Social Palimpsests - clouding the lens of the personal panopticon

Citation for published version:
Murray-Rust, D, van Kleek, M, Dragan, L & Shadbolt, N 2014, Social Palimpsests - clouding the lens of the personal panopticon. in K O'Hara, C Nguyen & P Haynes (eds), Digital Enlightenment Yearbook 2014 : Social Networks and Social Machines, Surveillance and Empowerment. IOS Press, pp. 75 - 97. DOI: 10.3233/978-1-61499-450-3-75

Digital Object Identifier (DOI):
10.3233/978-1-61499-450-3-75

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Digital Enlightenment Yearbook 2014

Digital Enlightenment Yearbook 2014
K. O'Hara et al. (Eds.)
IOS Press, 2014

doi: 10.3233/978-1-61499-450-3-75

Social Palimpsests – Clouding the Lens of the Personal Panopticon

Dave MURRAY-RUST [a,1], Max VAN KLEEK [b], Laura DRAGAN [b] and Nigel SHADBOLT [b]

[a] *Centre for Intelligent Systems and Applications, Department of Informatics, University of Edinburgh*

[b] *Web and Internet Science Research Group, Electronics and Computer Science, University of 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. This leads to a tension between a desire to engage with online society and enjoy its benefits on one hand, and a distrust of those with whom the data is shared on the other. In this chapter, we explore a set of obfuscation techniques which may help to redress the balance of power when sharing personal data, and return agency and choice to users of online services.

Keywords. obfuscation, data politics, personal data stores, social sachines

Introduction

In 1890, Warren and Brandeis called for consideration of new laws to better secure the 'the right to be let alone', as 'the private lives, habits, acts and relations of individuals' were perceived to be under threat by the emerging technologies of their time – including 'instant' photography, and early electro-mechanical recording devices [47]. Over a century and a quarter later, technology still continues to threaten individuals' right to be left alone, now at unprecedented scale and fidelity. In terms of technological sophistication, the massive computational capacity hosted in various data warehouses is being used to track individuals' daily activities, communications and locations. Such profiling comprises a race to build the most complete picture(s) of each person's life, so as to be able to apply this in-depth knowledge to precisely target them with, advertising and other behavioural manipulations – often in ways too subtle to be perceived (e.g. [28]).

Such behavioural manipulation, driven by omnipresent observation, brings to mind Bentham's thought experiment of the Panopticon, developed by Foucault, describing a prison in which inmates' lives are constantly surveilled as a means of discipline and

[1] Dave Murray-Rust, Centre for Intelligent Systems and Applications, Department of Informatics, University of Edinburgh.

exertion of control. In the world of technologically-driven data collection which we inhabit, this has effects both on people's behaviour, as they internalise the fact that they are surveilled, and on the way in which they are treated, and socially sorted by the gears of our algorithmic society [42].

Meanwhile, the rise of tracking has not been met with a corresponding increase in either public awareness of how and when observation is taking place, nor practical ways with which individuals can protect themselves from unwanted profiling and targeting. As a result, a great asymmetry has arisen between the information brokers with vast resources that track, and those end-user citizens who are being tracked with little or no choice. As if this asymmetry was not problematic enough, the few privacy-enhancing tools that have proven effective against surveillance have often been portrayed by mainstream news and media outlets as tools for terrorst and paedophiles further discouraging their use and adoption in the mainstream (e.g. [1,27]).

In this chapter, we are interested in practices which can cloud the lenses of the observers, to return to people the choice of how they are seen, and to help regain some control of the manner in which our lives are surveilled, both socially and otherwise. We discuss and analyse methods of privacy protection that advance beyond the current state of anonymisation tools which obscure the tracks of individuals, towards those that employ methods borrowed from information warfare [19,33], in order to allow individuals to regain autonomy from unsolicited tracking and behavioural control. We first discuss the rise of personal data economy, followed by a survey of lying and falsification in context-aware systems, current anonymisation and privacy tools. This is followed by an overview of strategies for obfuscation, some of which are currently implemented in either mainstream tools or proof of concept studies, and some of which are speculative, future possibilities.

1. Background

1.1. The Rise of Personal Data and Services Reliant on it

As we pass through the digitally-augmented world that we collectively inhabit, the set of actions with the potential to produce data grows year on year. Portions of this outpouring are kept and stored as *capta*, from *capere*: to keep [16]. From using an access card to unlock a door, right down to tracking individual footfalls, pervasive digital systems illuminate and annotate our physical activity. Accreting around this body of physical observation is an expanding sphere of mental observation and analysis. This can take the form of active practises around recording mental states, such as journalling, but it can also include computational inference, where frequency of posting on social networks becomes an adjunct metric for connectedness, and search terms indicators of intent. As such, the modes of collecting this information can range from explicit, user initiated submission of data, through consensual background recording, to invisible, asymmetric electronic surveillance.

Increasingly, in order to utilise services, we must provide our data to third parties. This ranges from mobile phone numbers being required for Yahoo accounts, to location data being shared with Foursquare or Grindr, to the NHS adding personal health information to centralised databases. The pervasiveness of computationally mediated action and

interaction in modern life means that for much of this data 'refusal is not a practical option, as data collection is an inherent condition of many essential societal transactions' [10].

This leads us to introduce the term *fiat data*—when an organisation uses its position to demand (by fiat) the disclosure of certain information in return for use of its services. Much as being a citizen of the United Kingdom requires paying taxes in pounds Sterling, if one wants to interact socially on Facebook, one must pay the personal data tax which they demand.

In some cases, the provision of personal data is a necessary requirement for the provision of a service, but in others it represents an attempt by the organisation to create a monetisable product from its users. Many pieces of data fit both goals – for example the addition of mobile phone numbers to user accounts may have security benefits for the users, but it also adds stronger, more persistent ties to offline identities.

When data is collected, there is a spectrum of approaches from outright demands, to asking or encouraging users to furnish their data. Increasingly, *gamification* is used to manipulate users into self-surveillance, by providing rewards – whether within the system or through the promise of self-improvement – for activities which require the sharing of data to function:

> Literally, within an hour of waking up, I am playing at least two games that promise to help me become a more productive worker and prolific writer. . . . I want to suggest two things: 1) that gamification is a form of surveillance; and 2) this surveillance is pleasurable [48].

Fitness apps, activity monitors, location based social networks require the user to hand over their location data in return for the promise of increased fitness, self awareness to the ability to connect with others. This user-driven data collection becomes a form of *participatory surveillance*:

> Online social networking can also be empowering for the user, as the monitoring and registration facilitates new ways of constructing identity, meeting friends and colleagues as well as socializing with strangers. This changes the role of the user from passive to active, since surveillance in this context offers opportunities to take action, seek information and communicate [2].

There have been calls to create privacy enhancing tools which minimise the amount of data that is collected to only that which is necessary. Data minimisation, combined with user control over the data and preservation of the context where the personal data was originally disclosed, make up the *Privacy 3.0* view of Borcea-Pfitzmann et al. [7]. As the ability of technology to record our lives evolves and improves, there is an increasingly acute need that the legislative side must catch up and provide the framework for the recognition of personal information ownership rights [41].

There is a divergence between the goals of privacy conscious users and organisations providing services based on personal information. Spiekermann and Novotny have developed a tiered model for privacy-aware information markets to elucidate the tension between these different actors [43]. The model divides the market into four spaces: customer relationships, organisational control, customer control, and safe harbours for big data. Each of these spaces has its own rules and responsibilities for those operating within it, and recommendations from the authors on what technical and legislative infrastructure is needed for their realisation. Some of the most important issues and solutions they

present include the separation of the service exchange from the information exchange, the mandatory privacy-friendly service option that providers must offer, and the lawful accountability of organisations for the personal information that they collect, including any subcontractors they might pass the information to.

In summary, while there are many situations where we are coerced, cajoled or manipulated into sharing our personal data, the cost of avoiding such sharing is increasingly becoming untenable for large sections of the population of the connected world. However, there are steps being taken in the direction of a privacy-aware market for personal information which protect the ownership rights of users over their personal data, and their right to privacy, while also allowing for the market to innovate and grow based on legally obtained and used, and of higher quality personal information.

