


Who Masks? Correlates of Individual Location-Masking Behavior in an Online Survey

Dara E. Seidl

Department of Geography, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182, USA


dseidl@sdsu.edu

 <https://orcid.org/0000-0001-8737-7115>

Piotr Jankowski

Department of Geography, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182, USA and Institute of Geocology and Geoinformation, Adam Mickiewicz University, Poznań, Poland

pjankows@sdsu.edu

 <https://orcid.org/0000-0002-6303-6217>

Abstract

Geomasking traditionally refers to a set of techniques employed by a data steward to protect the privacy of data subjects by altering geographic coordinates. Data subjects themselves may make efforts to obfuscate their location data and protect their geoprivacy. Among these individual-level strategies are providing incorrect address data, limiting the precision of address data, or map-based location masking. This study examines the prevalence of these three location-masking behaviors in an online survey of California residents, finding that such behavior takes place across social groups. There are no significant differences across income level, education, ethnicity, sex, and urban locations. Instead, the primary differences are linked to intervening variables of knowledge and attitudes about location privacy.

2012 ACM Subject Classification Security and privacy → Human and societal aspects of security and privacy

Keywords and phrases privacy, geoprivacy, geomasking, obfuscation, accuracy

Digital Object Identifier 10.4230/LIPICs.GIScience.2018.57

Category Short Paper

Funding This material is based upon work supported by the National Science Foundation under Grant No. 1657610. This work was supported in part by a American Association of Geographers (AAG) dissertation grant and an American Geographical Society (AGS) fellowship.

1 Introduction

While a large body of research is dedicated to protecting the privacy of human subjects, there has been less documentation on the efforts of individuals to protect their own privacy. The set of procedures known as geomasking typically refers to the alteration of point data to protect both spatial distributions and privacy of data subjects [2]. Common geomasking techniques include random perturbation [6], donut masking [4], and grid masking [12]. The typical use scenario for these top-down strategies is for researchers who wish to share geospatial data with others, but must protect privacy. Masking behavior at an individual level, such as by responding to location requests with false or imprecise address data, can also serve to protect an individual's geoprivacy. This study tests the correlates of bottom-up or individual-level



© Dara E. Seidl and Piotr Jankowski;
licensed under Creative Commons License CC-BY

10th International Conference on Geographic Information Science (GIScience 2018).

Editors: Stephan Winter, Amy Griffin, and Monika Sester; Article No. 57; pp. 57:1–57:6



Leibniz International Proceedings in Informatics

LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

location masking in an online survey, finding that intervening variables of hacking exposure, social media use, and geoprivacy knowledge and attitudes are significantly correlated with masking behavior.

1.1 Related work

In their geoprivacy manifesto, [5] argue that location privacy stands apart from information privacy, in part because of the range of inferences that can be drawn from location, the ubiquity of location-collecting technology, and the incentives which draw consumers to share their locations. Compared to information privacy, which has been reported on by [8] and [1], not much is known about specific geoprivacy-related attitudes and behaviors. Obfuscation at the individual level is characterized as an act of resistance to surveillance [3], an idea seconded by [13] and [14], who argue that Tor, an onion routing technology that masks location by altering IP address, is a prime example of resistance to geosurveillance. Compared to the technologically-advanced location masking of Tor, this study focuses on the masking behavior internet users exhibit when faced with the explicit location request: “Please enter your home location.” Both the precision and participant-reported truthfulness of entered location are collected as outcome variables in determining “who masks”.

2 Methods

An online survey testing location masking behavior was deployed to California residents between October 2017 and March 2018. Participants were drawn from two samples: a random address-based sample obtained from Survey Sampling International (SSI) and contacted by postal mail, and a non-probability online open sample, reached by paid ad placement on Facebook and free advertising on Craigslist. A primary concern in the survey design was to avoid social desirability bias, which results in inflated privacy concerns by participants in studies advertised as privacy-related [11]. Therefore, this survey was designed to omit use of the word “privacy” and to capture location masking as it might occur in a routine online setting. Participants were told they were participating in a study about “online information sharing” and were debriefed about the true purpose of the study at its conclusion, at which time, they were also given the option to withdraw their responses.

