Social Palimpsests - clouding the lens of the personal panopticon

. . . or "Suck my tailpipe" - how I learned to stop worrying and toxify my exhaust data

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 the benefits of sharing, and a distrust of those with whom we share our data.

Keywords. Obfuscation; Data politics; Personal Data Stores; Social Machines; Pants-on-fire;

Introduction

Privacy advocates have been accused of advocating methods to enable criminals, stalkers, harassers [], miscreants, and even terrorists [] to more easily conduct their misdeeds against society. The "I have nothing to hide; therefore I should not care" argument [] used often in defense of ignoring privacy in the early days of targeted advertising and the Web still re-surfaces today. But as theoretical arguments of what *may* become have been replaced by tangible examples of what has actually has, this, among other arguments have become thoroughly dismantled. The sophisticated methods being used to track and pinpoint details of individuals has not been met with a corresponding increase in counter-tracking and counter-surveillance tools; as a result, people are being subject to an unprecedented barrage of targeted attempts to behaviourally manipulate them - often in ways too subtle for them to perceive (e.g. [**?**]).

Thus, this chapter takes the position that privacy protection not merely as a concern for a minority with peculiar sensitivities towards being seen or exposed, but instead, as a set of capabilities necessary for preserving individual autonomy and well-being, in an economy where personal data has become the most precious digital commodity.

Bentham's thought experiment of the Panopticon, developed by Foucault describes a prison in which inmates lives are constantly surveilled as a means of discipline and 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[24].

We are interested in practises which can cloud the lenses of the observers, to choose how we are seen, and to help regain some control of the manner in which our lives are surveilled, both socially and otherwise:

> "We give up on privacy because we live in the Panopticon. A lot of people think there is an inevitability about this [loss of privacy]. But the availability of data does not sanction its use." Sir Nigel Shadbolt, quoted in the Wall Street Journal[23]

In this chapter, we discuss and analyse methods of privacy protection that advance beyond the current state of anonymisation tools that merely obscure the tracks of individuals towards those that employ methods borrowed from information warfare [], in order to allow individuals to regain autonomy from unsolicited tracking and behavioural control. We first discuss X and Y, followed by a survey of lying and falsification in context-aware systems, current anonymisation ad 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, and relation so surveillance

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 [11]. From using an access card to unlock a door at the workplace, 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 collection of 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"[8].

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. The term comes from a loose analogy with *fiat money*: currencies whose value is derived from the mandate of a government[1]. 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.

This use of authority to demand taxes in a fiat currency can be contrasted with modern cryptocurrencies such as Bitcoin; here, there is no central authority to backstop value and demand taxation—rather, the participants in the economy collectively decide what the units of currency are worth. The analagous approach to personal data would be a personal data economy in which participants decided what and how they wanted to share, where sharing was based on personal utility rather than centralised dictat.

In some cases, the provision of personal data is necessary requirement for the provision of a service, but in many cases it represents an attempt by the organisation to create a monetizable product from its users. There is a spectrum of approaches from outright demands to asking or encouraging users to furnish their data. Increasingly, *gamification* is be 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"[25]

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]

[1] Fiat money is, by definition *inconvertible*, and *intrisically useless*, and hence must derive its value from external forces, typically government decree, status as legal tender or by being the required currency of taxation. While the conditions of inconvertibility and uselessness are not necessarily applicable to personal data, the mechanism of central authority decree is a common defining factor

In summary, there are multitudinous situations where we are coerced, cajoled or manipulated into sharing our personal data, and the cost of avoiding such sharing is increasingly becoming untenable for large sections of the population of the industrialised world.

1.2. Sharing is a loss of control

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. **TODO:** *eMax: Do we want to mention DRM here as a futile attempt to control how data is used once it's sh*

TODO: *There is so much more we can say right here about this.*

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

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. A government may decide to share previously confidential data, as in the case of the recent *care.data* fiasco in the UK; a company can be bought and its assets acquired—the purchase of Moves by Facebook raised issues around the terms and conditions of data handling companies; and even without malice accidents 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. Essentially, once data is shared, the sharer has no control over what happens to it.

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 proportion 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?" [1]. Sharing data today cedes control to the entities of tomorrow, with their greatly enhanced capabilities.