1.2. Issues with Data Sharing, Control and the Future

Sharing data, by definition, is the entrusting of other parties with information; this necessarily involves relinquishing control over how it is subsequently handled and disseminated. While it is possible to attach terms and conditions to the data at the point of sharing, the sharer must rely either on the technological means or the compliance of the recipients. Technological means to control information usage – such as Digital Rights[2] Management (DRM) – are generally quickly circumvented [15, p. 6]. Research is being carried out into improving *accountability* (e.g. [18]), which can help to create a context which limits breaches of trust. However, this again relies on mechanisms outside the control of the sharer, who must rely on legal means and the evolution of social norms to ensure their desires are met. Essentially, once data is shared the sharer has no control over what happens to it. Once shared, the data becomes persistent. Bruce Schneier commented[3] that we are in the middle of a social change, where the ephemerality of action is being lost:

> Google doesn't only know what I think better than my wife does, it knows what I think better than *I* do, because it remembers all I said and did.

We are used to operating in a world where what we say will be forgotten sooner or later, where actions remain embedded in the context in which they occurred. The increasing persistence of record challenges this, as all utterances and behaviour becomes permanently availble for later scrutiny.

There are many issues with sharing data; we highlight four of them here

Sharing is persistent, while situations evolve; once data is shared, there is no technical means to revoke it. However, the context around its sharing and the organisations involved are subject to change. For instance:

- a government or other organisation may decide to share previously confidential data, as in the case of the recent *care.data* fiasco in the UK
- a company may be bought and its assets acquired – the purchase of Moves by Facebook and the potential for subsumption of location tracking into social network data raised issues around the terms and conditions of data handling companies

[2] Often called Digital *Restrictions* management by opponents, e.g. www.defectivebydesign.org

[3] As part of a panel at the "Don't Spy on Us" day in London, UK, on 7 June 2014.

- unseen events such as security breaches can expose vast swathes of personal data, or court proceedings may force private communications to become public – the Enron emails still represent the largest publicly available corpus of private emails.

Technology improves: what is safe to share now may not be in the future. Brad Templeton from the Electronic Freedom Foundation uses the analogy of 'Time travelling robots from the future': the information collected now will be subjected to increasingly sophisticated analysis techniques as time progresses, so the implications of sharing that information can be far beyond expectations. For example, in the future, it may be possible to carry out facial recognition on massive quantities of CCTV footage, and reconstruct the movements of a large number of citizens. This corresponds to the surveillance robots coming back in time and monitoring us now. In a similar vein: 'Would you have liked to be gay 40 years ago in a monitored society? Or an enemy of J. Edgar Hoover with modern tools in his hands?' [44]. Sharing data today cedes control to the entities of tomorrow, with their greatly enhanced capabilities.

De-identification is not a magic bullet: data is often shared subject to the condition that it will only be shared in an *anonymised* or *de-identified* form. Whilst sharing de-identified data aids in privacy preservation, it can also create a false sense of security, since it is generally impossible to rule out the chance of *re-identification*. This might be through the set of data released, or through the fusion of multiple data sources to provide higher discriminatory power. As a highly public example: Netflix challenged the public to create a better recommendation engine, based on a corpus of anonymised viewing histories. Subsequently, it was shown that many records within the database could be identified by comparison with publicly available sources [37], let alone access to other, non-public data. Narayanan and Felten's recent report [36] explains in a non-technical manner why de-identification of data remains problematic. Some data is particularly resistant to de-identification: for example, location data, where four data points are enough for re-identification in many cases [13].

Databases can be joined: as more databases of personal information become available, whether publicly or privately, the possibility to match, join, correlate, and share data increases, and the effects of single points expand well beyond the environment in which they were created or shared. In short, data are held in *leaky containers*: 'data move freely between different sectors of society with the result that information from discrete contexts, e.g. private life, work life and shopping, are being mixed rather than contained separately.' [32, p. 37–44]. This leads to unexpected and unforseeable consequences when sharing seemingly innocuous data points: we are not in a position to know the outcomes of our actions.

1.3. The Beneficial Uses of Ambiguity and Lies in Social Mediation

Studies in Computer Mediated Communication (CMC) and agent-based modelling of online communities have studied the use of deception both in real and simulated settings, revealing a contrast between *pro-social* forms of lying and deception, and *anti-social* forms [25]. A wide range of pro-social kinds of deception have been documented, primarily around coping strategies that people adopt towards dealing with the demands and complexities of social mediation in increasingly digitally-connected lives.

One strand of work pertaining to such coping strategies, is the exploration of how people leverage the 'space of ambiguity' created by 'low-bandwidth' computer-mediated communications channels, such as SMS (text messaging) or Instant Messaging, or even voice mobile phone conversations. Here, ambiguity about the channel is used to facilitate the use of pro-social deception for the purpose of efficient mediation. For instance, an individual might intentionally ignore an incoming call, or blame poor reception (where there is none) to terminate a call prematurely in a situation where they are with other priorities, or wish to avoid talking to someone for another reason. In work by Aoki and Woodruff [4], it is argued that such uses of ambiguity are beneficial to the actors who use them, providing effective avenues for exercising control while simultaneously mitigating potentially socially awkward or damaging situations, with a low social risk. As digital systems become increasingly sensor-enabled and aware of people's activities in the drive to provide more 'real-life' kinds of interpersonal interaction, the space for ambiguity collapses. This reduces people's ability to leverage ambiguity, potentially reducing individual autonomy. Aoki and Woodruff argue that it is essential for designers to design communications systems with some ambiguity left, so that users retain flexibility around coping strategies for social mediation.

The second strand pertains to the use of 'butler lies', which are a more explicit use of deception. Butler lies are used for the purposes of simplifying social channel mediation. For example, in order to terminate a conversation without leaving the other party feeling unimportant, one might use the excuse: 'I have a meeting with my boss', when in fact, no such meeting exists [21]. Reynolds et al. found that butler lies amounted to 27.1% of all communications mediation messages sent by participants who texted partners and co-workers [39]. The purpose of such messages was often to mitigate social awkwardness or reputational damage, and reduce attentional load. Additionally, they were used independently of the nature of the relationship among communicating parties. Such lies were particularly employed in awkward situations, such as when a speaker wishes to convey the that they are 'too busy to interact with the recipient, but not too busy to interact with others' [39]. A detailed analysis of the effects of having the lies exposed showed that the potential emotional impacts were overall small, and much less significant than for other kinds of lies.

The important role of these studies is to demonstrate the beneficial nature of some forms of deception and witholding of information as coping strategies to deal with the complexities of social coordination and an increasing demand for communication. The goals of *translucence* in communication systems [17] are noble, but they must be balanced against users needs for ambiguity and autonomy.

While lie maintenance required to avoid discovery may be trivial ('sorry, I'm hungry, have to go!'), it becomes complicated as lies extend over time, loose their ephemerality, and become woven into the social fabric. The ability to compare multiple accounts of history – especially once the time travelling robots are involved – means that dissonance within the social fabric is more obvious than weaknesses in a single thread.

1.4. Managing Multiple Identities

A natural part of online life is the ability to tailor the persona we present to different communities and contexts. An individual may want to disclose certain things to their professional colleagues, while presenting differently to friends and family or non-mainstream

friend groups. As an example, prolific content creators on Youtube have multiple online personae for different kinds of content—'official' channels, a personal one, and sometimes channels for the works of particular characters [20].

Pseudonyms are one way of representing personae. In [12] Dalton describes pseudonymity in social machines as a one-to-one or a one-to-many relationship – between the human and the pseudonyms. It is possible for a person to be consistently known by a pseudonym over several systems, or use many pseudonyms because they are easy to create and to maintain. However, the many-to-one relation is possible as well, as illustrated by Anonymous, or by Nicholas Bourbaki in the field of mathematics, where the same pseudonym is used by a group. The group can evolve in time while the pseudonym remains.

Social networks are prone to 'context collapse' [8,34] where the personae are forcibly merged. The spreading and linking of data across multiple sources – *leaky containers* – contributes to this contextual collapse. In particular, data which is rooted in physical fact provides multiple opportunities for joining up otherwise separate databases. These data leakages affect users in a manner which is at once invisible and corrosive to the construction and maintenance of personae.