2.1 Conceptualization

This study follows a knowledge-attitudes-behavior framework to predict participant location masking, a model commonly used to predict behavioral outcomes in health and environmental studies [9][7]. Hypothesized background variables included age, education, sex, income, ethnicity, and urban location. Given that previous negative privacy experience online increases perceived risk of sharing on social media [15], hypothesized intervening variables included recent identity theft or hacking, social media use, and employment experience with personal data. It was hypothesized that location masking behavior would be most closely correlated with high geoprivacy knowledge and concern for geoprivacy. Each of these variables was measured in a series of Likert-type questions in the survey.

2.2 Survey design

The primary test of location masking was participants’ response to “please enter your home location,” for which they were given text boxes for street, cross street, city, state, and zip code. If respondents entered a text-based location, they would then have the option to open

■ **Table 1** Differences between mail and online sample in Mann-Whitney U tests for background variables (* $p < 0.05$).

Variable	Mail Sample	Online Sample	Sig
Female	55%	76%	*
White	66%	55%	
College degree	69%	44%	*
Median age group	45-54	25-34	*
Median income tax bracket	25% (38,000–92,000)	15% (9,000–38,000)	*
Somewhat or very urban	62%	56%	
Total participants	113	101	

up a map and adjust a pin to their chosen coordinates. By default, the map pin was placed at the geocoded coordinates of the entered street address with the Google geocoding API. Respondents then selected their level of agreement on a five-point Likert scale (strongly disagree to strongly agree) to the statements, “I intentionally provided incorrect information on my home location” and “I intentionally moved the pin on the map away from my home location.” The remainder of the survey tested geoprivacy knowledge, attitudes, and the other background variables with similar Likert-type items, asking participants to respond with their level of agreement. The survey was hosted on the Qualtrics platform and fully encrypted.

2.3 Analysis

Differences between the two samples were analyzed with Mann-Whitney U tests, a non-parametric test for differences between two categorical variables [11]. Due to the ordinal nature of the majority of the study variables, Spearman’s correlations were calculated between each of the variables and tested for significance [10]. To determine geographical patterns, global and local Moran’s I were applied as tests of spatial autocorrelation for survey participation rates, location masking behavior, and geoprivacy-related attitudes.

3 Results

The questionnaire had a total of 214 respondents with 113 in the mail sample and 101 in the online open sample. The two samples differed significantly in age, income, education level, and gender composition, based on Mann-Whitney U tests (Table 1). The online open sample was more female, younger, and had lower education levels and incomes compared to the mail-based sample. The mail sample self-reported on average as more urban, though this did not reach significance. The mail sample was also significantly more likely to have employment experience working with personal data. In terms of location masking, the online sample was significantly less likely to provide a numbered street address for home location ($p < 0.05$), compared to the mail sample, although the majority of participants in both cases provided home location at this highest precision (73% of mail sample respondents and 56% of open sample respondents). When it came to factuality of reported home location, however, there were no significant differences between the two samples (Figure 1). About 15% of respondents somewhat or strongly agreed that they intentionally provided an incorrect home address, and 11% of respondents who interacted with the map function agreed that they intentionally moved the pin away from their home location.

When tested with global Moran’s I, there was no global clustering of the respondents from the two samples at the county level when normalized by population. This suggests that a randomly distributed sample was achieved in both cases. Location masking behavior was