De-identification doesn't work: data is often shared subject to the condition that it will only be shared in an *anonymised* or *de-identified* form. 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 [20], let alone access to other, non-public data. Narayanan and Felten's recent report [19] explains in a non-technical manner why de-idenfication of data remains problematic. This is also highly dependant on the data in question; location data is extremely difficult to anonymise, with four datapoints being enough for re-identification in many cases []

Databases can be joined: as more databases of personal information become available, whether publicly or privately, the possiblity to match, join, correlate and shara 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."[17, p.37–44].

TODO: *What you are sort of getting at as "rights" to data and the lack of regulation thereof; there is no notion of data own*

1.3. The case for lying and the importance of anonymity

White lies can aid in social network growth: [16]

Most, if not all, social interactions involve both strategic omissions and various kinds of lies and non-truths to manage the myriad conflicting social demands placed upon us.

Butler Lies: [15]

Translucence: "What we say and do with another person depends on who, and how many, are watching." - [12]

Contrast with Transparent society [7]; power imbalance between parties.

1.4. Multiple Identity

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. **TODO:** *Amy's Content creation on youtube [14]*

TODO: *Ben Dalton's thesis about Persona's throughout the agents [9]*

1.5. Why doing it socially is difficult

Sharing certain personal data is a barrier to anonymity and obfuscation; data which is rooted in physical fact provides multiple opportunities for joining up otherwise separate databases. The lie maintenance required to avoid discovery may be trivial ("sorry, Im hungry, have to go!") but may become significantly more complicated as lies extend over time, 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.

1.6. Why verification and provenance are better than 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 dana boyd notes: "Any

model of privacy that focuses on the control of information will fail.". This leads the teenagers that dana 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."[6]. Strategies like this work when there is a difference in understanding between the surveilled and the surveiller, and collapses as soon as the comprehension barrier is removed.

Validation, however is a more subtle tool: if a users 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 alsoverify 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 datastore 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.

2. Review of current approaches and tools

TODO: *Max to do some writing*

TODO: *Refs to work in...*

- Marwick, public domain: [18]
- Reigeluth, data traces: [22]
- Beer, algorithms and power: [4]
- Goldberg, public/virtual participation: [13]
- Rauhofer, Future Proofing Privacy: [21]
- Simon, panopticism [24]

- The revolution has started!
- tor, anonymous remailers, burner phones, gotta change up, yo!
- HTTPSEverywhere
- Surveillance and Society: `http://library.queensu.ca/ojs/index.php/surveillance-and-society`
- CVDazzle
- Heat-signature cloaking burqas, hoodies
- Bluetooth and MAC randomisation in iOS 8
- Silent Circle, Cryptocat
- DNT in IE10
- Adblock/Adblock Plus, Privacy Badger, Disconnect.Me
- HTTPsEverywhere
- VPNs

Open source and trustworthiness
Theory of obfuscation: Types of disinformation [3]:

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

Some existing stuff and the things we can link it to later

- TrackMeNot generates plausible google searches (Chaff)
- FaceCloak? Encrypts facebook data
- CacheCloak - sends a range of plausible future paths to location based services (Palimpsestification)
- Shopping card loyalty swaps (Account Sharing)
- DuckDuckGo - mixing up user searchers (Account Sharing, no cleverness)
- CVDazzle

3. A selection of obfuscation strategies

Alexander's taxonomy [3] 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. The social aspect inherent to modern communication tools increases the [possibility of] interaction with others, which in turn increases the possibility of exposure

TODO: *what's that word i'm looking for?? detection/ yearbook-obscurity/DEF-Obfuscation.tex uncover / disclosure/ reveal*

of the lies. However, we can also use [the socially expected] social interaction to our advantage, colluding to strengthen the obfuscatory practices. yearbook-obscurity/DEF-Obfuscation.tex obfuscation.yearbook-obscurity/DEF-Obfuscation.tex

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 oustisde the digital sphere.

For each strategy we discuss: *i*) what kind of alteration of baseline data is performed; *ii*) what the motivation and possible use cases are; *iii*) how some form of computational support aids in the deception; *iv*) how the strategy can be applied to other data *v*) 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:

> 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.

