Predictors of Social Networking Privacy Behaviour

Sam Kavanagh

October 23, 2013

Abstract

While the global uptake of social networking sites continues unabated, it is important to consider the effect that this increasing level of diversity (both cultural or otherwise) has on user privacy concerns and behaviour. Over the course of this paper a wide range of potential predictors of this behaviour are discussed, including the evaluation of two privacy-related paradoxes. There exists much need for future research into several of these predictors, especially into the direct relationship between user privacy *concerns* and *behaviour*. Nonetheless, the findings outlined over the course of this paper demonstrate the existence and diversity of an array of internal and external factors which greatly impact upon the privacy behaviour of social network users.

1 Introduction

The rapid uptake of social networking site usage is both widely documented and near incomparable. Facebook alone currently boasts a user base reaching comfortably into the billions, with over 70% of their users now coming from *outside* of the USA [4]. Despite its global popularity, Facebook is in fact not the most popular social network in many large countries. The social networking sites Orkut, CyWorld and MeinVZ dominate the marketplaces of Brazil, Korea and Germany respectively [4, 3].

Social networking sites exist in what has been termed a 'nonymous' environment whereby relationships are anchored, as is the data and information that is

posted [3]. Thus, the privacy risks associated with the usage of social networking sites has been the subject of increasing attention in both scientific journals and mainstream media [3]. Unfortunately however, despite the aforementioned widespread *global* uptake of social networking this attention is typically targeted towards a small (western) subset of users; with studies showing that the privacy issues associated with the checked and unchecked publishing of data on social networks has received little-to-no attention in developing countries [4]. Moreover the information that is published generally does not take into account the wide variety of cultural and other external factors that greatly impact upon a persons conception of what actually *constitutes* private information [2].

Over the course of this paper, we investigate which factors are the major predictors of privacy attitudes and behaviour on social networking sites. We first discuss a small subset of the commonly cited dangers associated with the usage of social networking sites, before attempting to discern measurable bounds for which cultural differences can be observed. There exists two distinct and widely discussed privacy-related paradoxes in existing literature. The control paradox refers to the idea that by *increasing* a users control over their privacy, they often reveal *more* information than they would with lesser control. The privacy paradox pertains to the belief that users often *do* care about their privacy, however they do *not* in practice do anything about it. The accuracy of these paradoxes are also investigated over the course of this paper in order to determine their respective feasibilities as predictors of social networking privacy behaviour.

Finally, a wide range of non-cultural factors that are intuitive predictors of privacy behaviour are discussed, for we cannot accurately provide privacy awareness and education until we truly understand at an individual or at least cultural level what actually constitutes private information.

2 The Risk

Before discussing the likely predictors of social networking privacy behaviour, it is important to first briefly describe what is actually at risk (i.e. the potentially adverse consequences) when posting information on social networks.

There exists a variety of traditionally discussed and intuitive negative consequences that can be occur as a result of publishing information on social networks [3, 2]. These range from social networking attacks such as imperson-

ations or even identity theft, to more technical attacks such as the extraction of exif data from images. The risks associated with the checked and unchecked publishing of data on social networks are not purely limited to the attacks of malicious users however, but instead there exists an increasing trend for adverse consequences occurring as a result of legitimate entities (such as prospective employers) obtaining 'private' information [1].

Nonetheless, as both the quantity of social networks users - as well as their respective information publishing continues to increase, so too does the complexity of the attacks (both social networking and otherwise) employed by malicious entities. In order to represent the great deal of potential dangers now associated with this information publishing, Creese et al. [2] developed what they termed a data-reachability matrix for elucidating privacy information on social networks (Figure 1).

