SN Computer Science (2020) 1:205
https://doi.org/10.1007/s42979-020-00206-y

ORIGINAL RESEARCH

Organisational Security Dependent on Individual Privacy

Vitalian Danciu[1]

Received: 15 April 2020 / Accepted: 25 May 2020 / Published online: 18 June 2020

Abstract
The global proliferation of cloud computing, smart homes, the internet of things and machine learning requires a novel view on the flow of confidential information and its classification. The security of an organisation is affected by the privacy enjoyed by its members. Sufficient data on those members can be leveraged in a so-called abduction attack aiming to extract confidential information from the organisation. To illustrate it we develop a model of actors and data flows and discuss three scenarios in which the confidentiality achievable by an organisation is limited by the privacy of its members. We support the model and the attack scenarios by reviews of the most prolific sources of personal data, its handling and its perceived value to the individuals it pertains to.

Introduction

The last decades have seen the proliferation of ubiquitous information services provided to the general public without direct financial cost to the individual user. The operators of these services routinely collect data about their users, with the intent to generate revenue, e.g. by serving targeted advertisements. The landscape of IT services is growing and includes services increasingly critical to the general public. This development has been accompanied by diverse malicious activities, including bulk unsolicited email (spam), theft of information, theft of identity, denial of service, defacing of web property and fraud. Such incursions are met with stepped-up security devices and policies.

Such instances of the *innovation–exploitation–reaction* process show that, as an information society, we are still in an exploratory "shake-down" phase, despite what the disruptive effect of the intensive introduction of information processing and digital communication on public and private life might suggest. In many countries, the reactions have been underpinned by laws, in an attempt to suppress attacks or privacy breaches originating in their legal domain. They include the prohibition to tamper, remotely, with others' IT infrastructure as well as rules and obligations with respect to the handling of personal data. However, there are also indications [26] for privacy regulations remaining unenforced for the benefit of data analysis revenue.

The meaning of the terms "security" and "privacy" have been consolidated during this period, driven by events as those described.

Information security concerns itself with the protection of stored data and communication processes. Its scope is the *"preservation of confidentiality, integrity and availability of information"*[4],[1] as well as authenticity, accountability, non-repudiation and reliability.

Privacy is a multi-faceted term (cf. e.g. [22]). It has been discussed long before the advent of computers in law research [27] as the rights of a person with respect to normally confidential information obtained without the person's consent. Modern interpretations of the term include rights even after consent has been given, in the sense of *data protection*, concerning itself also with the collection, processing and exchange of data pertaining to a (natural) person. In public discourse, the terms "security" and "privacy" are often found as contrary weights of the balance between the security of a country (claimed to require information about its citizens) and the privacy of the individual. Such arguments typically discuss the amount of privacy that is sacrificed with the aim of improving security.

This article is part of the topical collection "Future Data and Security Engineering 2019" guest edited by Tran Khanh Dang.

✉ Vitalian Danciu
danciu@nm.ifi.lmu.de
http://www.mnm-team.org/~danciu

[1] Institut für Informatik, Ludwig-Maximilians-Universität München, Oettingenstraße 67, 80538 Munich, Germany

[1] Numeric references to sections of the original text removed from quotation.

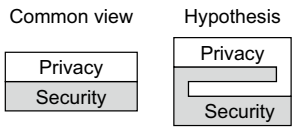

Fig. 1 Inter-reliant disciplines

Privacy has been argued to be important to society for reasons of the self-development of its individual members [9]. This reason is valid in the context of the formulation of law and the shepherding of culture. It does not seem compelling in a setting where benefit is a function of revenue or of services consumed. Therefore, in this paper, we explore a reason that might have more appeal in that setting, by illustrating how organisational security can be undermined by the choice to waive privacy.

Hypothesis

To control access to undisclosed, i.e. private information, information security techniques (specifically cryptography) are commonly used to protect data locally or during transmission. Therefore, privacy is often perceived by computer scientists as an application of information security to the personal sphere.

In this paper, which extends my work in [11], I argue that the reverse holds true also: that privacy of personal data is a prerequisite of effective information security. Figure 1 offers an illustration of these two views.

In particular, we explore how the security of an organisation is limited by the privacy of its members, hence: the security of the organisation as an application of the privacy of the individual.

Synopsis

The intention of this paper is to raise awareness for attacks on the confidentiality of the information held by an organisation by means of indirectly observing the behaviour of its members.

