Privacy Albeit Late

Gustavo Rauber
UFMG
Brazil
rauber@dcc.ufmg.br

Virgílio A. F. Almeida
UFMG
Brazil
virgilio@dcc.ufmg.br

Ponnurangam Kumaraguru
IIIT-Delhi
pk@iiitd.ac.in

ABSTRACT

Online Social Networks (OSNs) such as Facebook and Twitter have experienced exponential growth in recent years. Users are spending more time on OSNs than on any other sites and services on the Internet. Users post and share a lot of personal information on these sites without being aware of privacy implications or simply not caring much about them, what turns to be a treasure for marketing companies and cyber criminals. Characterizing the privacy awareness of users is important to design technologies and policy solutions. Users expect the OSN to provide good privacy protection or controls so they can make informed decisions about their privacy. This paper investigates the privacy awareness of users on Facebook using real-world data (not self reported). The main findings are: only a low percentage of users change the default privacy settings; a large percentage of users expose their gender publicly; women are more concerned about disclosing personal information online; many users share their photo albums and links (content) to everyone; users exercise more control over content that are more potentially dangerous to their reputation. The present study is one of the first to characterize the privacy awareness on OSN through a real world experiment. Implications of the study are discussed.

Categories and Subject Descriptors

H.1.2 [**Models and Applications**]: User / Machine systems—*human factors, human information processing*; K.4.1 [**Computers and Society**]: Public Policy Issues—*Privacy*; K.6.5 [**Management of Computing and Information Systems**]: Security and protection

General Terms

Experimentation, Measurement, Human factors

Keywords

Privacy, Online social networks, Real-world experiments

Webmedia '11 Florianópolis, SC, Brazil

1. INTRODUCTION

Popular destinations such as search engines, news media, social networking, video sharing, online games on the Web attract hundreds of millions of users every day, where they interact with different kinds of services. On one hand, these interactions yield valuable data that can be used to personalize the user's web experience. On the other hand, these interactions always leave data crumbs that can be used to breach user's privacy. Also, these destinations can share sensitive or personal user information with other users or third parties without proper user consent. Search engines can consciously or inadvertently also build user profiles, store user IP addresses, or collect any other information that could ever tie a particular search to a specific user [15].

Social networking sites offer attractive means of online social interactions and communications, but also raise privacy and security concerns. Facebook is the number one network in the world, except for a few countries, like Brazil, Japan and China.[1] Web surfers now spend more time socializing on Facebook than searching with Google.[2] Facebook has more than 500 million active users at any given point in time and 30 billion pieces of content (hyperlinks, notes, photos, etc.) are shared by its users each month. Facebook supports more than 70 languages, what makes it a huge global digital space. Like the Web itself, Facebook is a powerful technology to increase connection between people separated by borders of nation, language, religion and culture. With an estimated 65 billion friendships, it is important to study how this crucial technology is perceived across different cultures [9] and understand user's privacy awareness.[3] Facebook has also been revising its privacy policy and settings from the day of inception, what directly affects a large population in the world. The focus of this research is to study Facebook users' privacy awareness / carelessness around the globe, and in particular, in Brazil and India.

To the best knowledge of the authors, this paper is the first study to analyze and compare the privacy awareness of Facebook users in Brazil and India through a real world experiment. It has used real-world data (not self reported) for studying privacy preferences. Understanding users' behavior in real world settings is critical to develop any technological or policy solutions [19]. The findings from this paper can be useful for other OSNs and not just Facebook.

[1] http://www.brentcsutoras.com/2010/09/02/top-social-networks-top-internet-countries/
[2] www.comscore.com
[3] http://www.huffingtonpost.com/2010/04/27/facebook-changes-raise-pr_n_553129.html

The main contributions of this paper are:

- Investigate privacy awareness of Facebook users using real world data
- Show that the majority of the users are oblivious to privacy and reveal a lot of personal information on Facebook
- Show that users from two different countries have different perceptions about the desired levels of privacy.

The remainder of the paper is organized as follows: in the next section, other related privacy studies on Facebook are discussed and some literature on cultural studies on privacy is given. In Section 3, the study setup is presented, along with participant demographics and the data set that was collected for this study. In Section 4, the hypotheses evaluated using the data collected are drawn. In Section 5, the results of the analysis are provided, demonstrating that users share a lot of personal information on Facebook. In Section 6, the implications of the results are portrayed and finally, the limitations and future work of this research are discussed.

2. BACKGROUND

In this section is presented a brief background on various studies (in particular, privacy) that have been done on Facebook. It also describes some results from studies which have analyzed cultural aspects of privacy to provide a background on the comparison that is made between users from Brazil and India.