1.5. Verification and Provenance as an Alternative to Sharing

Sharing is a crude mechanism. Once data has been shared the originator can no longer exert control over it and must rely on the behaviour of the recipient, which, as noted, may fail to meet user expectations. As danah boyd notes: 'Any model of privacy that focuses on the control of information will fail'. This leads the teenagers that boyd studies to engage in *social steganography*, manipulating messages so that 'Only those who are in the know have the necessary information to look for and interpret the information provided.' [9]. Strategies like this work when there is a difference in understanding between the surveilled and the surveiller, and collapse as soon as the comprehension barrier is removed.

Validation, however, is a more subtle tool: if a user's personal dataset can be made sufficiently questionable as to be useless on its own, then locus of control shifts to the user choosing to validate parts of the dataset which can be performed in a more nuanced, contextualised manner. If a user is the final arbiter of trust, they can decide to i) sign parts of their record, so that it is verified public fact; ii) co-sign it with another entity, so either can also verify it but not anyone else; iii) verify it through an anonymous channel, so that the entity to whom they provide verification cannot propagate the claim further. This verification can be carried out entirely separately from the data store itself, allowing for the presentation of different datasets as valid in different contexts, as well as unorthodox methods such as using the Bitcoin blockchain to notarise datasets, so that they can be verified in the future without revealing them as true at the time. A conceptual move towards verification and provenance-based approaches underpins the obfuscation strategies outlined in the rest of this chapter.

2. Review of Current Approaches and Tools

Privacy tools for end users of the Web have focused on approaches to allow people to cloak their originating location (IP address) and identity online, as they access web sites

that may be instrumented with code from any number of advertising networks and tracker agents. The most basic of such tools constitute simple browser add-ons that explicitly block the execution and downloading of web beacons, tracking pixels and other tracking agents, through use of a dictionary-approach (of known tracking agents) (e.g. DisconnectMe[4] or AdBlock Plus[5]), or using adaptive algorithms to infer cross-site tracking (e.g. Privacy Badger[6]). Similarly, plugins such as HTTPSEverywhere[7] force the browser to communicate with all web sites using encrypted HTTP, to prevent ISPs and other intermediaries from eavesdropping on traffic between the end user's computer and the end-user host. The use of HTTPS also thwarts some deep-packet inspection (DPI) based tracking methods [29] – such as those employed by public WiFi access points at international Starbucks coffee shops, which are used to determine customers' interests and also to filter and re-rank search results to benefit Starbucks.

In the middle have been a handful of application- or service- specific anti-tracking tools, including TrackMeNot [24], a plugin which aims to reduce tracking and personalisation done by search-engines by diluting the user's query within a flood of other, random search queries. Similarly, CacheCloak [35], intercepts requests for a user's location from various location based services, and returns plausible nearby, but inaccurate, results.

Peer to peer systems such as Bittorrent, Freenet and others pose additional challenges for privacy because a user's client typically connects directly with a large number of other individuals' clients. This means the client's location (IP address) is typically highly visible within the network. However, in some cases, depending on particular details, approaches like Tor (discussed next) can be used to disguise location. Work on disguising an individual's interests within this network – to promote plausible deniability of intent – includes projects such as SwarmScreen [11] and OneSwarm [26]. Other approaches, including encrypting all transferred queries and data have been taken by systems such as Mega.[8]

Perhaps the most sophisticated current tool developed for obscuring network activity and traffic origin is Tor [14], which works at the network-level to hide the origin of packets when communicating with a website or other third party. Tor works as a peer to peer overlay network that routes Web and other network requests through a randomly selected circuit of hosts on the network using the onion routing technique, which makes it intractable to deduce the origin of a particular packet. In conjunction with end to end encryption (such as over HTTPs), Tor has been shown to effectively thwart eavesdropping and DPI methods.

Even with such tools, ISPs, mobile broadband providers and other last mile internet access providers can collect a significant quantity of low-level data pertaining to the physical place and times that a person accesses sites, the quantity of data exchanged, and the potential destinations. To thwart this level of tracking, several hardware vendors and operating systems are starting to incorporate the ability to perform identifier randomisation (such as MAC-address and Bluetooth hardware address randomisation.[9] Such features

[4] DisconnectMe – https://disconnect.me/

[5] AdBlock Plus – https://adblockplus.org/

[6] Privacy Badger – https://www.eff.org/privacybadger

[7] HTTPS Everywhere – https://www.eff.org/Https-everywhere

[8] Mega – www.mega.com

[9] Such features are expected to be introduced as a standard feature to the consumer market for the first time by Apple in iOS 8 – http://appleinsider.com/articles/14/06/09/mac-address-randomization-joins-apples-heap-of-ios-8-privacy-improvements

are standard for an emerging group of speciality *security hardened* devices, including security-enhanced mobile phones and tablets (such as the Black Phone[10]).

While these tools are dedicated to obfuscating action and identity for actions in the digital realm, automatic tracking of people and their activities in the physical world has also increased, thanks to improvements in facial recognition, the introduction of digital tokens and keys for granting people access to physical spaces, membership cards (loyalty programmes), credit cards, and so forth. A number of projects and approaches have similarly been introduced to reduce trackability in this space. In terms of protection against facial recognition algorithms, for example, Harvey et al introduced a special make-up technique based on reverse engineering the most popular face detection algorithms, called CV Dazzle [22]. To reduce infrared camera based person tracking Harvey also introduced heat signature cloaking burqas and hoodies in a line of clothing called *Stealth Wear*.[11]

To reduce city-wide tracking using transport cards such as the London Oyster card and NYC MetroCard, as well as purchase tracking via loyalty card schemes, a common practice that has arisen in several cities has been to host social meetups where people regularly gather and swap their valueless cards. This allows them to confound data mining algorithms by eliminating the assumption that a single card will be owned and used by a single person by having it used instead by a constantly evolving group [31].

3. A Selection of Obfuscation Strategies

The aim of this chapter is to set out some strategies by which the user can add some disinformation into digital social networks in order to regain some measure of privacy. Since every situation is different, from the operation of the network to the desires of the participants, it is necessary to develop several strategies, and understand under what circumstances they may be appropriate and what tensions between utility and privacy arise from their use.

As a starting point, Alexander's taxonomy [3] discusses several types of disinformation which relate to modifying single messages along schema such as:

redaction, where some or all of the information in a message is hidden;

airbrushing, where some of the information is changed: this can take the form of *local crowd blending*, where small alterations are carried out so the new message is similar enough to be plausible, or *global crowd blending* where messages are heavily altered in order that they resemble plausible, but quite different messages;

curveball, where extra, distracting information is added to messages which pushes them into a low density area of message space.

However, due to the pervasiveness of modern communications, we are concerned with modifying message *streams*, where a trace of multiple values must be considered. The social aspect inherent to modern communication tools increases the range and frequency of interaction with others, which in turn increases the chance of lies being exposed, while at the same time opening up the possibility of colluding to strengthen the obfuscatory practices.

[10] Black Phone – https://www.blackphone.ch

[11] Stealth Wear – ahprojects.com/projects/stealth-wear

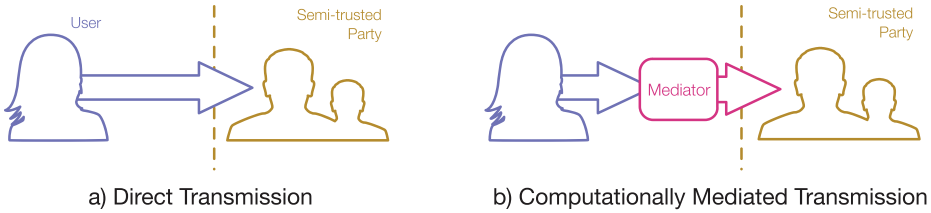

Figure 1. Models of interaction with semi-trusted services. a) Direct transmission of information; b) computationally mediated transmission, where some friendly computational machinery is enlisted to aid in obfuscatory processes.

In this section, we present a range of obfuscation strategies, some of which are speculative, but many of which are drawn from existing examples both inside and outside the digital sphere.

