Who Is Concerned about What? A Study of American, Chinese and Indian Users' Privacy Concerns on Social Network Sites (Short Paper)

Yang Wang, Gregory Norcie, Lorrie Faith Cranor

School of Computer Science
Carnegie Mellon University, U.S.A.
{wang,ganorcie,lorrie}@cs.cmu.edu

Abstract. We present a study that investigates American, Chinese, and Indian social networking site (SNS) users' privacy attitudes and practices. We conducted an online survey of users of three popular SNSs in these countries. Based on 924 valid responses from the three countries, we found that generally American respondents were the most privacy concerned, followed by the Chinese and Indians. However, the US sample exhibited the lowest level of desire to restrict the visibility of their SNS information to certain people (e.g., co-workers). The Chinese respondents showed significantly higher concerns about identity issues on SNS such as fake names and impersonation.

Keywords: Social Networking Site, Privacy, Trust, Culture, Survey

1 Introduction

Social networking services (SNSs) have become a global phenomenon. For instance 70% of Facebooks 500 million users are located outside the United States [8]. Other SNS sites tend to dominate various parts of the world, such as CyWorld in Korea, and Orkut in Brazil. Meanwhile, privacy issues in SNS have been hotly discussed in public media, particularly about Facebook in the US media [9]. Despite the steady rise of SNS worldwide, there is still little understanding of SNS privacy in other countries, especially non-Western developing countries.

Several studies have shown that general online privacy concerns [3], SNS usage patterns [12], and even privacy policies of SNS platforms [4] vary across different countries. We hypothesize that cultural differences may affect how SNS users perceive and make privacy-sensitive decisions. While we recognize that culture is fluid, dynamic, and often difficult to define, we chose to take Hofstede's approach [10] in using country as a proxy for culture in our study. This is because users from the same country usually use the same SNS. We conducted a multi-national survey to investigate SNS users' privacy attitudes and practices in three countries: China, India, and the US. To the best of our knowledge, this is the first empirical study that investigates users' attitudes about SNS privacy across countries.

2 Survey

2.1 SNS Sites and Respondents

We chose three SNSs for this study: Facebook, Renren, and Kaixin001. According to Alexa (as of August 3, 2010), Facebook has the highest traffic among SNS sites in both the US and India, while Renren.com and Kaixin001.com (two domestic Chinese SNS sites) are the top two SNS sites in China [2]. Since they were very close in terms of traffic, we decided to include both Chinese sites in our study. The three selected sites share several common features such as profiles, walls, photo sharing, shared links, and games. Also the sites are primarily geared towards personal or leisure use; and they support third-party application development on their platforms.

We recruited Facebook users who reside in the US or India, and Kaixin001 and Renren users who reside in China. We recruited only users who were 18 years or older.

2.2 Survey Design

The survey was designed to gain a better understanding of SNS users' demographics and SNS usage patterns, to elicit their attitudes towards sharing information on the Internet and on SNS, and to investigate their privacy-related experience and behavior on SNS. The survey has 10 questions covering demographic information and SNS usage patterns, 2 multiple-choice questions, 8 open-ended questions, and 69 Likert-scale questions covering privacy-related attitudes and practices with SNS.

2.3 Survey Administration

The survey was developed in an English master version. The master version was then translated to Simplified Chinese. We deployed three versions: an English version for American and Indian Facebook users, and two Simplified Chinese versions for Kaixin001 and Renren users. All three versions were hosted on SurveyGizmo for about two weeks in July 2010. We recruited our participants from crowd sourcing sites. American and Indian participants were recruited from Amazon's Mechanical Turk (MTurk). An accepted valid response received 50 US cents. Since the survey would take about 10-15 minutes to finish, our compensation rate was about $2-3 per hour, which is on a par with the normal hourly pay on MTurk. Similarly, we recruited our Chinese participants from a Chinese crowd sourcing site zhubijie.com (ZBJ) and each accepted valid response received 3 RMB (roughly 50 cents).