The background of the hypothesis, that security is contingent on privacy, is the plethora of personal data being collected and traded. We examine the reasons for data collection in Section "Background: A Situation Analysis" and establish that detailed data on the individuals and their behaviour is available. To lead up to examples on how an organisation's confidentiality can be compromised, we formulate a model in Section "Model". The model reduces real-world complexity to a small set of actors, their inter-relationships and the constraints under which they act. We introduce in Section "Abductive Attack Outline" three simple scenarios in which an organisation wishes to retain secrecy of some internal information while an adversary would try to acquire this information. For each, we outline how the desired information can be inferred from legally available data. We discuss the implication of this ability and a selection of countermeasures in Section "Discussion" before summarising our observations in Section "Conclusion".

Background: A Situation Analysis

The abduction attack described in this paper is contingent on the availability of personal data pertaining to members of the organisation targeted by the attack. Thus, in this section we establish this availability by examining the data available for collection and the limitations imposed on the collection.

Data Proliferation

Despite the public outrage following some of the leaks of personal data in the last years, the privacy policies of many corporations offering "free" services assert that collected data may be passed on to third parties. Thus, the public outcries appear to be more motivated by the use that data have been put to, such as in the Cambridge Analytica affair [8], than the fact that it has been processed by third parties.

Our interaction with services, appliances and with each other in the course of everyday life is subject to recording by the providers of services, the producers and maintainers of appliances and the communications systems and services we employ.

Data sources

Proliferation of data sources in the personal and utility domains of a person is ubiquitous, due to the pervasive use of personal network devices (phones, tablets, wearables, etc.) and the proliferation of "smart" appliances connected by the Internet of Things. The privacy implications of this emerging network have been discussed [1–3, 7] but mostly in the sense of requiring consent from the users whose privacy may be implicated by the dissemination of the data.

The data potentially collectable by different devices that we use in daily life either in the personal or the utility domain. The data collection within the private domain is triggered explicitly, and the user has some control over it. Table 1 describes exemplary data sources from this domain.

A device in the utility domain is under the control of a provider (e.g. the electricity grid operator): data collection may trigger implicitly, without the user interacting directly with the device. Examples for such devices are given in Table 2. The tables list examples of a device class, the kinds of data collectable by an operator/maintainer of the device and the information potentially inferable about the device user.

Table 1 Exemplary data sources within the personal domain of an individual

Class	Example	Data kind	Inferable information
Portable	Smartphone, Smart watch, GPS navigation	Contacts, location, speed, information access, communication events, communication content, time of events	Home, work and other frequented addresses, daily schedule, entertainment preferences, travel, social circle, health concerns
Appliance	TV, dishwasher, washing machine, air conditioner, vacuum robot	Usage, program	Current occupation/activity, habits, entertainment interest, home layout
Voice activation	Echo and similar, toy/baby-phone	Ambient audio, speech recognition result	Environment audio, voice print, number of persons present, daily schedule, queries

Table 2 Exemplary data sources within the utility domain of an individual

Class	Example	Data kind	Inferable information
Smart home	Lights, blinds, temperature, locks	Presence, environment configuration	Current occupation/action, daily schedule
Smart grid	Electricity, gas, water	Usage events, usage volume, concurrent usage	Operated device, persons in household, daily schedule, current occupation/action
Payment	Credit card, online payment, phone payment	Time, volume, recipient, credit	Economic status, shopping habit, gender, age, location/travel
Private transport	Emergency locator, diagnostics	Location, usage	Itinerary, occupation, travel habits
Public transport	Personal ticket, face recognition	Journey events	Itinerary, home and work area
Public space security	Face recognition, radio fingerprinting	Enter/exit detection zone	Itinerary, home and work area

A comprehensive collection of data sources would necessarily include as a focus common information services. Communication services (email, voice/video call, chat) and social network services are well known to handle and collect data about their users. Indirect exposure is inherent to many services. Person A posting a photo of Person B on social media, or sending Person B an email from a "free" mail service divulges to those service providers information about Person B without his knowledge or consent.

Economy of Indirect Exchange

The collection of personal data requires in many legislations the consent of the person it belongs to. We will review how consent is obtained from a perspective of economic exchange.

Figure 2 shows three different exchange scenarios. The traditional exchange is that a person will pay for both services used and goods purchased, as shown in the uppermost part of the diagram.