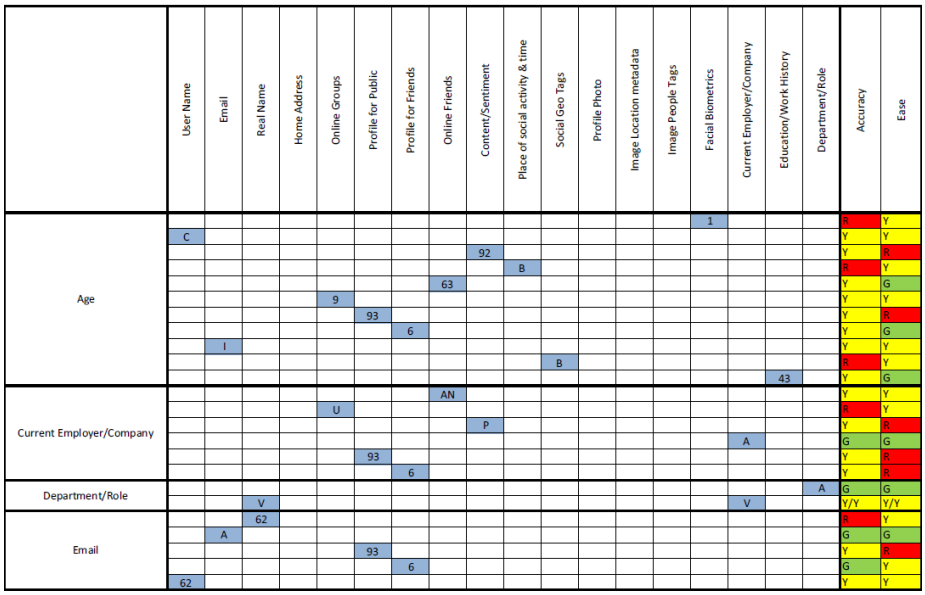

Figure 1: A small subsection of the data-reachability matrix developed by Creese et al. [2]

This matrix attempts to capture the vast array of information that it is potentially possible for attackers to *derive* given only a small set of initial data, as well as the accuracy and ease associated with doing so. Creese et al. believe that the information contained within this matrix is so compelling that it could

potentially be '...a key piece in the puzzle of making these risks more tangible to Internet users'. For this purpose, the authors plan to build a user-friendly website to display the information to users, however they neglect to consider whether people actually care enough about the data contained within this matrix to do anything as a result of it. Moreover, they fail to consider the fact that cultural perceptions of what actually constitutes private information varies greatly. A large study by Wang et al. [4] (discussed in more detail in Section 3) for example found that given a large list of personal attributes (including street address, phone number, employer and email address) Indians only considered personal phone numbers to be sensitive information. Thus, before we attempt to educate people as to what private information they are potentially leaking, we need to first understand what private information *is*.

3 Cultures

It is difficult to define cultures, especially considering their dynamic and inherently varying natures. For the purpose of simplicity, this section (like existing studies) will mostly discuss cultural differences across *countries*, whereas other factors (such as social norms) will be discussed in further detail in Section 6.

3.1 General Patterns Across Countries

In research by Wang et al., Chinese, American and Indian (representing the worlds three largest populations) users of social networking sites were surveyed in attempt to discern differences in usage patterns. The authors utilized crowd sourcing sites to recruit users, providing monetary reward for participation in their surveys. While Wang et al. acknowledge that this method is subject to self-selection bias, they undertook a variety of measures in order to 'clean' the provided data.

The authors found a wide variety of privacy behaviours and attitudes, including (when controlled for other variables) a statistically significant *generalizable* trend amongst the three countries (Figure 2). It was found that generally social networking users from the USA were the most privacy concious, followed by China and then India. Given the frequent statistical significance of the results when considering the country variable, the authors instead decided to focus primarily on privacy factors that *deviated* from this trend.