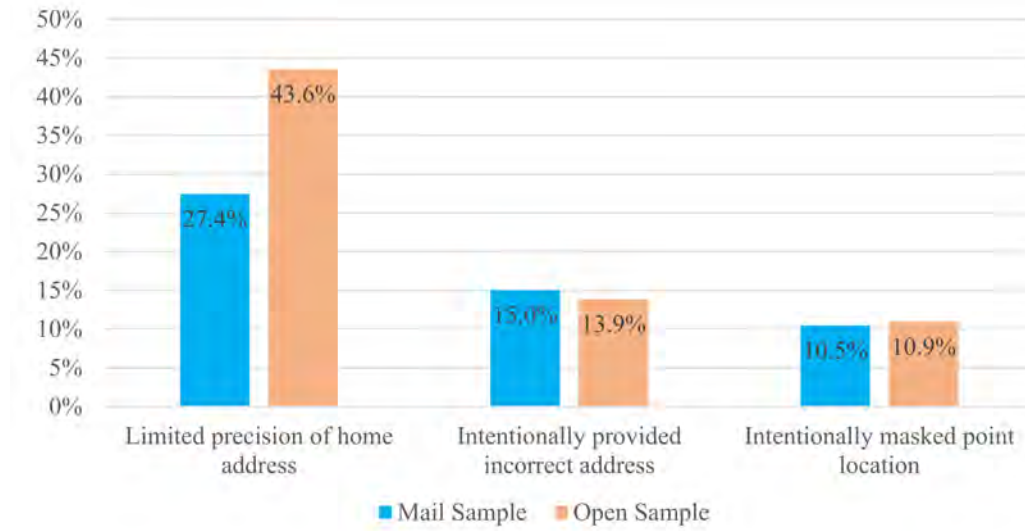


Figure 1 Results by sample for three location masking behaviors.

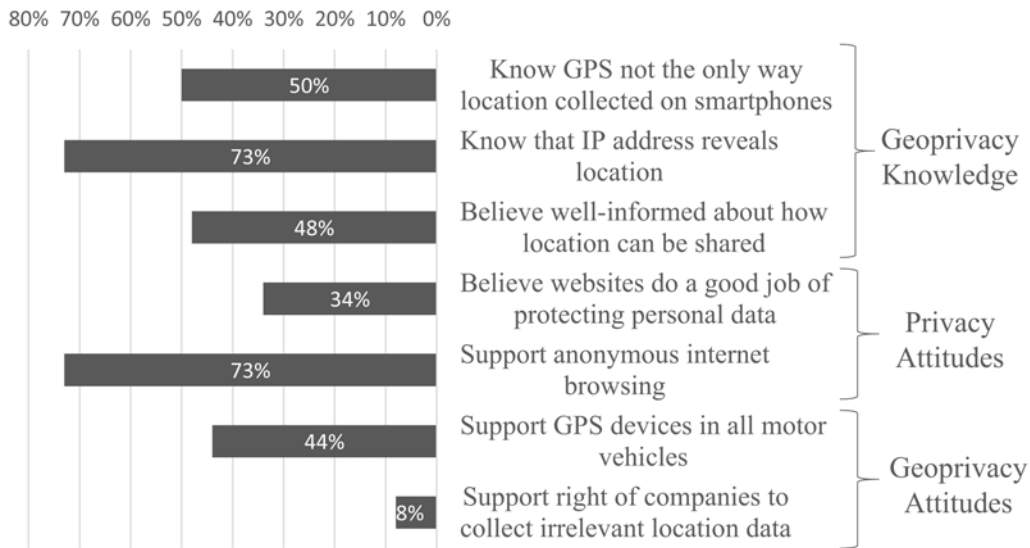
not globally clustered when tested with Moran's I, however, two of the attitude variables, trust in websites to protect personal data and support of GPS devices in all vehicles, were globally and locally clustered ($p < 0.05$).

Overall knowledge about location privacy was low to average, with just 50% aware that smartphones collect location outside of GPS, and 73% aware that IP address reveals location (Figure 2). Self-reported knowledge about how location is shared was also low, with 48% believing themselves to be well-informed. The attitude results demonstrated overall concern for privacy, with only 34% believing websites to do a good job of protecting personal data, and just 8% supporting the right of companies to collect irrelevant location data.

The Spearman's tests (Table 2) revealed that no demographic background variables were significantly correlated with the three indicators of location masking. Location precision had the highest frequency of significant correlates. Respondents were more likely to mask location by providing lower address precision if they were part of the open sample, if they had a recent hacking experience, if they had more knowledge about smartphone location collection, and if they did not trust websites to protect their personal data. Lower precision was also correlated with other masking behavior, including use of technology to alter IP address and provision of incorrect address information to retailers. The two intentional masking outcome variables were not correlated with knowledge or attitudes, but again with other location masking behaviors. Enjoyment of social media was the one intervening variable significantly correlated with providing accurate home location.

4 Conclusion

With 15% of participants admitting to providing incorrect address information, location masking behavior is a small but present minority among participants, and it takes place across demographic lines. The precision of location respondents provide appears to be dependent on context, trust, and knowledge, rather than background variables. The open online sample, respondents who do not trust websites to protect their personal data, and respondents who know that location can be collected in smartphones outside of GPS were more likely to



■ **Figure 2** Percent of participants exhibiting geoprivacy-related knowledge and attitudes.