Since many years now, information services are being provided to users without direct financial compensation. It is also common that the provider requires consent for the use of data in exchange for the use of the services, including the data processed by the service (e.g. messages) and the metadata (addresses, interactions with the user interface). Service

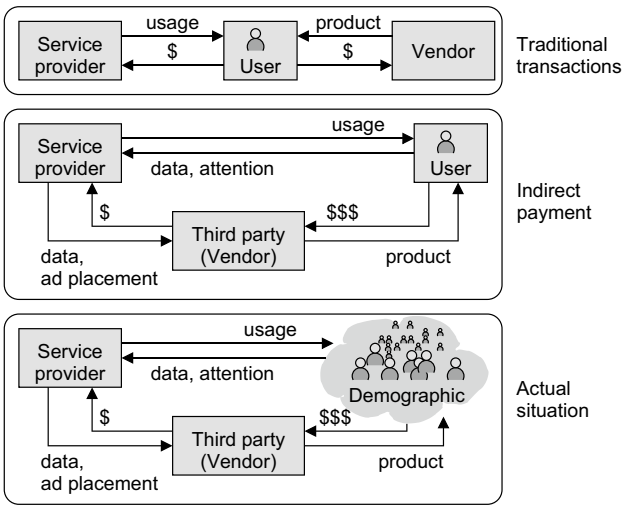

Fig. 2 Exchange of services and compensation

providers generate revenue by "harvesting" attention (for advertising to the general public) and data (for sale or for advertising to a specific group or to the individual). Vendors leverage this attention to place advertisements targeted by (profile) data about the user.

The connection between service use and indirect payment is dissolved when it becomes unimportant if a certain user

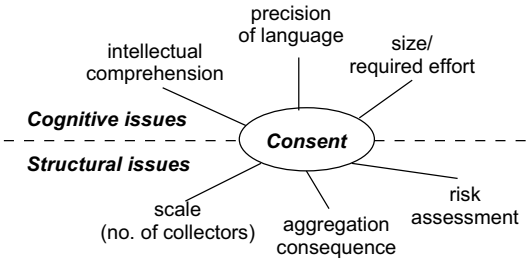

Fig. 3 Reasons for unwilling or uninformed consent according to Solove [23]

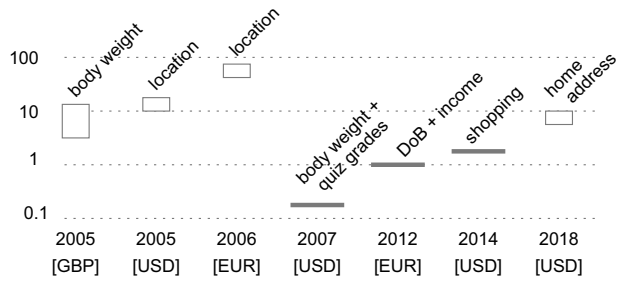

Fig. 4 Value of privacy

will spend his money on the advertised goods, as long as a sufficient number of users from the targeted demographic do, as suggested in the lower part of Fig. 2. The strong growth of the "Internet companies" suggests that the combined worth of attention and personal data is much higher than the cost to provide and deliver the services.

Besides targeted advertising, the personal data can be leveraged to infer financial, risk, health and other information about a person. Such information can be employed outside the information services. Despite these issues, customers continue to agree to the exchange because (a) the cost to them is hidden ("just ads"), (b) they perceive that the cost is to a group and not personal c) the services have become a necessity and (d) they are uneducated with respect to alternative, less invasive services. In the case of devices, the use of the device may necessitate consent to the collection of data by an associated online service. We review in the following the issues of consent (Section "Consent") and the valuation of personal data by the owner himself (Section "Valuation of Personal Data").

Consent

Solove [23] discusses the reasons for why consent is currently ineffective as an instrument for privacy. He identifies the conceptual and structural issues illustrated in Fig. 3, which appear consistent with the results of surveys on privacy administered to members of the general public.

The cognitive issues reflect the difficulty to read and understand ("comprehension") a frequently large ("size/required effort") body of text. This text may be written in the style of a legal text and be intended to cover a multitude of cases for present and future data use, resulting in vague and general statements about what data may be collected and how it might be treated subsequently ("precision of language").

The structural issues include the difficulty for the individual to manage the agreements with a large number of data collectors, which exacerbates the effort required to manage consent (grant, withhold or withdraw) based on informed decisions.

The remaining two structural issues reflect the lack of insight on the part of the data owners with respect to the aggregation of multiple data sets on one hand and regarding the risk to them if isolated or aggregated items of personal data are collected and disseminated. The former would require knowledge about data analysis techniques that allow the extraction of information beyond the literal data about a person. The latter would require knowledge about the leverage afforded to entities acquiring such data with respect to the data owner. Hence, in both cases the decision on consent would need to be informed on the basis of specific technical knowledge that is not available to the average data owner. It seems probable that these issues in particular lead to a low valuation of personal data.

