[17] R. Dey, C. Tang, K. W. Ross, and N. Saxena, "Estimating age privacy leakage in online social networks," in IEEE INFOCOM: Mini-Conference, March 2012.
[18] K. Thomas, C. Grier, and D. M. Nicol, "unfriendly: Multi-party privacy risks in social networks," in Privacy Enhancing Technologies, 2010, pp. 236–252.
[19] K. Lewis, J. Kaufman, and N. Christakis, "The taste for privacy: An analysis of college student privacy settings in an online social network," J. Computer-Mediated Communication, vol. 14, no. 1, pp. 79–100, 2008.
[20] "Demographic, social, economic and housing profiles of each borough in nyc," October 10, 2011, available at: http://www.nyc.gov/html/dcp/download/biggapps/DCP_acs_nyc_boros_001.xls.
[21] "Understanding your privacy controls on facebook," May 27, 2010, available at: http://blog.facebook.com/blog.php?post=394231632130.
[22] R. Gross, A. Acquisti, and H. J. H. III, "Information revelation and privacy in online social networks," in WPES, 2005, pp. 71–80.
[23] A. Acquisti and R. Gross, "Imagined communities: Awareness, information sharing, and privacy on the facebook," in Privacy Enhancing Technologies, 2006, pp. 36–58.
[24] K. Strater and H. R. Lipford, "Strategies and struggles with privacy in an online social networking community," in BCS HCI (1), 2008, pp. 111–119.
[25] F. Stutzman, R. Capra, and J. Thompson, "Factors mediating disclosure in social network sites," Computers in Human Behavior, vol. 27, no. 1, pp. 590–598, 2011.
[26] Y. Liu, K. P. Gummadi, B. Krishnamurthy, and A. Mislove, "Analyzing facebook privacy settings: User expectations vs. reality," in IMC, 2011.