For each strategy we discuss: *i*) what kind of alteration of baseline data is performed; *ii*) what the motivation is for doing it; *iii*) what the possible use cases are; *iv*) how some form of computational support aids in the deception; *v*) some of the systems (if any) which do this currently.

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 chapter is as follows.

> 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 wish to verify that they were at a particular location.

The service is hence *semi-trusted*: there are some benefits which the user wishes to accrue, but there are aspects of the service which make the user unwilling to entrust their complete location history to it. We have chosen location as a clearly understandable facet of personal data, and one which can be easily used to re-identify individuals from anonymised datasets [13].

The standard model of interaction (Figure 1a) involves the user submitting their data directly to the service; for our obfuscatory techniques, we would like to enlist computational support (Figure 1b). This is typified by, but not limited to, mediation from a personal data store, which acts on behalf of the user to modify the data which they provide. In Figures 2 and 3, we plot a fictitious one-dimensional location measurement against time in order to give a sense of how obfuscations unfold across time in multiple locations.[12] We show the individual's true location as a continuous line, along with the measurements made by their device; we then overlay the points which would be submitted on their behalf after obfuscation.

3.1. Strategies for the Lone Obfuscator

Figure 2 lists a collection of possible obfuscation strategies, with Figure 2a showing the true location as a curve, and the dots representing reports of this position to the location-aware service.

[12] While a two dimensional, map-based representation would be more immediate, it is then difficult to clearly show temporal aspects.

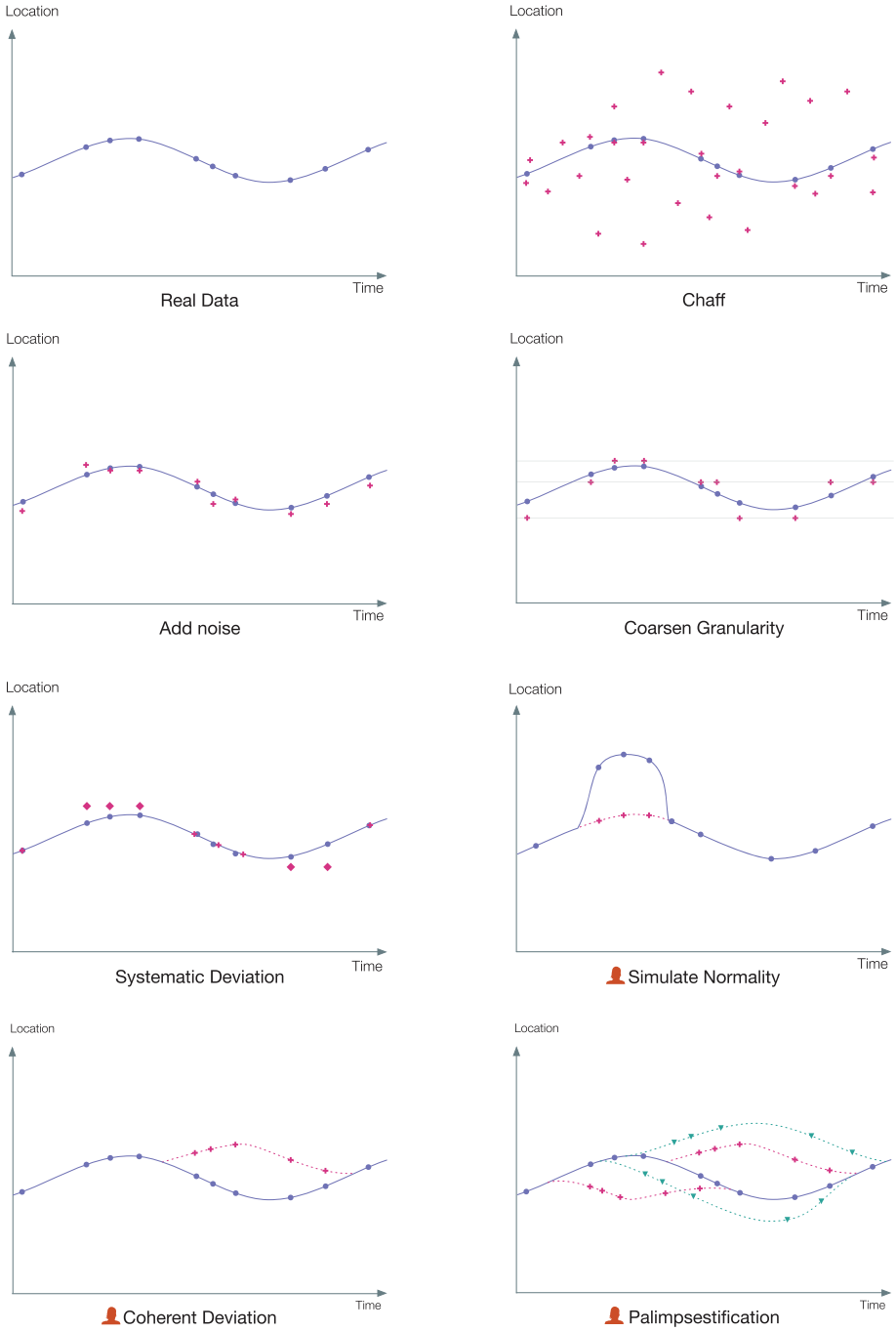

Figure 2. Obfuscation strategies for the lone agent. Location has been reduced to a single dimension so it can be plotted against time. Real data is shown in purple with circular points. Constructed or manipulated points have different shapes. Schemes which require a simulation of human behaviour are marked with a head icon (👤), and where an alternative model of the world is created, it is shown as a dotted line.

3.1.1. Chaff

World War II fighter planes would emit clouds of radar reflective *chaff*, which created multiple traces on the screens of radar operators, and hence disguised the true position of their aircraft. In a similar manner, we can add in multiple location data points alongside the real ones, making it difficult to determine what the true values are. This is the one of the few methods where the complete, accurate data stream is stored. Hence the user can still access any benefits which rely on accurate information. However, adding a multitude of randomised points to a service which expects a single contiguous trace is both easily detectable, and may break functionality – a run tracking application would be likely to give unreliable distance estimations in the presence of chaff.

3.1.2. Noise Injection

The most computationally simple form of obfuscation is the addition of 'noise' to the reports which are sent to the semi-trusted party. Here, the points which are submitted deviate from the true values in a random manner. This allows the user to conceal their exact location, while giving a broad indication of where they are. Depending on the level of noise, this can allow the use of location based services without revealing much about actual behaviour. For example, it might reveal your location on the high street so you can arrange to meet friends, but without revealing which shops you were visiting. This is compatible with services which expect coherent location data, and may be indistinguishable from the inaccuracies of the location sensors. One downside is that the "true" location traces are not present in the record of the service. For example, TripAdvisor can still provide a good enough list of recommended attractions around the given 'noisy' location, however a navigation application will not be able to provide reliable directions.

3.1.3. Coarsened Granularity (or Quantisation)

Rather than adding noise to the data being sent, it can instead be quantised to a coarser granularity, akin to blurring, or zooming out on a map. Again, this is a technique which may help to derive useful information from the service without revealing more than is necessary: using a service to find friends in the same city should only require city level information to be shared. An example of this can be found in Android's permission system, which has separate controls for `ACCESS_COARSE_LOCATION` versus `ACCESS_FINE_LOCATION`.

3.1.4. Systematic Deviation

In some cases, it may be possible to introduce systematic deviations into the digital record. In order to do this, the user needs to be able to define which points to alter, and what to replace them with. One possibility would be thematic replacement – 'hide the times I went to the pub by saying I was at a cafe'. Another would be to disguise the user's home and work locations – places where they are less likely to require location based searches, but which make it very easy to re-identify them from anonymised data. It is likely that this will require some form of computational support to i) identify targets for replacement as they occur and ii) find suitable replacements. Using this technique, some, but not all of the true data is stored; however, derived information – such as beverage preferences in the example above – can be wildly and purposefully distorted. The nature and fact of the distortions may be hard to uncover, as no simultaneous traces or strange movement patterns are produced. Depending on the domain, subtle alterations may have large effects.