We acknowledge that this methodology is subject to self-selection bias and any bias that may reside in the recruiting sites. Therefore, we cannot make claims about whether our sample is a representative sample of the SNS users in the three countries.

2.4 Data Cleaning and Analysis

By July 27, 2010, we received 343, 354, and 355 complete responses from China, India, and the US, respectively. We used a combination of measures to help determine

whether our participants from these crowdsourcing sites were taking their tasks seriously. We paid attention to unusually short completion time (two standard deviation from the mean), inconsistent answers (we asked the same underlying question with slightly different wordings), and verbatim repetition or nonsense free-response answers. After manually checking the answers and filtering out suspect participants, we were left with 321 valid responses from the US, 312 from India, and 291 from China. Our Chinese sample consists of 138 Kaixin001 users and 153 Renren users.

We analyzed the data using a series of Analysis of Covariance (ANCOVA) with country as the independent variable. The dependent variables were the answers to the 69 privacy-related Likert scale questions. Age, gender, educational level, computer/IT career experience, the SNS site used, and the frequency of visiting the SNS site were used as covariates. ANCOVA was used because our dependent variables are intervals (Likert scales) and we have both categorical (e.g., country, gender) as well as continuous independent variable or covariates (e.g., age).

3 Survey Results

Overall, participants from the three countries exhibited very different privacy attitudes, concerns, and behaviors on SNS. Since the three country samples also differed significantly in terms of their demographics, we controlled for individual characteristics such as age and gender. On nearly every question, the results differed significantly by the country variable. In general, US users tend to be the most privacy concerned among all, followed by Chinese users. Indian users were the least privacy concerned. Compared with the US and Indian samples, the two Chinese sub-samples (Kaixin001 and Renren users) were more similar than different for most measures. Therefore, we do not report their results separately, but instead report them collectively as the Chinese sample. Since we had a relatively large sample and many measures exhibited statistically significant difference by country, we paid most attention to measures where the results were particularly interesting, results diverted from the general pattern (US $>$ China $>$ India), or measures where country was not the most important predictor of the results. More detailed results of this survey can be found in [13].

3.1 Demographics

Valid answers from 924 respondents were used for the analysis. Table 1 presents the demographic make-up of our sample.

3.2 Privacy Attitudes of Personal Content on SNS

We asked how comfortable participants would be with everyone on the Internet seeing content from their SNS profiles such as their walls and status updates (on a 7-point likert scale). Figure 1 shows the percentages of respondents who had each level of comfort with everyone seeing these types of content for our three country samples. Content types listed from top to bottom in the figure generally followed an increasing order of privacy sensitivity. We can observe that the three country samples largely agreed on

Table 1: Demographics of our study participants:

Note: *, **, *** statistical significance at p<.05, .001, .0001

		China	India	US
Sample size		291	312	321
Gender ***	Men	56.4%	60.9%	36.4%
	Women	43.6%	39.1%	63.6%
Age ***	Mean	23.5	27.1	31.4
	SD	3.8	6.7	11.0
IT education or career ***	IT	41.6%	65.7%	12.1%
	Non-IT	58.4%	34.3%	87.9%
At least some college education		88.3%	90.7%	86.0%

the privacy sensitivity ranking of different types of content. For nearly all items, US repondents were the most privacy concerned, followed by the Chinese and the Indian repondents, and we found statistically significant differences (at least p<.001) among the three countries (US > China > India). Phone number, residence street address, email address, photo, employer were considered as privacy sensitive by more than half of both the US and Chinese respondents. However, only phone number was considered privacy sensitive by more than half of the Indian respondents. The list of privacy sensitive items considered by our US sample is similar to prior research [1].

