Social Palimpsests - clouding the lens of the personal panopticon

Dave MURRAY-RUST [a], Max VAN KLEEK [b] Laura DRAGAN [b]
[a] *Centre for Intelligent Systems and Applications, Department of Informatics, University of Edinburgh*
[b] *Southampton*

Abstract. The use of personal data has incredible potential to benefit both society and individuals, through increased understanding of behaviour, communication and support for emerging forms of socialisation and connectedness. However, there are risks associated with disclosing personal information, and present systems show a systematic asymmetry between the subjects of the data and those who control and manage the way that data is propagated and used. In this chapter, we explore a set of techniques for ameliorating the tension between the desire for the benefits of sharing and a distrust of those with whom we share our data.

Keywords.

Introduction

0.1. The rise of personal data and services reliant on it

0.2. Data, Capta, Fiat Data, Exhaust Data - and the impossibility of limiting data diffusion

"Refusal is not a practical option, as data collection is an inherent condition of many essential societal transactions" [2]

0.3. The case for lying and the importance of anonymity

0.4. There's lots of great stuff out there!

Types of disinformation [1]:

redaction is hiding some or all of the information in a message
airbrushing is changing some of the information. *local crowd blending* means change it to a nearby message likely to be plausible. *global crowd blending* means change it to a message in a dense part of the space.
curveball add extra distracting information, push message into low density space

0.5. Why doing it socially is difficult

0.6. Scenario

Lets take the scenario of modifying location data; might want to:

- Hide the fact we're away from home so we don't get robbed - create fictitious data that looks like it came from home town. Enlist friends.
- Avoid paparazzi - generate chaff so real location is obscured
- Disguise the fact we're going to ballet lessons - systematically lie, saying we're going to the boxing gym instead

Alexander's taxonomy [1] discusses several types of disinformation which relate to modifying single messages. In contrast, due to the pervasiveness of modern communications, we are concerned with modifying message *streams*, where a trace of multiple values must be considered. Additionally, there is the possibility of interaction with others, whether it is collusion to strengthen obfuscatory practices, or the addition—purposeful or otherwise—of information which exposes the obfuscation.

It is problematic to consider the obfuscatory tactics here without a sense of the scenario in which they are to be deployed. Our scenario in this paper is:

> The user wishes to make use of services which expect location information; the location information provided is shared publicly and is almost certainly stored indefinitely. At times, the user may want to draw on location based information—such as restaurant recommendations or directions—and there may be times when they with to verify that they were at a particular location.

In order to illustrate the effects of these different strategies, we plot a one dimensional representation of "location" against time.

1. Tools

1.1. Personal Data Stores - allies on the intimacy battleground

1.2. Verification and provenance

References

[1] J. M. Alexander and J. M. Smith. Disinformation : A Taxonomy. Technical report, University of Pennsylvania Department of Computer & Information Science, 2010.

[2] F. Brunton and H. Nissenbaum. Vernacular resistance to data collection and analysis: A political theory of obfuscation. *First Monday*, 16(5):1–16, 2011.