3.1.5. Pretend to be me

With increasing computational support, it becomes possible to create a model of the user which outputs plausible 'normal' data. This can then be used to replace periods of abnormal behaviour, or even replace normal behaviour with statistically similar but untrue data. An early example is when neighbours (or automatic switches) are employed to turn lights on and off in a home which has been vacated for the holiday, disguising the true anomalous data of a dark, empty house with the appearance of normal occupation. Similarly, one might avoid making Facebook posts which indicate an absence to avoid burglary. This kind of deception can be difficult to achieve. However computational systems are emerging which can aid users, for example Beyer's digital alibi system [5].

3.1.6. Coherent Deviation

As the converse of simulating normality, the user may wish to pretend to be somewhere they are not. This is similar to creating systematic deviations, but on a grander scale; the user would like to create a narrative for the deviation, and then have suitable data points constructed. For example, the user might pretend to be on holiday, or at a conference, and would like location traces which match that narrative to be created, such as going to the convention centre in the day, and returning to a hotel at night. This requires a computational model of user behaviour which can be applied to new locations – a non trivial task. However, there is the potential to create obfuscated data which is difficult to distinguish from standard behaviour. Services like Please Don't Stalk Me[13] enable geotagging tweets with fake location information, and the Chrome browser supports faking location information in the browser, through the Chrome Developer Tools emulator. These tools can be used for some of the previous strategies, for noise injection by using random locations, or for systematic deviation, when only the location of certain tweets is modified.

3.1.7. Palimpsestification[14]

Taking the idea of coherent deviations a stage further, and combining with the idea of *chaff*, the user could create multiple overlapping traces; each trace would be locally coherent and plausible, but someone inspecting the data would have no way of knowing which was the real one. This is similar to the strategy of CacheCloak [35], which continuously generates sheaves of probable future behaviour and searches location based services relevant to each path. The computational support required is similar to the coherent deviation example – to be able to run a model of the user's behaviour in novel locations – although more coordination might be required between the stories. The trade-off is that while the true location data can be entered along with the generated points, the deception is obvious, and location based services may become upset at the multiple paths.

[13] http://pleasedontstalkme.com/

[14] 'Palimpsest' originallyer referred to a document where orignal markings have been abraded so that it may be re-used, but still contains some traces of the original, so the texts are superimposed. Modern usage covers a range of other situations when multiple activity traces are overlaid.

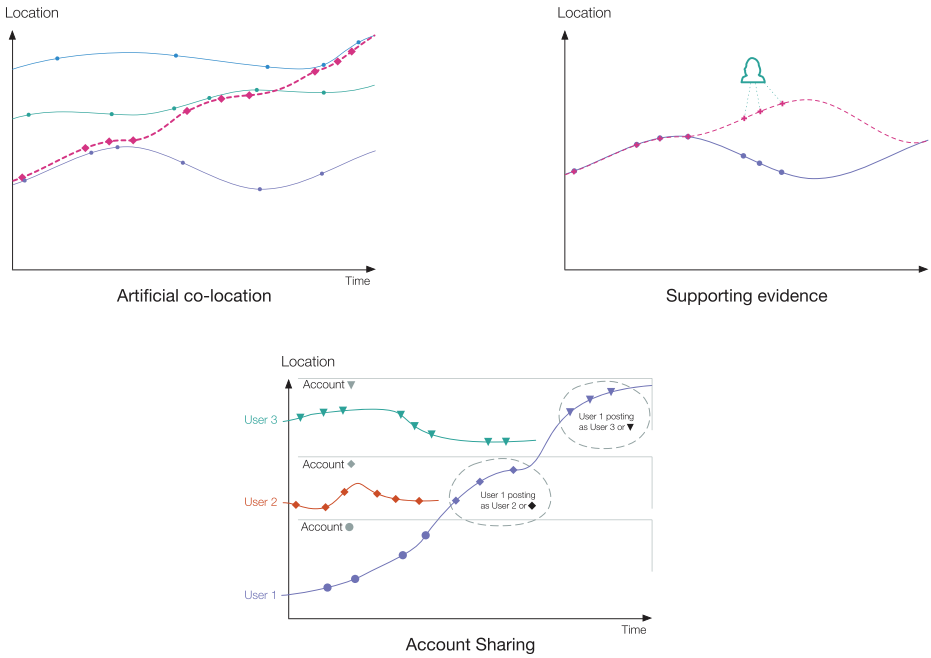

Figure 3. Three possible multi-person obfuscaction strategies. As before, modelled behaviour is shown as dotted lines. In the *Artificial co-location* strategy, circles represent real datapoints, diamonds are the ones that the obfuscating user submits. In the *Account sharing* strategy, the shape of points represents the account through which they are submitted.

3.2. Collaborative Obfuscation

Including others in the obfuscation challenge opens up a range of new strategies, where collusion can aid in the creation of otherwise unachievable data streams or increase the veracity of artificially created data. Generally the possibilities in computational systems are analogous to pre-computational possibilities; computational support tends to be in the form of coordination to find collaborators or and check coherence of data points. The ideas outlined here are more speculative, as few computational systems of this type exist. There are aspects which make these strategies harder to pull off as coherence is required across multiple different accounts; however the counterpoint is that if successful, the obfuscation is better supported and harder to detect.

3.2.1. Artificial Co-location

One way to obtain a realistic but untrue location trace is to re-present the trace of a collaborator. This can look like relatively natural behaviour; two people meeting up to carry out joint activities or socialisation. Computational support here can involve finding accomplices to ‘co-locate’ with – people who are willing to share their location, and are behaving in ways which match the desired story – as well as the technical details of transferring location devices between accounts.

3.2.2. Supporting Evidence

Co-located people often share the fact of their co-location, explicitly or implicitly, whether in group photos – 'Here's me and X on top of the Scott monument', broadcast messages – 'Just been hanging out with X at the coffee shop', or shared plans – 'Going to the cinema with X tonight – anyone want to join in?'. Enlisting collaborators to make these kinds of posts can help to add depth to a constructed trace, weaving it more tightly into the social fabric.

3.2.3. Account Sharing

In a similar manner to the swapping of loyalty cards discussed in [10], users of services can share accounts. This results in an account or set of accounts with more or less plausible activity, yet allows the users to remain unidentified. Much as the loyalty card swapper confounded efforts to inspect individual buying habits, or the 'Anonymous' movement aggregates the activities of a multitude of participants behind a stylised mask, services such as DuckDuckGo aggregate many people's search results, ensuring that the search providers cannot build up any identifiable user histories.

Computational intelligence can be enlisted to support many different ways of assigning people to accounts, such as the following.

Many to one schemes have a single account controlled by multiple people. This can result in a completely incoherent manner; DuckDuckGo's aggregated search makes no attempt to imitate individual behaviour. Alternatively, sharing can be tied into a coherent shared identity, where multiple people contribute to a single shared persona [12]. Here there is a challenge to maintain consistency: when multiple people control a call centre's chat avatar, they must ensure that the relevant information and state is shared [*ibid.*]. When multiple users control a single game character, the game world enforces consistency, and the community must produce coherent action streams in order to progress.

Randomised schemes allocate accounts to people without a discernible order. When loyalty cards are mailed between anonymous participants, there is an explicit desire to produce implausible data in order to confound analysis. Online accounts can be similarly shared, leading to traces which are unlikely to have been produced by a single individual. In our locative service example, this allows users to access benefits which do not rely on individual history while preserving some level of privacy. Computational support involves finding accounts to share, and ensuring that each account is only accessed by a single person at any given time.

Structured schemes allow for accounts to be used as appropriate according to some criterion. If a location service offers history based benefits (e.g. loyalty rewards or reputation) then it could be beneficial to borrow a local user's account when going on holiday – couchsurfing but with login credentials instead of flats. Infrastructure would be required to discover appropriate accounts, and mediate access.

4. Operationalisation – Managing Deception and Its Side Effects

4.1. Going Beyond Location

In Section 3, we discussed obfuscatory possibilities with respect to a location based service; however, this is a single application area, and the need for regaining informational autonomy is felt across a spectrum of data types and services, hence we must discuss how these techniques generalise.