2.1 Privacy on Facebook

Due to its immense popularity (over 500 million active users at any given point in time), various research studies have been conducted on Facebook. Researchers have: analyzed the social network of Facebook users to find different patterns [20]; analyzed the impact of Facebook applications and games [24]; used it to study statistical sampling of participants on the Internet to generalize the result from the analysis [12]. One key topic that has been studied with respect to Facebook is privacy; also, in general, privacy has been an important topic of study on Online Social Networks [13, 26].

There have been many instances where users have consciously or inadvertently shared personal information on Facebook that have later become an embarassement for the users involved.[4] It has also been found that government and council employees in the U.K. are using social networking sites from office, where they are also exposing private or classified information.[5]

Privacy settings on Facebook have been on scrutiny for some time and various factors related to privacy settings on Facebook have been studied.[6] Privacy settings of Facebook has evolved over time.[7] Boyd et al. showed that both frequency and type of Facebook users as well as Internet skill are correlated with making proper modifications to privacy settings. They have also observed a few gender differences in how young adults approach their privacy configurations, which is notable, given that gender differences exist in so many other online domains [3].

Facebook has been used to study the Personally Identifiable Information leakages online. Krishnamurthy et al. analyzed Facebook for profile and friends to be viewed by others and found varying level of public exposure ranging from 76% to 99% of the users among different regional networks worldwide [15]. Gjoka et al. while analyzing Facebook for unbiased sampling showed that majority of users (84%) did not change their default privacy settings and only 7% of global users hid their friends from strangers [12] .

Many factors seem to influence the privacy awareness of users – geographical location, ethnicity, node degree and even the privacy awareness of friends. Gjoka et al. found that users around the globe were split between two extremes of privacy settings, which is inline with literature (Individualist and Collectivist society) [14]. Chang et al. using the Facebook data from the U.S. users showed that ethnicity of users impacted their privacy preferences. For example, Hispanic users share more photos than the average U.S. citizen user [4]. Gjoka et al. showed that users with low degree nodes tend to have stringent privacy settings while users with high degree nodes tend to be liberal in their privacy settings. This is counterintutive, as one would imagine that users with high degree nodes would be more aware of privacy settings and therefore would have changed it to being stringent. They also showed a positive correlation between one's privacy awareness and their friends' privacy awareness [12].

2.2 Privacy Studies in Brazil and India

Very little research work has been done in studying privacy perception or awareness in Brazil and India. Studying privacy awareness of users in these countries will help in decision making of technologies and policies for the use of the Internet. Countries like Brazil and India are expected to play a central role in the world of 21^{st} century.[8]

A large amount of research is conducted in the U.S. [17] and Europe on various aspects of privacy. Due to cultural background, there is a large difference in privacy perceptions among different parts of the world [2]. Hofstede has classified societies around the world into many categories and the two extremes are individualist and collectivist [14]. According to Hofstede both Brazil and India are collectivist societies. People in Brazil and India are unaware of various privacy issues both in the online and offline worlds [7, 16, 18].

2.3 Growth of Online Social Networking in Brazil and India

In Brazil, traffic to social networking sites grew 51% in 2009, reaching more than 36 million visitors aged 15 and older in August 2010. Facebook experienced triple-digit growth, increasing its audience 479% in 2009. However, according to Comscore, in Brazil, Orkut ranked as the most-

[4]http://www.dailymail.co.uk/news/article-1197562/MI6-chief-blows-cover-wifes-Facebook-account-reveals-family-holidays-showbiz-friends-links-David-Irving.html

[5]http://in.news.yahoo.com/139/20101010/882/twl-using-too-much-facebook-and-twitter.html

[6]http://www.guardian.co.uk/technology/2010/apr/26/facebook-privacy-hole

[7]http://www.economist.com/blogs/babbage/2010/10/facebook_and_transparency

[8]The combined BRIC (Brazil, Russia, India, and China) economies by 2050 is expected to be more than combined economies of richest countries in the world [25].

visited social networking destination, reaching 29.4 million visitors, followed by Windows Live Profile with 12.5 million visitors. Facebook secured the third spot with nearly 9 million visitors while Twitter had 8.6 million visitors, but with the highest Internet user penetration reach in the world, 23%. [6].

In India, more than 33 million Internet users aged 15 and older visited social networking sites in July 2010, representing 84% of its total Internet audience. India now ranks the seventh largest market worldwide for social networking, after the U.S., China, Germany, Russia, Brazil and the U.K. The total Indian social networking audience grew 43% in the past year, more than tripling the rate of growth of the total Internet audience in India. Facebook has the top spot among social networking sites, with 20.9 million visitors. Orkut ranks second with 19.9 million visitors (up 16% from past year), followed by BharatStudent.com with 4.4 million visitors [5].

Facebook achieved an astonishing growth on its active user base in the past months in Brazil and India. By keeping track of the data provided through Facebook Ads [8], it was seen that Brazil registered a growth of 167% on its reported active users base, while India registered a growth of 211%, both during the period comprehended between October 2010 and May 2011, as depicted in Figure 1. During the same period, Facebook also jumped from the 15^{th} to the 4^{th} spot in Brazil top sites by audience provided by Alexa[9], and also consolidated the 3^{rd} spot in India.