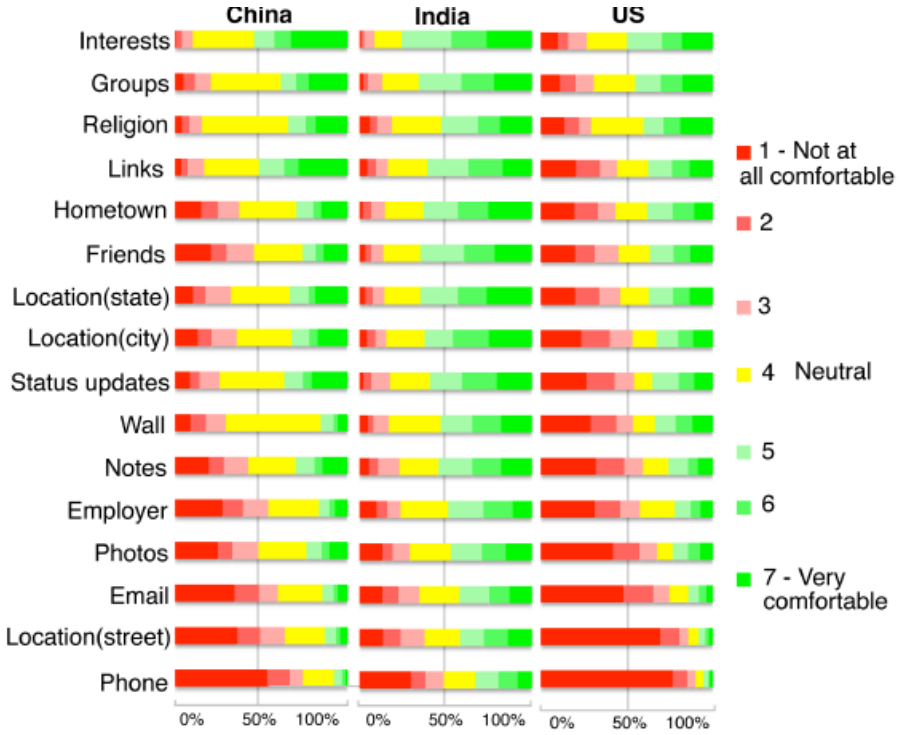

Figure 2: Privacy Attitudes of Users Across Countries [4]

3.2 Other Privacy Differences by Country

An area receiving increasing attention in mainstream media (especially in terms of targeted advertising) is the trustworthiness of the social networking operators themselves. This issue was investigated by Wang et al. demonstrating a strongly statistically significant difference amongst the three cultures. As with the aforementioned general privacy attitudes discussed above, American's exhibited the highest lack of trust in these operators followed by China and then India. Though not explained by the authors, this pattern can likely be attributed simply to the varying levels of attention privacy has been given by the media in these countries.

An important privacy concern on social networking sites is the restriction to the access of personal information to a desired subset of users. Interestingly, it was found that cultural opinions pertaining to this factor deviated greatly from the aforementioned generalizable relationship. Chinese users have by far

the highest desire to restrict the access to their social networking information, followed by Indians and then Americans. Possible cultural explanations for this are provided in Section 3.3.

Another divergent trend in cultural privacy differences exists in the concern of impersonation and resultant usage of fake names. While American's were generally neither concerned about impersonation nor did they utilize fake names, users from other countries felt differently [4]. Chinese users in particular demonstrated this point, with nearly half of all users utilizing fake names on social networks (once again possible explanations for these differences are provided in Section 3.3 below).

3.3 Possible Cultural Explanations

Wang et al. cite earlier work into cultural differences between these countries. This work states that the USA is a more 'individualistic' society and thus places more value on personal privacy. Chinese culture conversely is a much more 'collective' society while India sits somewhere in between the two.

As discussed above, Chinese social network users are however much more wary of who has access to the information they are publishing and also have a much higher likelihood of utilizing fake names. Wang et al. attribute these differences to the existing high levels of government censorship and monitoring of social networks. Additional earlier work identified by Wang et al. discusses how Chinese social networking users tend to utilize these networks not purely as a means of maintaining existing relationships, but instead as a way of creating *new* relationships. This could additionally explain why Chinese users are much more willing to be forthcoming with greater levels of (arguably) personal information on social networks. An alternative/additional explanation for American's general higher levels of privacy *concern* may simply be the aforementioned higher levels of media attention privacy has been given in comparison to non-western developing countries [4].

While the causes of cultural differences in both privacy attitudes and behaviour may never be definitively concluded in their entirety (especially when one takes into consideration their diverse and dynamic nature), what *is* important to emphasize is that the differences are major, and that they *do* exist.

4 The Privacy Paradox

In the following two sections both the 'privacy' and 'control' paradoxes are discussed. These paradoxes both hold the potential to be considered meta-predictors of privacy behaviour on social networks.