Location data is generally collected by a device which the user owns. In many cases, this is a smartphone, which uses a combination of GPS, cell tower triangulation and WiFi access point locations to determine a user's position in space. The user then has some level of choice about who to share the data with and how. This is not always the case, however: cell tower records can identify user's locations – and individuals can be picked out from very sparse histories [13]. There was a recent outcry[15] when it was discovered that smartphone manufacturers were collecting location data from their users without their knowledge, despite the possible mitigations of good intentions or technical reasons.[16]

Obfuscating data (location or otherwise) can be done only if the user is aware that the data is being collected in the first place. Some knowledge about the data collection process, such as what data is captured and when, can then help improve the obfuscation.

Some data, even when known to be collected, cannot be obfuscated. Examples include all official or administrative data: census data, police reports, tax returns, health records, but also dealings with private organisations with a regulated status like banks and utility providers. Obfuscating these types of data in any way is often a criminal offence. For example 'obfuscating' one's source of income by adding chaff or noise may be construed as tax evasion, theft, money laundering, etc.

For data that can be obfuscated without legal consequences, and for which we are aware of the data collection, we have the opportunity to use some of the strategies described in Section 3, although they must be adapted to fit the data and the situation at hand.

For a photo sharing application or microblogging tool, adding *chaff* could mean posting more content – but fabricated. It can be realised by posting public domain photos on Instagram from a randomly chosen topic, or flooding Twitter with randomly generated text, older posts, or fragments from existing texts like web pages or books. For browsing history in an online shop, we can use bots or browser extensions, or tools like Selenium[17] to simulate browsing of various items by randomly requesting pages beside the page we are interested in. While it does not hide the fact that a person visited the shop, it can obfuscate their interest. This can be useful when dealing with Amazon, or another large scale marketplace, but it can be less useful when the online shop is catering for a small niche market. More subtly, chaff can be added to the tags attached to photos or posts, to make categorisation more difficult. The goal of chaff is not to be believable on deeper inspection, or even consistent, so the algorithms for choosing what to post and where to browse need not be too complex.

Prepaid, disposable mobile phones ('burners') are used in order to avoid surveillance, as they break the link between a person's identity and their cellphone number. Similarly, ghost email addresses can be used to obfuscate online activity, as it prevents the email

[15] O'Reilly Radar: http://radar.oreilly.com/2011/04/apple-location-tracking.html
[16] Apple's response: http://www.apple.com/pr/library/2011/04/27Apple-Q-A-on-Location-Data.html
[17] http://docs.seleniumhq.org/

address from being used as a global identifier for the person. For websites which require registration, disposable email provider services[18] mean that a different email address can be used for each site, while still receiving the email confirmations in a 'real' mailbox. These email addresses can still be linked back to the true user if the service is compromised in some way; for truly disconnected disposable email addresses, services like GuerillaMail, 10 minute mail, etc. provide temporary email addresses with a short life span.

Applying *noise injection* or *coarsened granularity* strategies can limit the benefit of some services, and in some cases defeat the purpose of using them altogether – like injecting noise in the images posted on Instagram.

These strategies can, however, be used successfully with optional data that is required by some services but which do not impact on the service provided. Randomly changing your date of birth or hometown, or even your name on Facebook or Twitter, will still allow you to use the service to connect with friends or post updates respectively. Changing (injecting noise into) the date of birth can be used as the real date of birth approaches, to avoid the congratulatory messages, and reverted back to once the date has passed. A different type of changing granularity is withholding the year from the date of birth, thus allowing the actual date be know, without revealing the age. Any functionality of a service that is regarded of low importance can be used as a place where noise can be injected, or where the granularity can be coarsened. This of course leaves part of the information unobfuscated.

Changing the declared date of birth can also be *systematic*, a user can choose an untrue date which to reuse in *all* services that require it. Using the same completely fabricated persona in all online communication is supported by services like Fake Name Generator[19] which will not only create geographically localised names, but also provide valid additional information like address, credit card, email address, phone number, mother maiden's name, username and password, employment information, weight and height and blood type, and even a favourite colour. Using the persona systematically and consistently enables a complete separation between the real person and their presence online.

Several personae can be created and used to simulate collaborative obfuscation strategies, where the fake characters mimic the existence of supporting evidence for the lie. *Astroturfing* uses multiple accounts to simulate widespread support for a topic, organisation or political message, creating the illusion of a grassroots movement. It can, however, be used for the obfuscation of individuals' data, by corroborating claims, and adding depth to the constructed scenarios created.

4.2. Personal Data Stores – Allies on the Intimacy Battleground

Personal data stores (PDS) represent a vital component in the issue of online presentation: having trusted, user controlled repositories for data enables a more user-centric approach to management of *capta*, filling in the 'Mediator' component of Figure 1. Bridges can then be built between personal data stores and the rest of the world in order to support the connected, networked interactions which users now expect. If these bridges simply share the data, even in a controlled manner, nothing has been gained; hence the bridges become conduits for manipulating truth and constructing falsehoods. As personal data

[18] A comparison between some disposable email providers is available on LifeHacker http://lifehacker.com/5306452/how-do-you-keep-your-email-address-private

[19] http://www.fakenamegenerator.com/

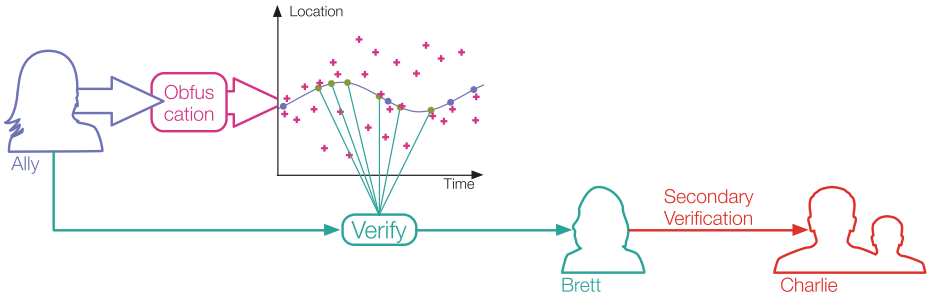

Figure 4. Example verification scenario. The user (Ally) provides a set of real data, plus *chaff* to a location aware service. A third party (Brett) then requests verification of some of the points, which Ally provides. Brett then wishes to share the data with Charlie, which requires Brett to verify to Charlie that the data are correct.

stores accumulate more real-time contextual data about the individual, as well as about the individual's social connections, PDSs can provide support for the often stressful and mentally burdensome task of lie maintenance, for example: i) identifying when a person's real activities or whereabouts contradict a lie, and might be discovered; ii) identifying indirect social channels that could expose a lie (e.g. through friends of friends); iii) suggesting appropriate lies to use which are least likely to be detected; iv) suggesting individuals to lie to to support lie maintenance (e.g. friends of the person being lied to).

4.3. Verification and Provenance Mechanisms

In Section 1.5 we suggested that verification is a more nuanced mechanism than control over sharing. One of the effects of the obfuscation strategies discussed previously is that it becomes impossible to know which parts of the user's data-stream are grounded in reality, and represent 'true' values. This means that if someone wishes to engage with the data and have an expectation of accuracy, they need to ascertain which parts of the record are correct. This shifts the locus of control from the process of sharing to the process of verification – the user can make claims about subsets of the data points currently attributed to them.

Let us consider a scenario where Ally has some personal data, which Brett would like to make use of. Brett also wants to sell Ally's data to Charlie.

There are a range of statements which Ally can make, including: 'this subset of data points is mine', 'these points are within 50 m of my true location', 'these points are representative of my general behaviour' and so on. The choice of which point to claim can be negotiated in the context of the question being asked, and Ally can determine what is and is not acceptable.

If the external agency wishes to disseminate the users data, it becomes an issue of propagating the trust which the user has given them – essentially, Brett must say to Charlie: 'Ally has verified these points to me, and now I am verifying them to you'. The manner in which the initial verification was carried out now becomes critical:

- if an email or similar communication is used, Ally simply declares 'these points are mine', then the secondary verification is only as strong as trust in the communication chain – Brett must convince Charlie that the email or message is genuine and emails are easy to fake.

