Measuring Snooping Behavior with Surveys: It’s *How* You Ask It

Diogo Marques
LaSIGE, University of Lisbon
Ed. C6, Piso 3, Campo Grande
1749-026 Lisbon, Portugal
dmarques@di.fc.ul.pt

Tiago Guerreiro
LaSIGE, University of Lisbon
Ed. C6, Piso 3, Campo Grande
1749-026 Lisbon, Portugal
tjvg@di.fc.ul.pt

Luís Carriço
LaSIGE, University of Lisbon
Ed. C6, Piso 3, Campo Grande
1749-026 Lisbon, Portugal
lmc@di.fc.ul.pt

CHI 2014, Apr 26 - May 01 2014, Toronto, ON, Canada
ACM 978-1-4503-2474-8/14/04.
http://dx.doi.org/10.1145/2559206.2581240

Abstract

In close relationships, snooping on another’s mobile device is commonly regarded as an invasion of privacy. The prevalence of such behavior is, however, difficult to assess. We compared two in-person survey techniques, one in which the question about snooping behavior is posed directly, and one in which strong anonymity controls are employed. Results (n=90) reveal that, while in the first case 10% of respondents admitted to snooping, in the second the estimate was 60%. This shows that, although there is a potent social desirability bias at play, strong anonymity controls do improve estimates. Furthermore, it suggests an alarming prevalence of snooping behavior among the target population.

Author Keywords

Privacy; Mobile Devices; Methodology; Social Desirability Bias

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Introduction

Personal mobile devices are a growing source of concerns relating to privacy. To some degree, people have assimilated the threat of tracking and dragnet surveillance [1]. However, targeted invasion of privacy by socially close adversaries, or insiders, has received less attention. As personal mobile devices store more personal information, this threat becomes of added importance. Recently, efforts to understand the risk of snooping behavior by non-technical adversaries have emerged [8]. One issue in this line of research is that the prevalence of the phenomenon is not easily measurable.

Asking people about snooping behavior runs into a number of issues. If asked directly, one may be reluctant to self-incriminate. If, instead, we ask whether one was a victim of other's snooping through their device, only the times when the victim knows about the breaching will be captured. In any case, the question itself can be construed as an invasion of privacy [10]. In this paper, we show that the prevalence of snooping behavior can be estimated through surveys, but appropriate methods are required. We propose the use of strong anonymity controls, namely using a voting procedure for in-person surveys, coupled with the unmatched count technique (UCT). While voting has for long been a staple of anonymity protection, the UCT has recently also been shown to be appropriate for sensitive questions [4].

To validate this approach, we conducted surveys with two competing procedures: one in which a question about snooping was posed directly, and one were the anonymity controls were used. Results show that the anonymity-preserving methodology leads to a much higher epidiomological estimate (60% vs. 10%), likely due to the relief of the social acceptability bias. Furthermore, they suggest that snooping behavior among the target population is alarmingly prevalent.

Related Work

Users′ security concerns

There is abundant evidence that users worry about security in their mobile devices. Even in the pre-smartphone era, a study of factors influencing the choice of a handset indicated that security considerations ranked second, only after battery life [3]. User concerns over security are preventing them to take full advantage of the technology at their disposal: as many as 70% of users are reluctant to perform many privacy-sensitive tasks on mobile devices [2].

The risk of "forced exposure" ranks high among security concerns. Personal mobile devices store a great deal of information that is considered sensitive, including passwords, files, contacts, emails, text messages, call logs, location traces, schedules, pictures and videos [2,8]. Exposure of this information to insiders can be particularly harmful, given the potential to damage social relationships [6].

A recent survey of crowdworkers found that 9% admitted to snooping through someone else's device. However, 12% indicated knowing that they had been a target of snooping [8]. We build on these findings and hypothesize that strong anonymity controls can produce more encompassing estimates, that do not rely on occurrences that may be unknown to the user.

The unmatched count technique

The unmatched count technique, first proposed by Raghavarao and Federer [9], is a survey instrument specifically designed for sensitive questions. The procedure requires that participants are randomly assigned to two groups, to which two different variants of a survey question are assigned. Individuals in the control group are shown a list of behaviors and asked in how many (not which) they have engaged. The other group has an extra item in the list, which is the sensitive behavior. If there is a difference in the mean number of reported behaviors between groups, it follows that it can be explained by the sensitive item. Thus, the difference in means is a base rate estimate for the behavior.

This technique has been shown to produce better estimates of a wide array of sensitive behaviors, including drug use, sexual practices and racial discrimination [4]. To reduce measurement error due to social desirability bias, the technique should be coupled with additional best practices, accounting, for example, with issues of third party presence, interviewer training, wording, and expected variance in the non-sensitive items [7]. The next section addresses how these best practices were put in place, in conjunction with a voting procedure that enhances anonymity.