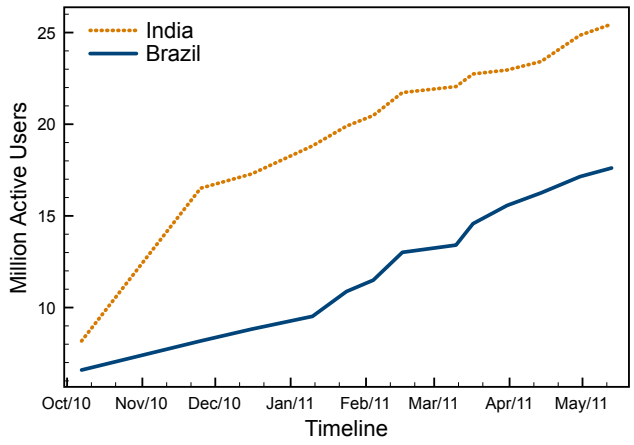

Figure 1: Reported growth of Facebook active users in Brazil and India [8].

3. METHODOLOGY

In this section it is presented how participant recruitment was promoted for the Facebook (FB) application developed for this particular study. The application itself is also depicted. The data collected is described in detail along with the demographics of participants and their friends.

3.1 Recruitment and demographics

To recruit participants to the study, emails and Facebook messages were sent, fliers were affixed on university notice boards around the globe, much word-of-mouth was made and also a couple of media posts were published. Most of the campaigning about the study was done in Brazil and India. A domain was registered *http : //www.theprivacystudy.org/* to host all information about the study. The website allowed participants to spread the word about the FB application through several online channels such as *Twitter* and *Google Buzz*. In the final count of the data that is hereby analyzed, the application received 540 likes on Facebook out of 664 users; 77% of the participants had at least one friend who also joined the study. Prizes were offered through raffle to participants, comprising one high-end mp3 player of 32 GBs and five games for PC/Mac. It was also noted "Get your friends to install the application and increase your chances to win a prize!" on the website to help attract more participation in the study.

[9]http://www.alexa.com

3.2 Study setup

Six hundred sixty-four participants installed the application in their Facebook account. When participants installed the *Privacy Study* application[10], they were presented with the study privacy policy which explains what kind of data will be collected from them and for what purpose. The application worked in conformity with the Terms of Service of Facebook and ethical ways of studying social media [10]. Participants were then invited to authorize the application access to their Facebook data and some pieces of information from their friends as well, as depicted in Figure 2. The following data items were collected from participants who installed the application:

- User and friends basic information (sex, birthdate, timezone, current city, hometown)
- User content privacy settings
- User content meta information (eg: album size, creation date, type, tags)

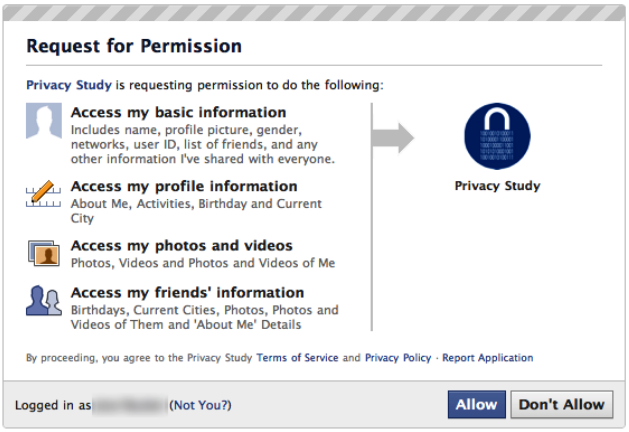

Figure 2: Permission dialog box presented to participants while installing the application. It is requesting the participant for accessing the account information. Blurred the user ID in the figure.

[10]http://www.facebook.com/apps/application.php?id=144781782200917

After installing the application, users were presented with a breakdown of the visibility of the content (photo albums, videos and links) that they have shared on Facebook (as shown in Figure 3). A pie chart shows the percentage of what is visible to *Everyone*, *Friends and Networks*, *Friends of Friends*, *Friends Only*, what has *Custom* visibility and what is available to *Self* only. Users can see their personal breakdown for photo albums, videos and links by clicking on the radio buttons placed beneath the pie chart. The application also displays the number of friends who have currently installed the application and the quantity of coupons for the prizes draw earned so far.