Users of social networks often suffer as a result of two opposing forces; on the one hand users do (generally) care about their privacy, but on the other - the two main reasons users utilize social networking sites are to maintain relationships, and *present themselves to others* [3]. The concept of deliberately representing oneself is referred to as *impression management*. By impression management we do not refer to users who are narcissistic in nature (though the effect of narcissism on privacy is discussed in further detail in Section 6.1), but instead simply the desire held by *most* users of social networks (and indeed humans in general) to present themselves well to friends, family, employers and other acquaintances [4, 3]. Thus the term *privacy paradox* is used to capture the concept that while users *say* they care about their privacy, they continue to publish potentially private information [3]. Unsurprisingly, work by Utz and Kramer [3] found that there existed a statistically significant correlation between privacy behaviour and impression management. In other words, impression management is yet another predictor of privacy behaviour on social networks. This finding does not entail the existence of the aforementioned privacy paradox however, thus Section 4.1 below briefly discusses further work by both Utz and Kramer as well as other research into current privacy trends in social networking behaviour in an attempt to discern its validity.

4.1 The Privacy Movement

Whether a result of the increasing levels of media attention or simply improved accessibility, there exists a globally increasing trend in the modification of default privacy settings on social networks [3, 4]. Further study by Utz and Kramer concluded that unlike only a year earlier, the vast majority of social network users (up to roughly 90%) in their study were now modifying their default privacy settings. Utz and Kramer went on to conclude that although it may have been true in the past, their research provides further evidence demonstrating a trend away from the existence of a privacy paradox. This additionally further

verifies the existence of a movement towards a future whereby social networking users not only have more control over their privacy settings, but are simply more privacy concerned and proactive [3, 4].

Though an appealing trend on the surface, an important question that resultantly arises is '*will average users actually benefit from increasing levels of privacy* ***control*** *on social networks?*'

5 The Control Paradox

In their paper *'Misplaced Confidences: Privacy and the Control Paradox'* Brandimarte, Acquisti and Loewenstein [1] provide a metaphorical comparison of the effect that seatbelts have on drivers, to the effect increasing the levels of privacy control have on social networks users. Risk homoeostasis refers to the effect whereby when provided with increased safety measures, users will adjust their behaviour to counteract them (e.g. in the case of the aforementioned example, drivers with seatbelts tend to drive more recklessly).

Similarly, additional related research shows that when people *feel* they are in a position of control, they tend to be more willing to take risks - and they consider these risks less severe [1]. Brandimarte et al. hypothesize that this may be because when provided with the ability to control some elements, users often neglect to consider others.

5.1 The Effect of Control on Privacy Behaviour

Based on the aforementioned prior research, Brandimarte et al. were interested in whether it could be inferred that when a user is given more control (perceived or otherwise), they will reveal more information than they would had they been given lesser control. Additionally the authors posit that when given *lesser* control (once again perceived or otherwise), users will be less willing to reveal personal information. Finally, the authors hypothesize based on the aforementioned risk-neglect characteristic that when provided with increasing levels of control over their privacy settings, users of social networks will neglect to consider other external factors (for example what will their contacts or other third parties do with their information).

Brandimarte et al. found over the course of several studies that the aforementioned hypothesis were all accurate. When provided with higher levels of

control over their privacy settings users often revealed much more information than what would be typically be given, moreover the information that was revealed was often of a much more sensitive nature.

In an interesting additional study performed by Brandimarte et al., users were divided into two groups and asked a series of personal questions. Both groups were told that the information would be used to create a new social network, however one group was told there was only a 50% probability that their information would be posted. This was done in an attempt to represent the level of *perceived* control held by users. Interestingly (albeit as per their hypothesis) Brandimarte et al. found that the group with only a 50% probability of their information being posted were much *less* forthcoming with private information. This is despite the fact that in reality the users in this group had a *lower* likelihood of the information ever being seen.

Despite the increasing privacy movement outlined in Section 4.1 and the widespread call for increasing control to be given to users over what information is made available on social networks; while arguably what 'the people' want, this may in fact not be in the average users best interests [1].

Thus while neither an individual nor cultural characteristic, an additional predictor of privacy behaviour on social networks is the level of control provided to users.