Valuation of Personal Data

Figure 4 shows the approximate range of bids accepted by survey participants for different kinds of personal data: The studies show disclosure of location information for around 10 GBP [12], disclosure of location information (1 month) for 1–150 EUR, with 10% of the bids below 1 EUR [10], and willingness to have their quiz results and body weight disclosed for a benefit of 0.25 USD [17], disclosure of date-of-birth and income for a benefit of 1 EUR [6], disclosure of shopping choices for 2 USD [5], disclosure of the date-of-birth and body weight for bids of USD 6–12 [18], sale of private data attributes for bids around 10 USD [25].

It is fair to note that, the studies that produced the data all had different methods and sample sizes, they were conducted in different countries and in different peer groups. While the comparison of numerical results is awkward, their collection does allow a tentative quantification for the monetary privacy valuation of individuals. These numbers suggest a rather high willingness to exchange personal data for some benefit, while the price/benefit extracted by the (former) owner of the personal data remains rather low. Despite privacy incidents, this relationship appears to persist over time.

We can conclude that, as a rule, consent to use and disseminate private data is given due to its low valuation compared to the high cost of refusing the use of common services.

De-anonymisation

Collectors and brokers of data may be required to anonymise the data they collect, thus disconnecting it from the known identity of the user. However, de-anonymisation has been shown to be effective with only few data points, e.g. for credit card data [13], for web browsing data [24], and for video-on-demand ratings [21]. In all these cases, de-anonymisation was performed by using multiple data sets with an overlap of attributes and reasoning over the cases where the overlapping attributes matched. Zhou et al conclude [28] that anonymisation techniques for social network data should be developed reactively to be effective. Such findings suggest that mandatory anonymisation may increase the effort of the adversary but not obviate the achievement of his goal.

Behaviour interpretation

The behaviour of people can be observed indirectly by evaluating effects in their environment. Data from utility sensors as well as communication meta-data represent such effects. It can be procured *passively*, by trading for data already in the market or *actively*, by stimulation. Active acquisition entails pro-active requests for consent to a candidate group of individuals, observation of the reaction of users to time- or place-constrained prompts/offers, etc.

Summary

The findings from this section can be summarised as follows:

- There is a huge and growing number of data sources.
- Individuals perceive the value of their personal data as low, compared to the services they take.
- Individuals are under pressure to use IT services (peer pressure, perceived lack of alternatives).
- Individuals default to granting consent.
- Data collector and service provider roles (often the same entity) profit from this situation, which motivates them to maintain it.
- Re-identifying single persons from anonymised data sets is possible if the data sets are combined.

From this description of the environment, it is possible to create a model centred on the data owner.

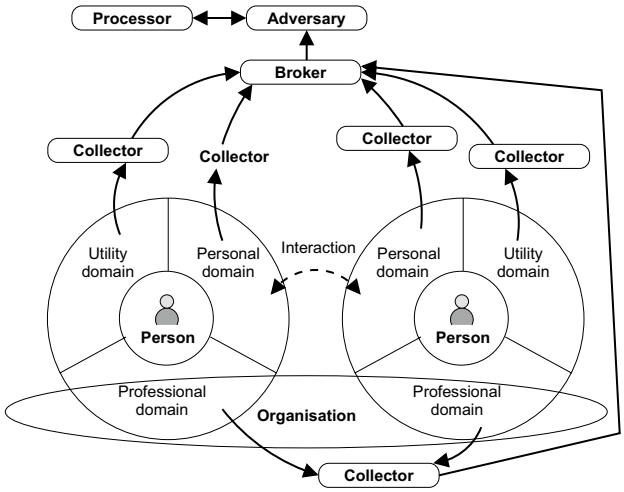

Fig. 5 Actors, domains and data flow

Model

We develop a model consisting of definitions of roles, their relationships, as well as the domains of data relevant to individuals. Within the model, we formulate a number of plausible assumptions as to the behaviour of the different roles. The purpose of the model is on the one hand to constrain the complexity of the real world to a manageable level (while retaining plausibility) and on the other hand to establish a terminology for use in the remainder of the paper.

Actors

We identify the following six roles, sketched in Fig. 5 as relevant to our analysis.

The Organisation is a company or other organisation that wishes to keep its internal information a secret.

The Adversary wishes to acquire some secret information of the organisation. The adversary cannot be a member of the organisation, nor the organisation itself. The adversary may be an organisation in the general sense, a group of people or a single individual.