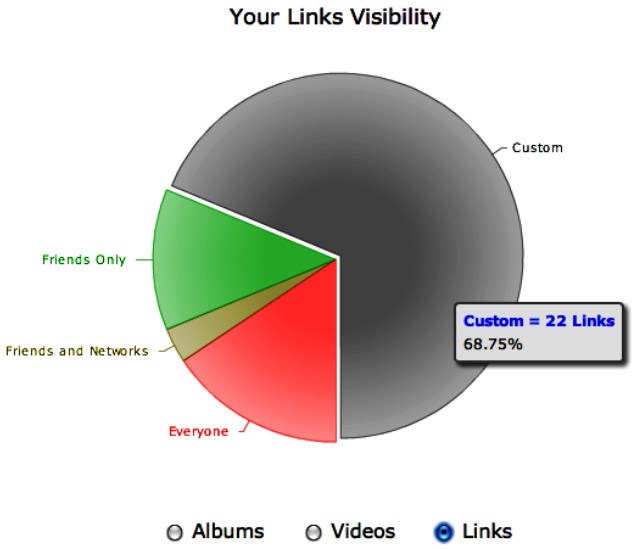

Figure 3: The Privacy Study application overview. It presents a breakdown of the visibility of the contents (photo albums, videos and links) that the participant has shared on Facebook. The application also portrays the number of friends who have installed the application and the quantity of coupons for the prizes draw earned (cropped from image).

3.3 Data set and Demographics

Two sets of data were collected, one about the participants who installed the application and another one about the friends of participants. Data set P represents the users who installed the study application and is summarized in Table 1. Both in Brazil and India male participants were more than female participants, which is opposite to the data that Boyd et al. used in their study [3]. According to Facebook Ads[8], there are far more male users (70%) in India compared to female users (30%), whereas there are more female users (54%) in Brazil than male (46%). An important percentage of participants (53%) in the study belonged to the age group 18 - 25, which is likewise the most active age group on Facebook. Young adults in this specific age strip are the most prevalent users of most Online Social Networks on the Internet.[11] Data set PF represents the participants along with their friends and its details are given in Table 2.

[11]http://social-media-optimization.com/2008/05/social-network-user-demographics/

Table 1: Demographics of the study participants (data set P). The data set comprehends participants from 25 countries. Data was not available for some users, presented as N/A in the table.

	Total N = 664	Brazil N = 341	India N = 235	Others N = 88
Gender (%)				
Female	25.15	29.33	18.3	27.27
Male	65.06	60.7	71.49	64.77
N / A	9.79	9.97	10.21	7.95
Age (%)				
Under 18	0.9	0.29	1.7	1.14
18 - 25	53.16	45.16	69.79	39.77
26 - 35	24.55	30.5	12.77	32.95
36 - 45	5.27	5.87	3.83	6.82
46 - 55	1.51	2.05	1.28	0.0
Over 55	0.45	0.29	0.85	0.0
N / A	14.61	16.13	10.64	19.32
User degree				
Average	207.17	155.2	252.39	287.56
Median	159	122	202	236
User content				
#Albums	3647	1512	1275	860
#Photos	74208	23632	20525	30051
#Links	18790	6444	8016	4330
#Videos	296	66	115	115

Table 2: Demographics of the study participants and their friends (data set PF). The data set comprehends users from 142 countries. 71379 users didn't reveal their country and are accounted in *Total* column only. Data was not available for some users, presented as N/A in the table.

	Total N=109832	Brazil N=9614	India N=16486	Others N=12645
Gender(%)				
Female	32.72	39.62	22.25	32.24
Male	54.74	50.77	70.01	52.00
N/A	12.55	9.61	7.74	15.75
Age(%)				
Under 18	2.36	1.03	3.05	2.56
18 - 25	36.39	27.92	50.87	25.6
26 - 35	17.98	26.05	6.32	16.36
36 - 45	4.34	5.87	1.17	4.36
46 - 55	2.01	2.43	0.73	1.69
Over 55	1.24	1.33	0.24	1.12
N/A	36.93	36.71	37.86	49.43

4. HYPOTHESES

Privacy settings is a means by which users set their preferences about how their profile or other information should be handled by the organization who is collecting the information (e.g. Facebook). Default settings are supposed to capture the most acceptable preferences so that users do not have to keep changing the default settings. It has been shown that 84% of the Facebook users did not change their default privacy settings [12]. It has been also shown that presenting information that is easily accessible to users can significantly aid the user to change their privacy settings on Facebook [21]. Acquisti et al. found that participants in their study had misconceptions about privacy on Facebook [1, 13].

Hypothesis 1: *The privacy settings of users is not significantly different from the default privacy settings in Facebook.*

From an individual standpoint, gender represents the least perilous personal information considered in the present study. Nonetheless, it represents an important piece for total information systems being built by marketeers, government agencies and criminal organizations. Gender is one piece of information (along with birth date and zip code) which can be used to identify a large percentage of U.S. citizens uniquely [23]. Gender can be fairly estimated using first name databases[12] and can even be predicted by image-based classifiers [11]. It has a specific privacy setting on Facebook, which is governed by a checkbox entitled "Show my sex in my profile" on the profile edit page. As one can guess, the default is set to visible. Users cannot control who can see the information, being it public available when disclosed.