We tested the inter-item reliability of the 16 content items using Cronbach's alpha [7]. The alpha value indicates to what extent questions measure the same underlying concept (or how they correlate with each other). Usually a scale is considered consistent or reliable if the alpha value is above .7. The alpha value of these 16 items is .94, indicating they are reliable in measuring the sensitivity of various content. We then computed a privacy sensitivity score for each respondent by averaging his or her answers to these 16 questions. The higher the score, the more privacy sensitive this person is with regard to the information she posts on the SNS. As Table 2 shows, the privacy sensitivity scores of the American sample (mean=4.7, sd=1.5) were significantly higher (ANCOVA, p<.0001) than that of the Indian sample (mean=3.3, sd=1.1) and Chinese sample (mean=4.2, sd=1.1). Notably, technical knowledge, gender, age and frequency of visit (in decreasing order of significance) were also statistically significant predictors. Although not universally, users without technical knowledge, female users, older users, and less frequent users are less likely to be comfortable with anyone seeing their data than their counterparts. For instance, we found that older users tend to be more uncomfortable if anyone can see their religious views.

3.3 Privacy Concerns on SNS

The privacy sensitivity dimension models how users control or decide what to post on SNS sites. To assess respondents' privacy concerns about what others can do with their data on SNS sites, we asked 15 questions such as whether the site has too much information about you, whether you are concerned that the site shares your information

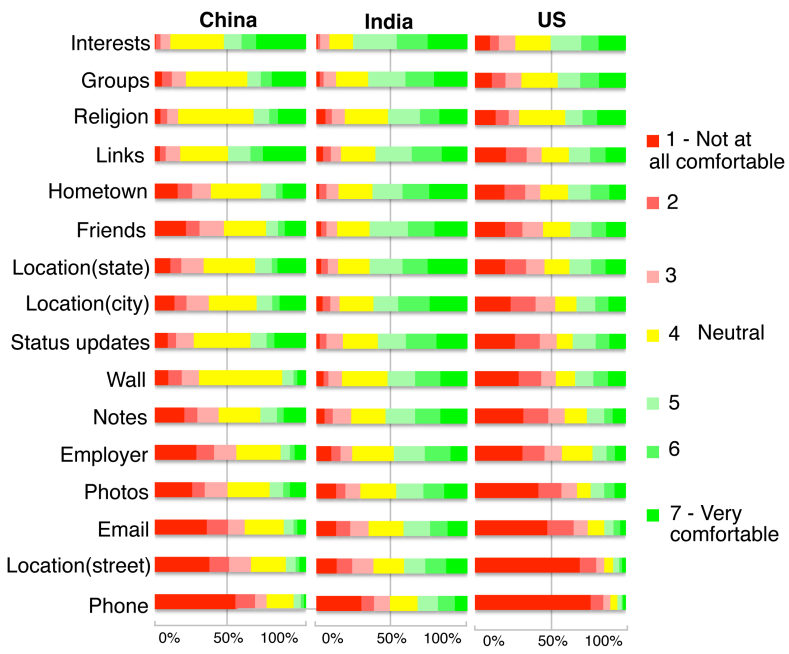

Fig. 1: Privacy Attitudes of Personal Content on the SNS (7-point likert scale)

with third parties, and whether your data on the site is secure. Again, we tested the inter-item reliability of the 15 questions using Cronbach's alpha, and the alpha value is .87 indicating they are consistent. We then computed a privacy concern score for each respondent by averaging his or her answers to these 15 questions. The higher the score, the more privacy concerned this person is with regard to SNS. As Table 2 shows, the privacy concern scores of the American sample (mean=5.0, sd=1.0) were significantly higher (ANCOVA, $p<.001$) than that of the Indian sample (mean=4.6, sd=0.9) and Chinese sample (mean=4.8, sd=0.9). While the same pattern seemed to apply to SNS privacy concerns (US $>$ China $>$ India), the Indian respondents were more worried about what others can do with their data (e.g., other people posting false information about the user) on the SNS than their own postings (i.e., having everyone on the Internet seeing their information) on the SNS.