6 Non-Cultural Predictors

The remainder of the body of this report will be divided into a brief discussion of other intuitive and/or major non-cultural predictors of social networking privacy behaviour.

6.1 Narcissism

A individual characteristic discussed over several of the research projects described in this paper was that of *narcissism* [1, 3]. As discussed earlier, one of the prime uses of social networks is in the promotion of self, i.e. impression management. It was therefore unsurprising that in the aforementioned study by Utz and Kramer, it was found that narcissistic users possessed a higher tendency to reveal personal information, and they did so to a higher number of users. Utz and Kramer went on to discuss that many users utilize social networks as a

means of representing their 'ideal' selves. This is an additional proponent that could influence narcissistic users to post higher quantities of information to social networks. Additional evidence for the negative influence of narcissism on privacy is provided in related work by Brandimarte et al., whereby the authors state that a persons inability to consider the negative consequences of others actions pertaining to their private information is directly correlated to their 'egocentrism'.

6.2 Trust

Utz and Kramer performed additional research into the effect of trust on privacy behaviour in social networks. Specifically the authors focused upon what is referred to as 'dispositional trust' (i.e. ones belief that others will act in their best interests). Surprisingly however, it was found that this had no statistically significant effect on user privacy attitudes and/or behaviour. Thus, though counter-intuitive (and with the exception of the *trust in social network operator* discussed in Section 3.2) it was found that trust is *not* a reliable predictor of social networking privacy behaviour.

6.3 Social Norms

The role of perceived social norms has long been confirmed in a variety of behaviours, ranging from the positive to more negative habits such as drinking and smoking [3]. Furthermore, existing research has specifically shown that social networking users are much more likely to set their profiles to private if there friends do so too [3]. Based upon this information Utz and Kramer hypothesized that a similar effect most likely exists within users social networking privacy attitudes and behaviour in general.

Once again, their hypothesis was (albeit partially) proven to be correct, whereby their research demonstrated that social norms were a prime predictor of social networking behaviour. Interestingly however, it was found that there did not exist a statistically significant correlation between social norms and privacy *attitudes*. This indicates that while social norms *are* a valid predictor of social networking behaviour, users nonetheless derive their own interpretations as to what actually constitutes private information.

6.4 Technical Knowledge

Research by Wang et al. into an investigation of the effect of a wide range of personal attributes including (but not limited to) gender, age and education interestingly found that technical knowledge was *the* most statistically predictor of social networking privacy behaviour. Users with higher levels of technical knowledge tend to be much more forthcoming with (arguably) private information on social networks.

Whether the importance of technical knowledge is itself an independent variable, or a result of the related findings pertaining to the existence of a the control paradox by Brandimarte et al. (and discussed in Section 5.1) is an area requiring future investigation.

6.5 Privacy Concern

Though arguably the most important predictor of social networking behaviour, there exists relatively limited emphasis placed on privacy concern. This is arguably because privacy concern is generally discussed from a consequential perspective, however more work is required to simply understand the effect that privacy *concern* has on privacy *behaviour*. As demonstrated by the above characteristics, there exists a variety of factors that would ensure this was not a direct 1:1 relationship.

Nonetheless, the limited research that has been performed into privacy concern has resulted in numerous interesting findings. As mentioned earlier, work by Utz and Kramer effectively demonstrated that in general privacy concerns *do* effectively measure privacy behaviour (and thus indicate a movement away from the aforementioned privacy paradox). Moreover the work by Wang et al. demonstrated that there exists great differences in privacy concerns across cultures, while work by Bradibart et al. exhibited similar differences as a result of a variety of non-cultural factors.

Regardless, when considering the obvious importance of privacy concern to privacy behaviour there is much work that needs to be done. In other words, future work is required to evaluate not what privacy concerns are affected *by*, but instead what is affected *by* privacy concerns.

7 Discussion

Our initial research into what is at risk with the checked and unchecked publishing of information in social networking sites has demonstrated that as the usage of social networks continues to increase (and with it the quantity of publicly available information), so too does the complexity of the methods employed by malicious entities [2]. It is the authors opinion that before 'cookie-cutter' solutions like those outlined in Section 2 are to be employed for the purpose of educating users as to the privacy dangers associated with social network usage, it is important to first understand both on an individual and cultural level what actually constitutes private information.