Hypothesis 2: *A high percentage of Facebook users expose their gender publicly.*

No day goes by without media coverage about someone having her reputation disputed on the Internet. The video shooted or the picture taken of someone's private life might become public and make the headline or reach an unattended audience one day. Watching this recurrent situation serves as an alarm for people to try to control what they expose and share online and specially with whom. In a decreasing scale of reputation endangerment certain contents can be shared on Facebook: videos, photo albums and links. Facebook allows users to fine tune their privacy settings for every piece of content shared over its network. Nevertheless, to achieve broader audience and network growth, in despite of user reputation preservation, the default visibility is configured so *everyone* in the social network can access these contents. Even so, users are given the option to customize the visibility of their content with the following options: *Everyone*, *Friends and Networks*, *Friends of Friends*, *Friends Only*, *Custom* or *Self*.

Hypothesis 3: *Photo album is one of the features on Facebook that endangers users reputation the most.*

[12]http://www.socialsecurity.gov/OACT/babynames/

5. RESULTS

The results presented in this section support Hypotheses 1, 2 and 3.

5.1 H1: Default privacy settings

For this analysis, the privacy settings S_u of each user u of PF data set were converted to a 3-bit word, where $S_u = 111$ represents public disclosure of the studied attributes, according to the encoding presented in Table 3. This approach gives an overall idea of how users changed the privacy settings and it has been used in the past [12].

Table 3: Encoding for basic privacy settings S_u of a user u. Bit 1 denotes the leftmost bit in the representation.

Bit	Attribute	Description
1	Date of Birth	=1 if full date of birth is visible
2	Location	=1 if current location is informed
3	Gender	=1 if sex is revealed

Facebook has two possible default settings: $S_u = 101$, obtained after filling out the initial registration form (i.e. DOB and gender visible to everyone), and $S_u = 111$, if the user informs his current location afterwards. As depicted in Figure 5, the majority (58.1%) of users keep the two aforementioned default settings ($S_u = 1*1$), where 39% of users remain with the initial registration settings ($S_u = 101$). Only 4.1% of all users appear to be really concerned of exposing their information and do take the time to conceal all the three pieces under analysis ($S_u = 000$).

In Brazil, the scenario exhibits 67.1% of users with one of the two possible default settings. In India, 22.1% of users disclose all the three attributes, while the equivalent statistic in Brazil is somewhat lower, 14.3%. The precedent difference is statiscally significant (Proportion Test, p-value $<$ 0.05). This supports Hypothesis 1 and reinforces the fact that simply providing a set of customization features does not ensure that users will take advantage of them [22].

Taking the gender influence into consideration the encoding possibilities are reduced by half, as the third bit is always 1. Among the remaining configurations it is considered the two extremes: $S_u = 001$, which stands for total concealment, and $S_u = 111$, which stands for total revelation. For all cases, women are significantly more conservative than men, as shown in Table 4. For instance, 24.1% of the women overall do not disclose their full date of birth neither their current location ($S_u = 001$). The same statistic reaches a lower portion of men, 17.2% overall.

Table 4: Reach across different genders for both extreme privacy settings S_u of users in PF data set with 95% confidence level.

	S_u	Female Avg(%) ± Std. Error	Male Avg(%) ± Std. Error
Brazil	111	13.89 ± 0.53	18.14 ± 0.55
	001	23.78 ± 0.65	15.19 ± 0.51
India	111	19.86 ± 0.76	26.54 ± 0.52
	001	25.3 ± 0.82	17.53 ± 0.45
All	111	18.15 ± 0.40	23.99 ± 0.34
	001	24.09 ± 0.44	17.19 ± 0.30

5.2 H2: Gender exposure

Using the PF data set, it was found that 87.5% of users revealed their gender information, the highest exposure level among factors studied in this research. Figure 4 presents the gender breakdown and exposure of participants and their friends. Concealment of gender information is low in all regions considered (< 16%). Such high levels of exposure obtained reinforce the assumption that users are diving in a sea of obliviousness towards privacy. Moreover, the information is publicly available, what corroborates the importance of default settings. Research has shown that most people rarely change them [3]. The results support Hypothesis 2.

5.3 H3: Content exposure

Facebook allows users to fine tune their privacy settings for every piece of content (e.g. photo albums, links, videos) shared over its network. In order to get this content to broader audiences and increase network value, the default visibility is configured to *Everyone*. Users are given the option to customize the visibility of the content with the following options: *Everyone*, *Friends and Networks*, *Friends of Friends*, *Friends Only*, *Custom* or *Self*. For the present analysis, user content privacy settings were obtained from P data set. As summarized in Table 1, only a few hundred videos were uploaded and shared on Facebook by the study participants, demonstrating that the feature is not so popular as in other video specific platforms (e.g. YouTube). For this reason, videos will be left out of the present article and focus will be given to photo albums and links.