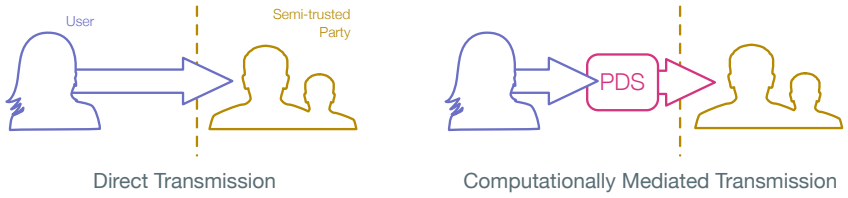

Figure 1. Models of interaction with semitrusted services. a) Direct transmission of information; b) computationally mediated transmission, where a personal data store is enlisted to aid in obfuscatory processes.

The service is hence *semi-trusted*: there are some benefits which the user wishes to accrue, but there are aspects of the service which makes the user unwilling to entrust their complete life 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[10].

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 typified, but not limited to mediation from a PDS 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[2]. 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. In all cases, a fictitious one-dimensional "location" measurement is plotted against time, to give a sense of how an individual's position in space changes. Figure 2a is the true baseline, with a curve indicating the continuous true position, and the dots representing reports of this position to the location-aware service. For each strategy we discuss: *i*) what kind of alteration of baseline data is performed; *ii*) what the motivation and possible use cases are; **TODO:** *I would split this in 2 - motivation, then use cases* *iii*) how some form of computational support aids in the deception; *iv*) how the strategy can be applied to other data **TODO:** *do we no do this in the next section for all strategies?* *v*) some of the systems which do this currently.

3.1.1. Chaff

World War II fighter planes would emit clouds of radar reflective sheets—*chaff*—which created multiple traces the screens of radar operators, and hence disguised

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

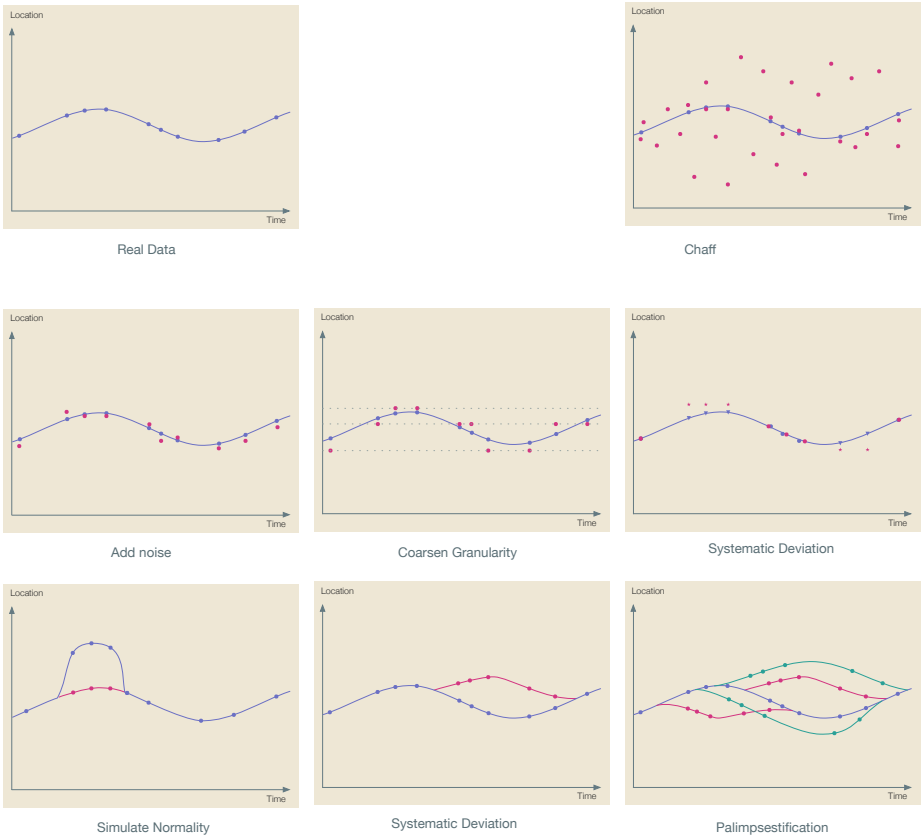

Figure 2. Obfuscation strategies for the lone agent

the true position of the aircraft. In a similar manner, we can add in multiple location datapoints alongside the real ones. This is the one of the few methods where the complete, accurate datastream 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, 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`; similarly, posted letters may be signed with a city rather than a street address.