3.4 Trust in SNS System or Operator

We examined to what extent SNS users trust the SNS system and operator. We asked four questions such as whether the particpant trusts the SNS with their personal information, and whether the participant feels the SNS employs trustworthy staff. We tested the inter-item reliability of the four questions using Cronbach's alpha, and the alpha value is .75 indicating they are consistent. We then computed a lack-of-trust score for each respondent by averaging his or her answers to these four questions. The higher the score, the less trust this person has with regard to the SNS system and operator. As Ta-

Table 2: Dimensions of SNS Privacy Attitudes (7-point likert scale)

Note: *, **, *** statistical significance at p<.05, .001, .0001

		China	India	US
Privacy sensitivity score ***	Mean	4.2	3.3	4.7
	SD	1.1	1.1	1.5
Privacy concern score ***	Mean	4.8	4.6	5.0
	SD	0.9	0.9	1.0
Lack-of-trust score ***	Mean	3.4	3.2	4.5
	SD	1.0	1.0	1.2
Desire-to-restrict score *	Mean	4.8	4.6	4.2
	SD	1.2	1.2	1.4

ble 2 shows, American sample had the highest lack-of-trust score (mean=4.5, sd=1.2) in the SNS system and operator, followed by the Chinese sample (mean=3.4, sd=1.0) and Indian sample (mean=3.2, sd=1.0). The differences in this score were strongly statistically significant (p<.0001) among the three country samples (US > China > India).

3.5 Desire to Restrict Information on SNS

To gauge user's desire to control the visibility of their information on SNS, we asked respondents five questions about whether they want to restrict some of their information so that certain people (parents, family, or co-workers) cannot see it or only certain people can see it. We tested the inter-item reliability of the five questions using Cronbach's alpha, and the alpha value is .77 indicating they are consistent. We then computed a desire-to-restrict score for each respondent by averaging his or her answers to these five questions. The higher the score, the more desire this person has to restrict the visibility of her information on the SNS site. Interestingly, we observed a different pattern (China > India > US) than from previous privacy measures. As Table 2 shows, the Chinese sample had the highest desire-to-restrict score (mean=4.8, sd=1.2), followed by the Indian sample (mean=4.6, sd=1.2) and American sample (mean=4.2, sd=1.4). The differences in this score were statistically significant (p<.05) among the three countries.

3.6 Fake Names and Impersonation

One of our research team members uses Kaixin001 and Renren, and has noticed that some friends use fake names. To investigate whether this is common in the Chinese SNS sites, we asked about instances of fake names and concern about impersonation.

Table 3 shows that the Chinese sample had the highest percentage of respondents that have friends who use fake names, followed by the Indian and US samples. Similarly, the Chinese sample had the highest percentage of respondents who were concerned about impersonation, followed by that of the US and Indian samples. Both differences were statistically significant (chi-square tests).

Table 3: Fake Names and Impersonation

Note: *, **, *** statistical significance at p<.05, .001, .0001

	China	India	US
Have friends use fake names ***	45.7%	39.3%	18.5%
Concerned about impersonation ***	36.3%	19.4%	28.6%

4 Discussion

Our American, Chinese, and Indian SNS respondents had significant differences in their privacy-related attitudes and behaviors. We observed a recurring pattern - the US sample was the most privacy concerned, followed by the Chinese and Indian samples. According to Hofstede's measurement on individualism almost three decades ago, the US is more of an individualistic society and thus values more personal privacy, whereas China is more of a collective society and India is somewhere in between [10]. While these individualism scores may help explain why American respondents were more privacy concerned than the Chinese respondents, they alone cannot explain why Indian respondents were the least privacy concerned. We also suspect that the recent intensive media focus on Facebook privacy raised American users' awareness of these issues and their concerns, but we could not measure this effect in the current study.