In order to investigate this we initially discussed research into the cultural differences held by users of social networks. As a result of the difficulties in attempting to capture or even define individual cultures, a study by Wang et al. into the cultural differences of social networking privacy concerns instead decided to simply represent cultures by countries. Nonetheless, their study found that there exists a generalizable trend in the privacy of the three countries investigated; whereby Americans are in general the most privacy concerned, followed by the Chinese and then Indians. The authors attributed these differences primarily to the individualistic society upheld in the USA, as opposed to the more collective society of China. Wang et al. were however unable to provide explanation as to why Indian social network users are the *least* privacy concerned.

Research into the validity of the *privacy paradox* demonstrated the increasing popularity of a 'privacy movement' promoting a future whereby average end users are not only more proactive and concerned with their privacy, but additionally have higher levels of control over their social networking behaviour [2, 3]. However, related work by Brandimarte et al. described the existence of a *control paradox*; demonstrating that giving users of social networks these increasing levels of control may in fact not be in their best interests.

Section 6 above discussed a variety of intuitive (though not *necessarily* accurate) non-cultural predictors of social networking privacy behaviour. Though counter-intuitive (and with the exception of the *trust in social network operators* discussed in Section 3.2) it was found that trust levels (both 'dispositional' or otherwise) had little to no effect on social networking behaviour, while technical

knowledge was cited as being *the* most important predictor of social networking behaviour [4, 3]. Though much work has been done on the effect of various factors (both on an individual and cultural level) into privacy concern, there remains a strong need for the relationship between privacy *concern* and *behaviour* to be directly evaluated.

8 Conclusions and Future Work

While the rapid uptake of social network usage across an increasingly connected world shows no signs of abating, it is important to understand the effect that this increasing diversity (both cultural or otherwise) has on privacy behaviour. The work discussed in this paper outlined a variety of likely predictors of social networking behaviour. However further investigation is required into several key areas.

Firstly, there exists relatively little discussion or targeted research into the privacy concerns and behaviours of specific cultures. While Wang et al. attempted to justify this by stating that '...culture is fluid, dynamic and often difficult to define', it is the authors opinion that from an anthropological standpoint, it is more than plausible to derive (non-geographical) cultures. Secondly, though Wang et al. attributed the differences in privacy attitudes between American and Chinese users to their individualistic and collective societies respectively, work is required to understand why Indian social networking users possess the least concern for their privacy. Additionally, though Wang et al. concluded that technical knowledge is the most statistically significant predictor of social networking privacy behaviour, future work is required to evaluate whether this is a result of the aforementioned effect of the feeling of control on user behaviour; or if it is itself a primarily independent variable. Finally, while much work has been done into the effect of individual and cultural factors on privacy concern; further work is required to directly investigate the effect of privacy concern on privacy *behaviour* [1, 3, 4].

While the causes of the aforementioned differences in social networking privacy behaviour may never be understood in their entirety, what is important is for both the educators of social networking privacy risks and the network operators themselves to realize that these differences *do* exist, and that they *should* be taken into consideration.

References

[1] L. Brandimarte, A. Acquisti, and G. Loewenstein. Misplaced confidences privacy and the control paradox. *Social Psychological and Personality Science*, 4(3):340–347, 2013.

[2] S. Creese, M. Goldsmith, J. R. Nurse, and E. Phillips. A data-reachability model for elucidating privacy and security risks related to the use of online social networks. In *Trust, Security and Privacy in Computing and Communications (TrustCom), 2012 IEEE 11th International Conference on*, pages 1124–1131. IEEE, 2012.

[3] S. Utz and N. Krämer. The privacy paradox on social network sites revisited: The role of individual characteristics and group norms. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 3(2):2, 2009.

[4] Y. Wang, G. Norice, and L. F. Cranor. Who is concerned about what? a study of american, chinese and indian users privacy concerns on social network sites. In *Trust and trustworthy computing*, pages 146–153. Springer, 2011.