3.1.4. Systematic Deviation

In some cases, it may be possible to introduce systematic deviations into the digital record. In order to 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. **TODO:** *examples?*

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].

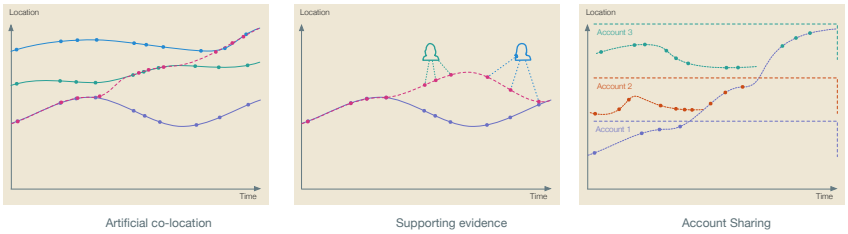

Figure 3. Multiplayer obfuscaction strategies: i) artificial co-location; ii) supporting information; iii) account sharing

3.1.6. Coherent Deviation

As the converse of simulating normality, the user may wish to pretend to be somewhere where they are not **TODO:** *more motivation for Alibot!* . 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 use 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. **TODO:** *examples*

3.1.7. Palimpsestification

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 to know which is the real one. This is similar to the strategy of CacheCloak **TODO:** *reference* , 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 tradeoff 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.

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 datastreams, or increase the veracity of artificially created data. Generally the possibilities in computational systems are analagous 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 - coherence is then required across multiple different accounts; however the counterpoint is that if successful, the obfuscation is better supported and harder to spot.

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 business of transferring location devices between accounts.

3.2.2. Supporting Evidence

TODO: *Maybe merge into preceding co-location section* 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 movies 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 [8], 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:

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[9]. 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 gameworld enforces consistency, and the community must produce coherent action streams in order to progress **TODO:** *cite Twitch!* .

Randomised schemes allocate accounts to people without a guiding principle; 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. **TODO:** *More explanation? Link to powerlevelling in WoW?*

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 spectrum of datatypes 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 **TODO:** *cite Nature article on cell location identification* .

4.2. PDSs to support obfusction

- create continuity
- act on your behalf

4.3. Personal Data Stores - allies on the intimacy battleground

Personal data stores (PDS) represent a partial solution to issue of presentation: having trusted, user controlled repositories for data enables a more user-centric approach to management of capta—those data which we choose to take and preserve. 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,

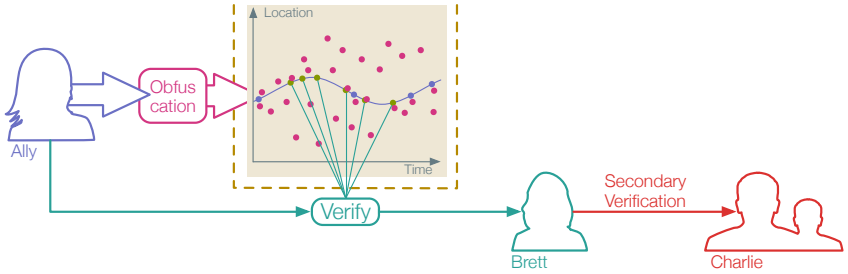

Figure 4. Example verification scenario. The user (Ally) provides a set of real data, plus *chaff* to a location aware service. A third party (Brett) then requests verification of some of the points, which Ally provides. Brett then wishes to share the data with Charlie, which requires Brett to verify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

nothing has been gained; hence the bridges become conduits for manipulating truth and constructing falsehoods. As personal data stores accumulate more real-time contextual data about the individual, as well as about the individuals social connections, PDSes 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)

4.4. Verification and provenance mechanisms

In the introduction we suggested that verification is a more nuanced mechanism than control over sharing, since sharing is impossible to control technologically. 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 datapoints 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 datapoints is mine", "these points are within 50m 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