Users have four different kinds of photo albums on Facebook: *profile pictures*, *wall photos*, *mobile uploads* and *normal ones*. The first three kinds are unique per user, while the latter serves for the general purpose and can be created at will. Although the visibility of *profile pictures* album can be configured in the same way as the others, the privacy data obtained through the Facebook API does not reflect that - being it always set to *Self*, and they are thus left out of the investigation.

The examination commences by taking an overall look on content visibility. For that, the entire content set was broken down by their current visibility setting, on a per content basis. For photo albums, the only regional statistically significant difference found was for the *Custom* setting (Proportion Test, p-value < 0.05), where India reaches 10% and Brazil stays around 5%. More important are the contrasts found between visibility levels, where the *Friends Only* setting appears as the first choice among all participants (Proportion Test, p-value < 0.05), showing a clear discontentment with the default exposure to *Everyone*.

As depicted in the upper half of Figure 6, the album visibility breakdown clearly diverges from its link counterpart. While the exposure of photo albums to *Everyone* reaches 35% overall, the equivalent statistic for links notches 56%. The situation is more accentuated in Brazil, where the gap attains 27% (42% for albums and 69% for links).

Another interesting way to look at the content exposure scenario is to analyze visibility usage reach. For that examination a bucket is created for each possible setting and each user is placed once in every bucket for which he has a content shared with that particular setting. So, for instance, if user u_1 has two albums shared with *Everyone* and two other albums shared with his *Friends Only*, he is placed in buckets *Everyone* and *Friends Only* a single time. The results for visibility usage reach are depicted in the lower half of Figure 6. The visibility for *Everyone* reaches 82.5% of all users for links and 55.8% for photo albums. Another clear display of preoccupation towards photo albums exposure is given by the *Friends Only* setting, where the usage reaches 55.5% for albums and 45% for links. The gap is even more prominent in Brazil and India, where it reaches 19.9% for the former and 29.6% for the latter. The precedent differences are all statiscally significant (Proportion Test, p-value < 0.05). These results support Hypothesis 3.

6. DISCUSSION AND IMPLICATIONS

Privacy can be understood as the ability to one person control the access to information about herself. In most cases portrayed throughout this paper these controls are present but go unnoticed or are often misunderstood by end users. People want freedom to express themselves in the digital era without having to deal with the enormous complexity of current privacy control designs. The default privacy settings of an Online Social Network plays a key role to preserve its users' reputation and shall not be subject to commercial interests solely. It was found that very few users change their default privacy settings. Thereupon, OSNs need to pay more attention when designing their defaults to best serve their users' privacy protection.

As it was demonstrated, only a small part of users appear to be really aware of the consequences of permissive settings and their pervasive consequences and therefore do not reveal any of the personal identifiable information under study. It was also shown that majority of the users reveal publicly their gender on Facebook.

Finally, it was shown that users exercise more control over content with more potential to endanger their reputation. For instance, the exposure of photo albums was oftenly configured to reach a narrower audience than links. An implication of that result would be to modify the OSN default visibility setting according to the content being shared, instead of having a single rule for every content kind. In other words, photo albums should not be made visible to *everyone* by default.

As in any real world experiment, a few limitations were present. The data was collected mainly through acquaintances and people whom it was able to reach through emails and fliers. Therefore, the sample obtained is a convenient

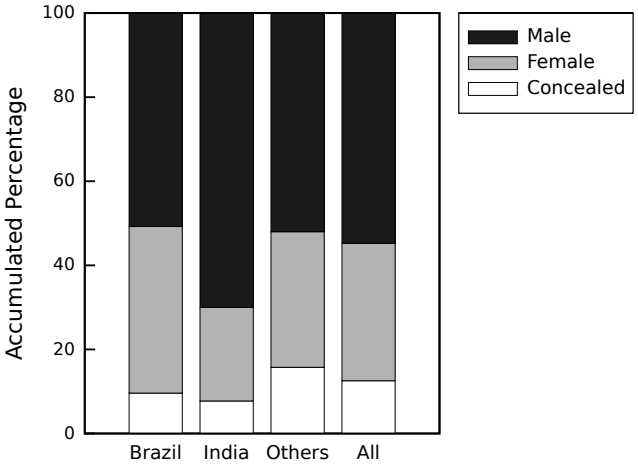

Figure 4: Gender breakdown and exposure of participants and their friends. About 87% of users expose their gender to their network. Used PF data set for the analysis.