- Ally could use a technique which would give Brett no future tangible proof of the verification – for example, a single use URL which lists IDs for the correct points; Brett would have no evidence with which to convince Charlie that Ally had verified the points, other than reputation alone.
- Ally can cryptographically sign the claim using public key cryptography. The claim is then essentially public knowledge, and anyone can check Ally's verification.
- Ally can sign the claim after Brett has; this means that Brett cannot hide the fact that they were the recipient of the claim, so it is impossible to propagate the claim anonymously.

All of these techniques relate to making the public record so unreliable that anyone who wants to use any of the data will need to separately establish a chain of provenance for certain parts of it. A related goal would be to make it illegal, or at least unacceptable, to use personal data without having a valid provenance chain for it. Essentially, in order to use anyone's data, Charlie would have to explain how they came to have it, and be able to prove that Ally had shared the data originally.

4.3.1. Notarisation

The verification examples above rely on Brett trusting Ally about which data points are correct. There may be times – e.g. when creating alibis – when Ally needs to have a stronger form of proof.

In this case, third party digital notarisation services can be employed.[20] These services take in some document or datum, and provide a certificate which can be used to verify that that piece of data was provided at a certain time. For example, if someone wants to make a prediction for the outcome of a football match, they could notarise that before the match, and then subsequently prove that they had made the prediction beforehand. It is generally not possible to prove that they only made a single prediction, so this technique is most suitable when the range of possible things to notarise is so large as to make notarising the entire space infeasible.

With regard to personal data, we can notarise our true data stream as we produce it. This means that we can prove that we had considered those points at the time, and if we say we were in a particular place, there is a high chance we were, however, it does not work in the complementary situation as producing a notarised point does not prove that we were not anywhere else.

Notarisation does not necessitate revealing the data itself. For instance, when submitting a location, a representation of the time and place could be hashed, and this hash notarised (Figure 5a). Additionally, points can be notarised in sequence, so that we can demonstrate contiguous sub-sequences of points as having been provided previously; by hashing the current location with the previous location, we can link the points together, to build up confidence in the notarised results (Figure 5b).

4.4. Effects and Ethics of Obfuscation

Brunton and Nissenbaum discuss the ethics of obfuscation at length in [10], and address issues like wastefulness, dishonesty, free-riding, pollution, and possible system damage. One answer is 'that obfuscation has no ethical or political valence of its own, only to the

[20] e.g. http://virtual-notary.org/, a free service hosted at Cornell University.

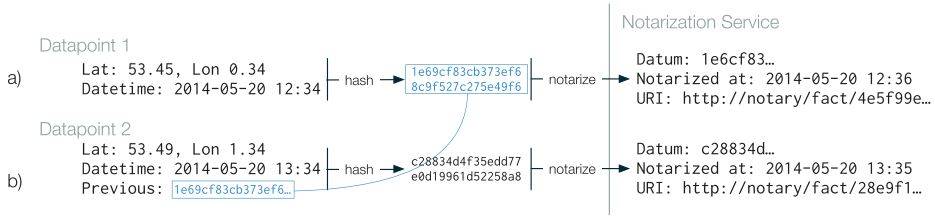

Figure 5. Notarisation of personal data. a) Data points and times are hashed, and the values sent to a notary service, which provides a URL to verify that i) the given data was supplied and ii) when it was supplied. Hashes are used so that the data is not publicly shared. b) If the hash of the previous submission is included, then sequences of consecutive points can be verified.

ends that it serves' [ibid.]. However, with the sharing of personal data, there are multiple overlapping contexts in which the obfuscation takes place, which should be dissected.

On a *personal* level, obfuscation changes the picture of us that a service has. For different services, different levels and types of distortion may be bearable or desirable: while adding noise to a location signal may still allow sensible location based recommendations, pretending to like random products might make a recommender service unusable. One of the reasons so many obfuscation strategies were presented is that each strategy preserves different qualities, and the choice of what and how to obfuscate must be made in context, as there is an unavoidable tradeoff between privacy preservation and maintaining the level of personalisation that makes a service worthwhile.

On a *social* level, being caught in a lie can be unpleasant; as Bok says:

> In practice, however, lying to enemies has enormous drawbacks... lies to enemies carry very special dangers of backfiring. All too often, the lie directed at adversaries is a lie to friends as well; and when it is discovered, as some always are, the costs are high [6, p. 141].

While the reputational damage from lying can be significant, some steps can be taken to mitigate this. Obfuscation techniques can be chosen which are clearly non-true to human observers, or to one's trusted circle of friends, echoing boyd's *social steganography* [9]. The alterations will then have the desired effect of confounding automated surveillance, without giving one's friends the sense of being deceived.

Additionally, systematic data manipulation can result in presenting an online persona which diverges from one's own. Recently, Mat Honan tried the experiment of 'liking' everything on Facebook [23]. After two days of adding this *chaff* to his activities, his public persona demonstrated radical far-right and far-left viewpoints, to the extent that his friends became concerned.

As well as presenting an untrue picture of oneself, systematically adding noise may also have unplanned side effects that might directly impact one's friends and connections. For example, Mat's friends suddenly found themselves deluged by the stream of liked articles he was generating, to the point where his junk data stream overwhelmed others' friends' legitimate statuses and impeded their ability to use the social network normally.

On a *societal* level, while obfuscation overlaps both 'pro-social' lying and 'butler lies', it must also be recognised that systematic untruths can weaken the social fabric, disrupting trust with friends and coleagues, as well as those carrying out surveillance. This adds friction to online interaction, as one must put the effort into constantly questioning

what data is true and false. On a broader scale, the social good which comes of having access to increasingly detailed personal data can be compromised if significant proportions of the data are untrue.

In terms of *accountability*, there is the danger of being unable to verify your real behaviour. While the possibility of notarisation (Section 4.3.1) can create support, any defence which relies on a stream of information which has been systematically manipulated is likely to be problematic. As online systems are increasingly being used to account for who you are and what you do, making your online persona questionable could have far reaching consequences. A complementary risk is that the constructed data might be accidentally incriminating: a spuriously generated location point might place a user at a crime scene, or a participant in an account sharing program could use a borrowed account while engaging in criminal activity. There is also a danger that the same tools that preserve personal privacy are applicable to covering up illegal or anti-social behaviour (e.g. cyber-bullying).

Finally, at a *systems* level, the context is dynamic, and services will respond to obfuscatory practices. Services may shut down or transition to a payment based model in response to the declining value of increasingly noisy data. Future services will be better at spotting obfuscated data, and may take steps to prevent it, analysing the coherence and plausibility of incoming data. Obfuscation systems will have to evolve along with the context in which they are used, creating an arms race between obfuscators and service providers.

These issues suggest firstly that obfuscation is a form of *free riding* [10]: if everyone did it the system would grind to a halt, but if a few people do it, they can take advantage of everyone else's truthfulness to preserve their own privacy. However, if the majority of the population engaged in obfuscation, this would be a clear signal that the balance of privacy was unacceptable to the general population, becoming a voice of protest.

Secondly, it is clear that more understanding is needed around what kinds of obfuscation to apply when, and how to create tools to enable these practices in the context of connected, visible, online social behaviour.

However, there is a view that anonymity is natural and healthy both on and off-line, and that methods that force full identifiability at each interaction are both difficult, fragile and unhealthy for long term growth. Proponents of this view (e.g. [9,30,45,46]) argue that obfuscation approaches are not just ethical, but necessary for individuals to maintain the autonomy and freedom from being entrapped into a single identity tied to their presence in the physical world. In particular, it has been argued (e.g. [46]) that the freedom to have multiple, flexible identities, including borrowed identities [12], group identities, and role-based identities, fosters the kinds of identity experimentation that leads people to learn how to cultivate healthy social relationships throughout their lives. This principle, known as the 'elastic self', has been responsible for key growth in online communities to date, and is what is being threatened by the surveillance and real-name policies mentioned previously.

The approaches outlined here, in concert with other tools to aid privacy, anonymity and ambiguity in communications, help to redress the current power imbalance and reduce the effects of caustic surveillance. By adding some elasticity to the system, uses regain some autonomy in the way they are seen and sorted online, and develop power over the constitution of their online identity.