■ **Table 2** Spearman’s rho between predictor variables and location masking behavior. Only significant correlations shown ($p < 0.05$).

	Correlates	Provided higher home location precision	Intentionally provided incorrect home location	Intentionally moved pin away from home location
Background	Sample (1=Mail Sample, 2=Open Sample)	-0.169		
Intervening	Enjoy contributing to social media		-0.227	
	Had unauthorized user on online account	-0.155		
Knowledge	Believe GPS only way location collected on smartphone	0.139		
Attitudes	Believe websites do a good job of protecting personal data	0.238		
Other masking behavior	Use technology to alter IP address	-0.169		
	Give inaccurate or misleading address information to retailers	-0.194	0.227	
	Turn location services off on smartphone			0.194
	Intentionally provided incorrect home location			0.406

provide a lower precision of home address. Location masking measured as truthfulness of location has fewer clear correlations with the hypothesized background variables than location precision does, but is significantly correlated with other location masking behaviors and lower enthusiasm for social media. The results demonstrate that in California, a U.S. state with a large high-tech sector, there is still relatively limited exercise of geoprivacy protection measures at an individual level.

References

- 1 Alessandro Acquisti and Jens Grossklags. Privacy attitudes and privacy behavior. In *Economics of information security*, pages 165–178. Springer, 2004.
- 2 Marc P Armstrong, Gerard Rushton, Dale L Zimmerman, et al. Geographically masking health data to preserve confidentiality. *Statistics in medicine*, 18(5):497–525, 1999.
- 3 Finn Brunton and Helen Nissenbaum. *Obfuscation: A user's guide for privacy and protest*. Mit Press, 2015.
- 4 Kristen H Hampton, Molly K Fitch, William B Allshouse, Irene A Doherty, Dionne C Gesink, Peter A Leone, Marc L Serre, and William C Miller. Mapping health data: improved privacy protection with donut method geomasking. *American journal of epidemiology*, 172(9):1062–1069, 2010.
- 5 Carsten Keßler and Grant McKenzie. A geoprivacy manifesto. *Transactions in GIS*, 22(1):3–19, 2018.
- 6 Mei-Po Kwan, Irene Casas, and Ben Schmitz. Protection of geoprivacy and accuracy of spatial information: how effective are geographical masks? *Cartographica: The International Journal for Geographic Information and Geovisualization*, 39(2):15–28, 2004.
- 7 Debra Siegel Levine and Michael J Strube. Environmental attitudes, knowledge, intentions and behaviors among college students. *The Journal of social psychology*, 152(3):308–326, 2012.
- 8 Mary Madden and Lee Rainie. *Americans' attitudes about privacy, security and surveillance*. Pew Research Center, 2015.
- 9 Susan Morgan and Jenny Miller. Communicating about gifts of life: The effect of knowledge, attitudes, and altruism on behavior and behavioral intentions regarding organ donation. *Journal of Applied Communication Research*, 30(2):163–178, 2002.
- 10 Susan A Nolan and Thomas Heinzen. *Essentials of statistics for the behavioral sciences*. Macmillan, 2010.
- 11 Erin Ruel, William Edward Wagner III, and Brian Joseph Gillespie. *The practice of survey research*. Sage, 2015.
- 12 Dara E Seidl, Gernot Paulus, Piotr Jankowski, and Melanie Regenfelder. Spatial obfuscation methods for privacy protection of household-level data. *Applied Geography*, 63:253–263, 2015.
- 13 David Swanlund and Nadine Schuurman. Mechanism matters: Data production for geosurveillance. *Annals of the American Association of Geographers*, 106(5):1063–1078, 2016.
- 14 David Swanlund and Nadine Schuurman. Resisting geosurveillance: A survey of tactics and strategies for spatial privacy. *Progress in Human Geography*, 2018.
- 15 Hongwei Yang and Hui Liu. Prior negative experience of online disclosure, privacy concerns, and regulatory support in chinese social media. *Chinese Journal of Communication*, 7(1):40–59, 2014.