Proposed methodology

To obtain a more accurate estimate of snooping behavior, we propose a methodology for face-to-face surveys. Here we presents the details of the methodology, which were refined through a pilot test (n=15, results not reported here).

The item list

We propose posing the question as follows:

Consider the following sentences relating to personal mobile device (smartphone or tablet) usage:

- (1) Since [DATE], I have received on my mobile device at least one unsolicited advertising instant message (SMS or similar).
- (2) Since [DATE], I have purposefully made phone calls or sent text messages to value-added numbers (excluding costumer support numbers and regular international calls/texts).
- (3) Since [DATE], I have used a device from someone I know without their permission to look into personal data (for instance, look through texts or call history).
- (4) Since [DATE], I have shared photographs taken with my mobile device with other people (for instance, with Instagram, Facebook, Twitter, SnapChat, WhatsApp, email, MMS, etc.).
- (5) Since [DATE], I have lost or had a mobile device stolen from me (even if you I have recovered it later).

How many of the previous sentences apply to you?

Item 3, being the sensitive issue, is not to be administered to the control group. All other items have low expected variance, with items 1 and 4 applying to a large majority of users, and items 2 and 5 applying to a large minority [5]. This minimizes the chances of respondents perceiving disclosure of engagement in the sensitive behavior. If, for instance, all but the sensitive item were very frequent, there would be reduced anonymity for those that engaged in the behavior, since the answer would frequently be "5".

The items are prefixed with a temporal frame to allow for cross-sectioning. This is important to this question because it is reasonable to believe that the sensitive behavior is changing throughout time, even within the same population. As personal mobile devices carry more information and there is greater awareness of this fact, snooping through someone else's phone may become of more value to potential intruders. Conversely, it also may lead to wider adoption of security mechanisms. Temporal cross-sectioning enables analysis of these dynamics in longitudinal studies, and further reduces the variance of the non-sensitive items.

The wording of the sensitive item is intended to be descriptive and non-judgmental, and is based on a previous survey [8]. Charged words, such as "snooping", are purposefully avoided. The remaining items also pertain to mobile device security, thus providing a context in which the sensitive issue is not out of place.

Mode of administration

This methodology is specific to in-person surveys. Therefore, measures have to be put in place to reduce nonresponse and misreporting. We propose a portable voting procedure for gathering responses that goes as follow:

- The interviewer shows the respondent a card with the item list on a clipboard and gives him/her a ballot.
- The interviewer instructs the respondent to write in the ballot the number of sentences that apply, never showing it.
- The interviewer shows the respondent a transparent ballot container and instructs the respondent to fold the ballot and place it in the container after responding.
- The interviewer looks away from the respondent and waits for the ballot to be cast.

To increase the feeling of anonymity, the ballot container should be seeded with blank votes.

Interviewers

Interviewers should be from the same group than the participants and trained in survey ethics and techniques (interception, gender balancing, etc.). Particularly, for this questionnaire, interviewers must be trained in intercepting individuals in way that reduces third party observation, since this is known to increase misreporting.

Interviewer guidelines, as well as the remaining materials used in the comparative study presented in the next section are available at http://diogomarques.net/snooping-survey.

Comparative study

To validate the proposed design, we conducted an experiment in which we administered both the proposed procedure (n=60, 30 in each group) and face-to-face direct questioning (n=30). The surveys were conducted in December 2013, and report to that year.

Research question

This study addresses a single research question: does the proposed methodology reveal higher prevalence of snooping behavior than direct questioning?

Methodology

The study employs a between-subjects experimental design. The independent factor is the survey procedure, with the two aforementioned levels. The response

Number of items selected	Frequency in 4-item group	Frequency in 5-item group
0	1	0
1	2	1
2	14	9
3	12	11
4	1	9
5		0
Total:	30	30

Table 1. Distribution of number of items selected by participant in the unmatched count groups. Each cell shows how many people selected the number of items indicated in the row header.

variable is the estimate of base rate among the population (the prevalence).

Apparatus

Materials for the procedures include: cards with the items for the UCT, ballot cards, a ballot container made out of a transparent plastic bag, and a survey records form for the interviewer. Copies of the paper materials are made available online at the aforementioned web address.

Interviewers and participants

Given that the UTC procedure requires a control group, who does not report on the sensitive question, 60 participants were recruited for that group, so that it is comparable to the 30 answering the direct question. Participants were recruited through street interception and weren't offered any compensation. Prior to administering the survey, they were screened for the following inclusion criteria: being between 15 and 34 years old, students, and regular smartphone or tablet users.