We observed this pattern (US $>$ China $>$ India) in many privacy aspects such as how comfortable respondents were with everyone on the Internet seeing their different types of information on SNS (privacy sensitivity score), how concerned they are about what other people can do about their SNS data (privacy concern score), and how much they trust the SNS system and operator (lack-of-trust score). However, this pattern was reversed when it comes to users' desires to restrict their information on SNS so that either certain people (friends, family, coworkers) cannot see or only certain people can see (desire-to-restrict score). Somewhat surprisingly, the Chinese sample had the highest level of desire, while the US sample had the lowest. One possible explanation is that American users may be more privacy concerned with regard to the site operator and businesses than with their interpersonal relationships. Another possible explanation is that American users post less sensitive information and therefore have less need for restrictions.Further work is needed to investigate these possibilities.

Prior research on Chinese SNS users suggests that they use SNS as a venue for meeting new people and for entertainment [5, 6]. This may explain why they are generally not very privacy concerned. However, our results show that they were particularly concerned about identity issues on SNS such as fake names and impersonation. We suspect this may be due to strict government regulations and monitoring and thus they tend to use anonymous or pseudonymous identities online. This in turn makes establishing online trust challenging. Our results with the Indian SNS users were largely similar to a previous study [11] that shows that Indian users are less privacy concerned than American users.

Our results suggests that different users may have varying priority or emphasis with regard to their privacy because of their cultural background and individual characteristics. Designers of privacy management tools need to take into account these factors.

Personalized privacy tools that learn a user's cultural background, characteristics and routine patterns of privacy decisions over time seem to be a promising direction.

References

1. Ackerman, M.S., Cranor, L.F., Reagle, J.: Privacy in e-commerce: examining user scenarios and privacy preferences. In: Proceedings of the 1st ACM conference on Electronic commerce. pp. 1–8. ACM, Denver, Colorado, United States (1999)
2. Alexa: Alexa: Top sites on the web (Aug 2010), http://www.alexa.com/topsites/
3. Bellman, S., Johnson, E., Kobrin, S., Lohse, G.: International differences in information privacy concerns: A global survey of consumers. The Information Society 20, 313–324 (Dec 2004)
4. Bonneau, J., Preibusch, S.: The privacy jungle: On the market for data protection in social networks. In the 8th Workshop on the Economics of Information Security (WEIS) (2009), http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.153.7796
5. China Internet Network Information Center: Survey report on Chinese SNS users. Tech. rep. (2009), www.cnnic.net.cn/uploadfiles/doc/2009/11/11/142713.doc
6. China Internet Network Information Center: Statistical survey report on Internet development in China. Tech. rep. (Jan 2010), http://www.cnnic.net.cn/uploadfiles/pdf/2010/3/15/142705.pdf
7. Cronbach, L.: Coefficient alpha and the internal structure of tests. Psychometrika 16, 297–334 (1951)
8. Facebook: Facebook statistics (Aug 2010), http://www.facebook.com/press/info.php?statistics
9. Gates, G.: Facebook Privacy: A Bewildering Tangle of Options. The New York Times (May 2010), http://www.nytimes.com/interactive/2010/05/12/business/facebook-privacy.html
10. Hofstede, G.H.: Culture's consequences: international differences in work-related values. Sage Publications (1980)
11. Kumaraguru, P., Cranor, L.: Privacy in India: Attitudes and Awareness. Proceedings of the 2005 workshop on privacy enhancing technologies (PET05) (2005)
12. Vasalou, A., Joinson, A.N., Courvoisier, D.: Cultural differences, experience with social networks and the nature of "true commitment" in facebook. International Journal of Human-Computer Studies 68(10), 719–728 (Oct 2010), http://www.sciencedirect.com/science/article/B6WGR-508K86F-1/2/c92ba5dbed69ed22b02e92f45f88511d
13. Wang, Y., Norcie, G., Cranor, L.F.: Who is Concerned about What? A Multi-National Study of Users' Privacy Concerns on Social Networking Sites. CyLab Technical Report CMU-CyLab-11-006 (2011), http://www.cylab.cmu.edu/files/pdfs/tech_reports/CMUCyLab11006.pdf