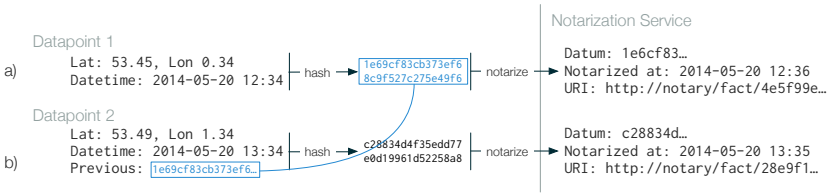

Figure 5. Notarization of personal data. a) Datapoints and times are hashed, and the values sent to a notary service, which provides a URL to verify that i) the given data was supplied and ii) when it was supplied. Hashes are used so that the data is not publicly shared. b) If the hash of the previous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

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 cryptographicaly 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.

4.4.1. Notarization

The verification examples above rely on Brett trusting Ally about which datapoints 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 notarization services can be employed[3]. 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,

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

if someone wants to make a prediction for the outcome of a football match, they could notarize 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, however, so this technique is most suitable when the range of possible things to notarize is so large as to make notarizing the entire space infeasible.

With regard to personal data, we can notarize 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 situtation as producing a notarized point does not prove that we were not anywhere else.

Notarization 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 notarized (Figure 5a). Additionally, points can be notarized 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 notarized results (Figure 5b).

5. Conclusion

- Translation to things that aren't location data; generalisability; can't add chaff to our bank accounts (or can we?)
- Viability - how do services react when we fill them full of noise? Plausible versions of these techniques
- Ethics - is this OK?
- Obfuscation evolves in lockstep with systems to see through it; future people will be better at spotting constructed points.
- In the short term services will start to become more suspicious about the data that goes into them; start rejecting points which represent causality violations.

References

[1]

[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] D. Beer. Power through the algorithm? Participatory web cultures and the technological unconscious. *New Media & Society*, 11(6):985–1002, Sept. 2009.

[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] D. Boyd. Networked Privacy. *Surveillance & Society*, 10(3):348–350, 2012.

[7] D. Brin. *The transparent society: Will technology force us to choose between privacy and freedom?* Basic Books, 1999.

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

[9] B. Dalton. Pseudonymity in social machines. In *WWW 2013 Companion*, pages 897–900, Rio de Janeiro, Brazil, 2013.

[10] 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*, 3:1376, Jan. 2013.

[11] M. Dodge and R. Kitchin. Codes of life: identification codes and the machine-readable world. *Environment and Planning D: Society and Space*, 23:851–881, 2005.

[12] 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):59–83, 2000.

[13] G. Goldberg. Rethinking the public/virtual sphere: The problem with participation. *New Media & Society*, 13(5):739–754, Dec. 2010.

[14] 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.

[15] 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*, pages 517–526. ACM, 2009.

[16] G. Iñiguez, T. Govezensky, R. Dunbar, K. Kaski, and R. A. Barrio. Effects of Deception in Social Networks. *arXiv preprint*, pages 1–19, 2014.

[17] D. Lyon. *Surveillance society: Monitoring everyday life*. McGraw-Hill International, 2001.

[18] A. Marwick. The Public Domain : Social Surveillance in Everyday Life. *Surveillance & Society*, 9(4):378–393, 2012.

[19] A. Narayanan and E. W. Felten. No silver bullet : De-identification still doesn ' t work. Technical report, 2014.

[20] A. Narayanan and V. Shmatikov. Robust De-anonymization of Large Sparse Datasets. In *IEEE Symposium on Security and Privacy*, pages 111–125, 2008.

[21] J. Rauhofer. Future-Proofing Privacy : Time For An Ethical Introspection ? *Surveillance & Society*, 10(2011):356–361, 2012.

[22] T. Reigeluth. Why data is not enough : Digital traces as control of self and self-control. *Surveillance & Society*, 12(2):243–254, 2014.

[23] B. Rooney. The Balance Between Open Data and Privacy. http://online.wsj.com/news/articles/SB10000872396390443884104577647600306243684#printMode, Sept. 2012.

[24] B. Simon. The return of panopticism: Supervision, subjection and the new surveillance. *Surveillance & Society*, 3(1):1–20, 2002.

[25] J. R. Whitson. Gaming the quantified self. *Surveillance & Society*, 11(1):163–176, 2013.