Three interviewers were recruited among M.Sc. students of our host institution. They were trained during a half-day session before the surveys were put in the field. Each one administered the same number of procedures to each experimental group, with gender balancing within each. The surveys were conducted in a nearby University, albeit not the same the interviewers attend (to avoid biases arising from participants knowing the interviewers).

Procedure

Upon interception, interviewers present themselves and asked for volunteer participation in a survey about security issues with mobile devices. The procedure for the strong anonymity surveys is as described in the former section. For the direct questioning survey, the following yes/no question was posed: "Since Jan. 1 2013, have you used a smartphone/tablet from someone you know without their knowledge, to look into personal data (for instance, look through texts or call history)?" Interviewers recorded the answer on a paper form.

Results

For the single dependent variable, results are as follows:

- Direct questioning procedure: 3 in 30 participants responded affirmatively, yielding a 10% estimate of prevalence of snooping behavior.
- Strong anonymity procedure: in the control group, the mean number of items selected was 2.33 (SD=.80), whereas in the other group it was 2.93 (SD=.87). The estimate is the difference in means, 60%. Table 1 shows the distribution of responses for both groups.

A two-proportion test indicates that the two estimates are significantly different at the 99% confidence level (Z=4.512). The estimate of 10% obtained with direct questioning approximates the one previously found in a study of crowdworkers [8]. Those results, however, are not comparable, since demographics and sample size are different, and temporal cross-sectioning isn't used.

Conclusion

The results show that, for face-to-face surveys, the proposed technique provides an estimate of prevalence that is significantly larger than direct questioning. This finding suggests a path for understanding the prevalence of security behaviors, even when they are sensitive. Furthermore, the estimate of 60% among the respondents is certainly a cause for concern. It suggests that the security mechanisms that are currently available are ineffective. They are either not being used, perhaps because they are inconvenient, or are not very secure against adversaries in close relationships (or both).

Limitations

Although the proposed technique reduces the effect of the social desirability bias, it cannot be said to be eliminated. Given the small sample size and cross-section of population it represents, the epidemiological finding is not generalizable. Nevertheless, analyzing a young and educated population can be of special interest for evolving technologies.

Future work

Having an encompassing estimate of this behavior requires observational studies at a larger scale. As face-to-face surveys are costly, we are engaged in developing a methodology for online surveys. Particularly, we are interested in ruling out that the high prevalence rate is a result of the materials used (e.g. the wording, the selection of non-sensitive items, the order in which they appear, interactions between them). We are also exploring ways in which to ask about the motivations and consequences of snooping.

Acknowledgements

We thank our volunteer interviewers, Rita Henriques, Joao Martins and Oliver Schnell. This work was supported by FCT (SFRH/BD/98527/2013, PTDC/EIA-EIA/117058/2010, PEst-OE/EEI/UI0408/2014).

References

[1] Acquisti, A. Nudging Privacy: The Behavioral Economics of Personal Information. IEEE Security & Privacy Magazine 7, 6 (2009), 82–85.

[2] Ben-Asher, N. et al. On the need for different security methods on mobile phones. In Proc. MobileHCI '11, ACM Press (2011), 465-473.

[3] Clarke, N. L., Furnell, S. M. Authentication of users on mobile telephones – A survey of attitudes and practices. Computers & Security 24, 7 (2005), 519–527.

[4] Coutts, E., Jann, B. Sensitive Questions in Online Surveys: Experimental Results for the Randomized Response Technique (RRT) and the Unmatched Count Technique (UCT). Sociological Methods & Research 40, 1 (2011), 169–93.

[5] Felt, A. P., Egelman, S., Wagner, D. I’ve got 99 problems, but vibration ain't one: A Survey of Smartphone Users' Concerns. In Proc. SPSM ’12, ACM Press (2012), 33-44.

[6] Johnson, M., Egelman, S., Bellovin, S. M. Facebook and privacy: it’s complicated. In Proc. SOUPS ’12, ACM Press (2012), 1-15.

[7] McNeeley, S. Sensitive Issues in Surveys: Reducing Refusals While Increasing Reliability and Quality of Responses to Sensitive Survey Items. Handbook of Survey Methodology for the Social Sciences, Springer (2012), 377–396.

[8] Muslukhov, I. et al. Know your enemy: the risk of unauthorized access in smartphones by insiders. In Proc. MobileHCI ’13, ACM Press (2013), 271-280.

[9] Raghavarao, D., Federer, W. T. Block Total Response as an Alternative to the Randomized Response Method in Surveys. J. Royal Statistical Society Ser. B 41, 1 (1979), 40–45.

[10] Tourangeau, R., Yan, T. Sensitive questions in surveys. Psychological bulletin 133, 5 (2007), 859–883