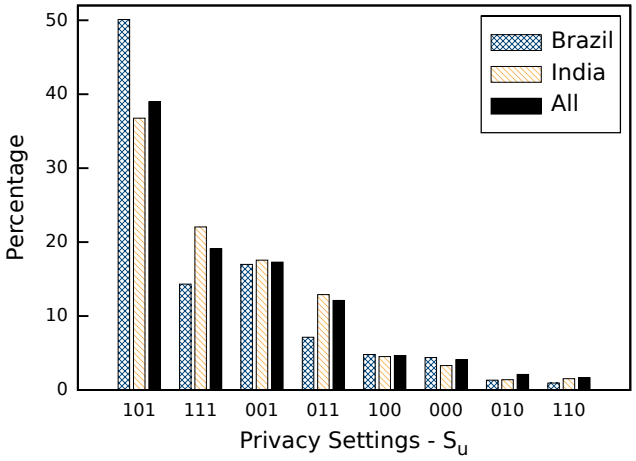

Figure 5: Distribution of the privacy settings S_u of users in PF data set; encoding used here is according to Table 3. Only 4% of the users conceal all the three pieces of information under analysis ($S_u = 000$).

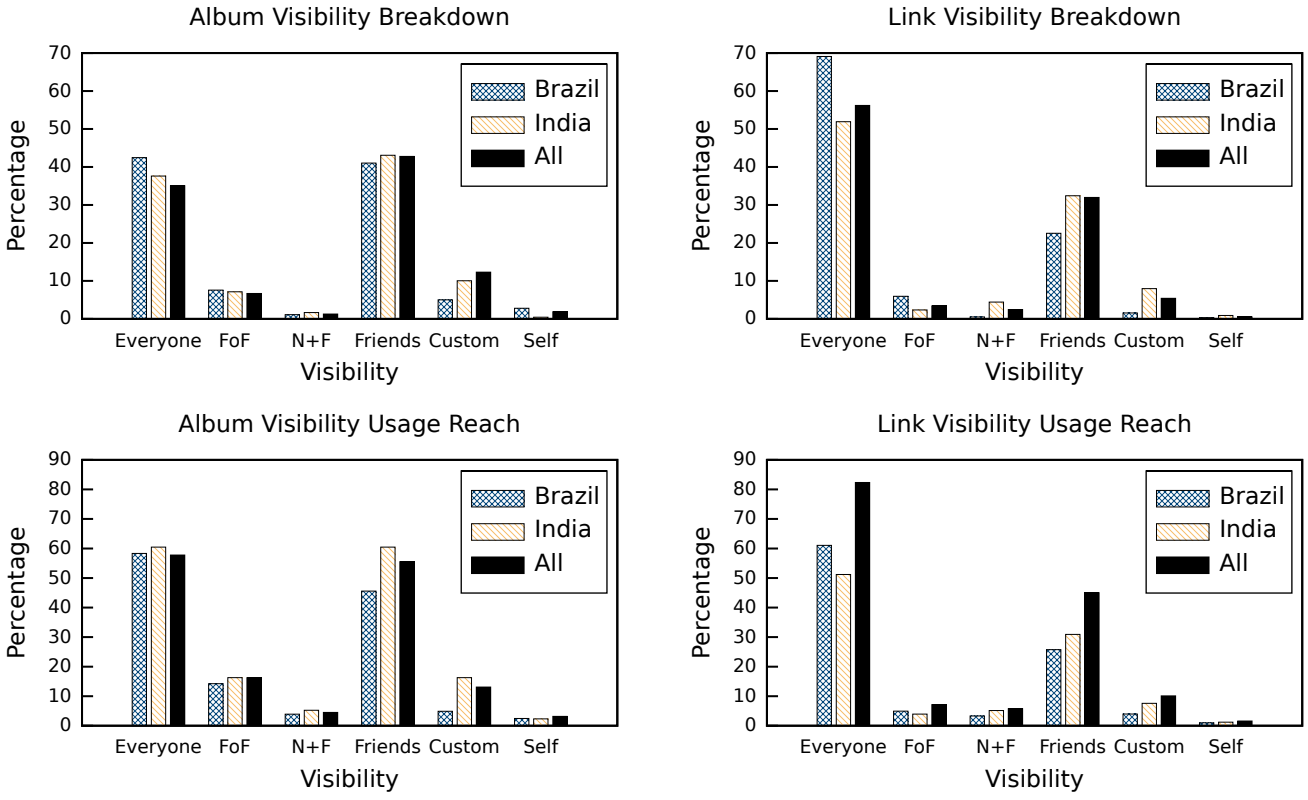

Figure 6: On all charts *FoF* is the acronym for *"Friends of Friends"*, *N+F* for *"Networks and Friends"* and *Friends* stands for *"Friends Only"*. The two charts at the top contrast the visibility breakdown between photo albums and links. As it can be seen, users are overall more concerned about their photo albums exposure, what is shown by higher peaks for *Friends*. This message is reinforced by the charts at the bottom, where is displayed the usage reach for every possible setting. The visibility to *Everyone* reaches more than 82% of the participants for links and 58% for albums.

sample, so the results may not be generalizable to all Facebook users. It is also understood that conducting such a study where the users are statistically representative of the country or group of interest is difficult to achieve. There is also a plan to study the longitudinal effects on the privacy settings on Facebook over the time.