5. Conclusion

In this chapter, we have presented a particular approach to shifting the boundary between privacy and use of data. We have taken a purposefully adversarial approach to the protection of personal data as an indication of potential steps which can be taken by individuals at a grass-roots level. These techniques are not the final word in personal data sharing, but sketch out a particular position in the journey towards a balanced societal attitude towards surveillance. In time we hope that stronger legal and social protections of personal data will be introduced, so that society can enjoy the benefits of data sharing while being respectful of the individual's right to privacy [38,40].

Acknowledgements

The authors would like to thank the anonymous reviewers for their insightful and constructive critiques, as well as Reuben Binns and Mark Hartswood for feedback and suggestions. This work was supported by the EPSRC SociaM project under grant EP/J017728/1.

References

[1] Communities against terrorism, https://info.publicintelligence.net/FBI-SuspiciousActivity/InternetJCafe.pdf
[2] A. Albrechtslund, Online social networking as participatory surveillance, *First Monday* **13**(3) (2008).
[3] J.M. Alexander and J.M. Smith, Disinformation: A Taxonomy. Technical report, University of Pennsylvania Department of Computer & Information Science, 2010.
[4] P.M. Aoki and A. Woodruff, Making space for stories: Ambiguity in the design of personal communication systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '05, pages 181–190, New York, NY, USA, 2005. ACM.
[5] S. Beyer, M. Mulazzani, S. Schrittwieser, M. Huber and E. Weippl, Towards fully automated alibis with social interaction. In *International Conference on Digital Forensics*, Vienna, Austria, 2014.
[6] S. Bok, *Lying: moral choice in public and private life*. New York: Pantheon Books, 1978.
[7] K. Borcea-Pfitzmann, A. Pfitzmann and M. Berg, Privacy 3.0:= data minimization+ user control+ contextual integrity, *it-Information Technology Methoden und innovative Anwendungen der Informatik und Informationstechnik* **53**(1) (2011), 34–40.
[8] D. Boyd, *Faceted id/entity: Managing representation in a digital world*. PhD thesis, Massachusetts Institute of Technology, 2002.
[9] D. Boyd, Networked privacy, *Surveillance & Society* **10**(3) (2012), 348–350.
[10] F. Brunton and H. Nissenbaum, Vernacular resistance to data collection and analysis: A political theory of obfuscation, *First Monday* **16**(5) (2011), 1–16.
[11] D.R. Choffnes, J. Duch, D. Malmgren, R. Guierma, F.E. Bustamante and L. Amaral, Swarmscreen: Privacy through plausible deniability in p2p systems. Technical report, Northwestern EECS Technical Report, 2009.
[12] B. Dalton, Pseudonymity in social machines. In *WWW 2013 Companion*, pages 897–900, Rio de Janeiro, Brazil, 2013.
[13] Y.-A. de Montjoye, C.a. Hidalgo, M. Verleysen and V.D. Blondel, Unique in the Crowd: The privacy bounds of human mobility, *Scientific Reports* **3** (2013), 1376.
[14] R. Dingledine, N. Mathewson and P. Syverson, Tor: The second-generation onion router. Technical report, DTIC Document, 2004.
[15] C. Doctorow, *Content: selected essays on technology, creativity, copyright, and the future of the future*. Tachyon Publications San Francisco, 2008.
[16] M. Dodge and R. Kitchin, Codes of life: identification codes and the machine-readable world, *Environment and Planning D: Society and Space* **23** (2005), 851–881.
[17] T. Erickson, W.A. Kellogg and I.B.M.T.J. Watson, Social translucence: An approach to designing systems that support Social Processes **7**(1) (2000) 59–83.

[18] J. Feigenbaum, A.D. Jaggard and R.N. Wright, Towards a formal model of accountability. In *New Security Paradigms Workshop*, pages 45–56. ACM, 2011.
[19] S. Giirses and J. Pridmore, Translating privacy into digital designs: Technical strategies to counter everyday surveillance.
[20] A. Guy and E. Klein, Constructed identity and social machines: A case study in creative media production. In *SOCM workshop at WWW2014*, pages 897–902, 2014.
[21] J. Hancock, J. Birnholtz, N. Bazarova, J. Guillory, J. Perlin and B. Amos, Butler lies: Awareness, deception and design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 517–526, New York, NY, USA, 2009. ACM.
[22] A. Harvey, Cv dazzle: Camouflage from computer vision', 2012.
[23] M. Honan, ILiked Everything I Saw on Facebook for Two Days. Heres What It Did to Me. http://www.wired.com/2014/08/i-liked-everything-i-saw-on-facebook-for-two-days-heres-what-it-did-to-me/, 2014.
[24] D.C. Howe and H. Nissenbaum, Trackmenot: Resisting surveillance in web search, *Lessons from the Identity Trail: Anonymity, Privacy, and Identity in a Networked Society* **23** (2009), 417–436.
[25] G. Iniguez, T. Govezensky, R. Dunbar, K. Kaski and R.A. Barrio, Effects of Deception in Social Networks. *arXiv preprint*, pages 1–19, 2014.
[26] T. Isdal, M. Piatek, A. Krishnamurthy and T. Anderson, Privacy-preserving p2p data sharing with oneswarm. In *Proceedings of the ACM SIGCOMM 2010 Conference*, SIG- COMM '10, pages 111–122, New York, NY, USA, 2010. ACM.
[27] L. Kelion, Nsa and gchq agents leak tor bugs alleges developers. Aug 2014.
[28] A.D.I. Kramer, J.E. Guillory and J.T. Hancock, Experimental evidence of massive-scale emotional contagion through social networks, *Proceedings of the National Academy of Sciences* **111**(24) (2014), 8788–8790.
[29] S. Kumar, J. Turner and J. Williams, Advanced algorithms for fast and scalable deep packet inspection. In *Proceedings of the 2006 ACM/IEEE symposium on Architecture for networking and communications systems*, pages 81–92. ACM, 2006.
[30] L. Laurenson, Towards a more pseudonymous internet. *The Atlantic*, August 2014.
[31] D. Lockton, The fight back: loyalty card subversion, Dec 2006.
[32] D. Lyon, *Surveillance society: Monitoring everyday life*. McGraw-Hill International, 2001.
[33] D. Lyon, *Surveillance studies: An overview*. Polity, 2007.
[34] A.E. Marwick and D. Boyd, I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience, *New Media & Society* **13**(1) (2010), 114–133.
[35] J. Meyerowitz and R. Roy Choudhury, Hiding stars with fireworks: Location privacy through camouflage. In *Proceedings of the 15th Annual International Conference on Mobile Computing and Networking*, MobiCom '09, pages 345–356, New York, NY, USA, 2009. ACM.
[36] A. Narayanan and E.W. Felten, No silver bullet : De-identification still doesn ' t work. Technical report, 2014.
[37] A. Narayanan and V. Shmatikov, Robust De-anonymization of Large Sparse Datasets. In *IEEE Symposium on Security and Privacy*, pages 111–125, 2008.
[38] J. Rauhofer, Future-proofing privacy: Time for an ethical introspection? *Surveillance & Society* **10** (2012), 356–361.
[39] L. Reynolds, M.E. Smith, J.P. Birnholtz and J.T. Hancock, Butler lies from both sides: Actions and perceptions of unavailability management in texting. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, CSCW '13, pages 769–778, New York, NY, USA, 2013. ACM.
[40] B. Rooney, The Balance Between Open Data and Privacy. http://online.wsj.com/news/articles/SB10000872396390443884104577647600306243684#printMode, Sept. 2012.
[41] P.M. Schwartz, Property, privacy, and personal data. *Harvard Law Review*, pages 2056–2128, 2004.
[42] B. Simon, The return of panopticism: Supervision, subjection and the new surveillance. *Surveillance & Society* **3**(1) (2002), 1–20.
[43] S. Spiekermann and A. Novotny, The vision: personal information markets and privacy.
[44] B. Templeton, A watched populace never boils. http://www.templetons.com/brad/watched.html
[45] S. Turkle, Life on the Screen, Simon and Schuster, 2011.
[46] T. Wang, *Talking to Strangers: Chinese Youth and Social Media*, 2013.
[47] S.D. Warren and L.D. Brandeis, *The right to privacy. Harvard law review*, pages 193–220, 1890.
[48] J.R. Whitson, Gaming the quantified self, *Surveillance & Society* **11**(1) (2013), 163–176.