Acknowledgments

This work was supported by the Indo-Brazil Science Council.

7. REFERENCES

[1] A. Acquisti and R. Gross. Imagined communities: Awareness, information sharing, and privacy on the facebook. In *In 6th Workshop on Privacy Enhancing Technologies*, pages 36–58, 2006.

[2] S. Bellman, E. J. Johnson, S. J. Kobrin, and G. L. Lohse. International Differences in Information privacy concerns: A global survey of consumers. *The Information Society*, 20:313 – 324, 2004.

[3] D. Boyd and E. Hargittai. Facebook privacy settings: Who cares? *Journal on the Internet*, 15(8), 2010.

[4] J. Chang, I. Roseen, L. Backstrom, and C. Marlow. ePluribus: Ethnicity on Social Networks. In *Proceedings of the Fourth International Conference on Weblogs and Social Media*, 2010.

[5] ComScore. Facebook captures top spot among social networking sites in india. Press Release, August 2010.

[6] ComScore. Orkut continues to lead brazil's social networking market, facebook audience grows fivefold. Press Release, October 2010.

[7] S. Diller, L. Lin, and V. Tashjian. The evolving role of security, privacy, and trust in a digitized world. pages 1213–1225, 2003.

[8] Facebook. Facebook ads. Website. http://www.facebook.com/ads.

[9] Facebook. Facebook statistics. Website. http://www.facebook.com/press/info.php?statistics, Retrieved Oct 11, 2010.

[10] D. Fisher, D. W. McDonald, A. L. Brooks, and E. F. Churchill. Terms of service, ethics, and bias: Tapping the social web for cscw research. *Computer Supported Cooperative Work (CSCW), Panel discussion*, 2010.

[11] A. C. Gallagher. Estimating age, gender, and identity using first name priors. In *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008.

[12] M. Gjoka, M. Kurant, C. T. Butts, and A. Markopoulou. A Walk in Facebook: Uniform Sampling of Users in Online Social Networks. Technical report, arXiv.org, 2009.

[13] R. Gross and A. Acquisti. Information revelation and privacy in online social networks. *Proceedings of the 2005 ACM workshop on Privacy in the electronic society*, pages 71–80, 2005.

[14] G. Hofstede. *Cultural and Organizations - Software of the Mind - Intercultural Cooperation and its importance for survival*. 1991.

[15] B. Krishnamurthy and C. Wills. Characterizing privacy in online social networks. *Proceedings of the first workshop on Online social networks*, (37–42), 2008.

[16] P. Kumaraguru and L. Cranor. Privacy in India: Attitudes and Awareness. In *Proceedings of the 2005 Workshop on Privacy Enhancing Technologies (PET2005)*, 30 May - 1 June 2005.

[17] P. Kumaraguru and L. F. Cranor. Privacy indexes: A survey of westin's studies. Technical report, Carnegie Mellon University, 2005.

[18] P. Kumaraguru, L. F. Cranor, and E. Newton. Privacy perceptions in india and the united states: An interview study. *In The 33rd Research Conference on Communication, Information and Internet Policy (TPRC)*, September 2005.

[19] P. Kumaraguru, S. Sheng, A. Acquisti, L. F. Cranor, and J. Hong. Lessons from a real world evaluation of anti-phishing training. *e-Crime Researchers Summit, Anti-Phishing Working Group*, October 2008.

[20] K. Lewis, J. Kaufman, M. Gonzalez, A. Wimmer, and N. Christakis. Tastes, ties, and time: A new social network dataset using facebook.com. *Social Networks*, 30(4):330 – 342, October 2008.

[21] H. R. Lipford, A. Besmer, and J. Watson. Understanding privacy settings in facebook with an audience view. In *UPSEC'08: Proceedings of the 1st Conference on Usability, Psychology, and Security*, pages 1–8, Berkeley, CA, USA, 2008.

[22] W. E. Mackay. Triggers and barriers to customizing software. In *CHI '91: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 153–160, New York, NY, USA, 1991. ACM.

[23] L. Sweeney. k-anonymity: a model for protecting privacy. *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, 10(5):557–570, 2002.

[24] X. Wei, J. Yang, and L. Adamic. Diffusion dynamics of games on online social networks. *3rd Workshop on Online Social Networks*, 2010.

[25] D. Wilson and R. Purushothaman. Dreaming with brics: The path to 2050. Technical report, Golman Sachs, 2003.

[26] B. Zhou and J. Pei. Preserving privacy in social networks against neighborhood attacks. In *ICDE '08: Proceedings of the 2008 IEEE 24th International Conference on Data Engineering*, pages 506–515, 2008.