A Person is a member of the organisation (e.g. an employee of a company) that works on site but does not reside on site .The person uses common information services to reduce chores, make payments, participate in social life, for entertainment and so on. Persons are the sources of personal data.

A Collector directly collects data from persons. Collectors include providers of pure information services (electronic content providers, search engines, communication

applications), network service providers, utility providers (electricity, heating, gas), vendors of IoT-based devices but also vendors of enhanced traditional products (cars, domestic appliances) and financial companies.

A Processor processes data for further use. Processing includes aggregation, profiling, indexing, tracking of use and analysis for a specific purpose (e.g. the assessment of usage of a service over time).

A Broker or Trader facilitates the exchange of data sets and the sale of data sets for profit.

Domains

We differentiate between three disjunct domains, which denote authority over their own set of functions and the data they process: the professional, personal and utility domain. The domains structure the devices employed by persons, allowing them to generate data within the professional and personal domains and to trigger data generation within the utility domain.

The professional domain is that in which the person acts on behalf of the organisation. It encompasses a person's interaction with communication and information systems provided by the organisation and using utilities at the organisation's site (e.g. network connectivity).

The personal domain encompasses the functions that pertain to the private sphere of the person. That includes private communication, entertainment, private transport, home automation.

The utility domain encompasses the person's use of devices whose function requires information services outside the personal domain and utility services (network connectivity, electricity, gas, public transportation, home delivery services, etc.), which may generate data collectable by the respective utility provider.

Abductive attack outline

Data collected from members of the organisation may be employed by the adversary to discover information confidential to the organisation. The adversary uses the data available to infer the confidential information of the organisation. We call this an *abductive attack*, since the adversary's result is acquired by the abductive inference and constitutes a best interpretation that cannot be verified without knowledge of ground truth. In this section, we examine a number of such scenarios as thought experiments.

For a successful abduction attack, the following are necessary:

1. *A goal* a specification for the information that is normally confidential to the victim and that the adversary attempts to acquire
2. *A correlation model* the information may be inferred from data not directly related to it. The model specifies the relationships between different data, that allow to infer intermediary information leading to the goal, or that allow to derive the goal itself. For example, the model may be chosen as a Bayes Network, i.e. a directed acyclic graph with nodes representing the incidence of different kinds of data and links denoting their relationship.
3. Observation data satisfying all the kinds of data required by the model.
4. An inference procedure that operates on the model.

Assumptions

Given the model described in the previous section we make the following assumptions about the behaviour of the model roles and their environment.

Times and places We assume that the organisation operates a single site and a single network. Constraining the model simplifies the demonstration of the analysis without invalidating its approach in principle. At any given point in time, a person is located:

- at work, at the organisation site
- at their home
- en-route
- at a third location from a (small) set (e.g. store, gym, restaurant) that are visited regularly.

We differentiate between working hours, spent at the organisation's site and leisure time, spent at home or at a third location.

Fair play We assume that, laws and policies are respected by all roles. While this may not always be the case in reality, the purpose of our analysis is to show how confidentiality may be broken without resorting to illegal means. Thus, we assume the following:

- Data are being collected, processed and traded with consent.
- Collectors perform (pseudo-)anonymisation if required by law.

- Organisation policy prohibits their members to disseminate information pertaining directly to the organisation's concerns.
- The adversary refrains from any direct attack on the organisation, including direct social engineering targeting single persons.

Data availability We have established in Section Data Proliferation the existence of a rich, growing and free data market. Hence, we assume that exchange of data and remuneration is not technically limited, i.e. if a data point is collectable, then it is available for analysis.

Example scenarios

Consider the following three scenarios in which the adversary attempts to acquire knowledge about undisclosed, internal aspects of an organisation. For each scenario we develop a procedure usable by the adversary to acquire the desired information starting only with the information presented in the respective scenario in addition organisation's physical location and its network address range. All scenarios are framed within the bounds of the model and constitute cases in which the adversary is apt to break the confidentiality of the organisation.

Current production load. A production site, a factory, wants to keep confidential the production load from their (overseas) customers. During times of high production load, the factory employs additional workers in order to fill three shifts. Disclosure of the production load may have influence on the company's stock value or lead to the loss of potential contracts, should the apparent situation contradicts the available capacity advertised to customers.

Persons who travel to the organisation's site and arrive there within a time interval, then leave the organisation's site approximately eight hours later, may be workers. The number of time intervals for arrival should indicate to the adversary the number of shifts currently employed by the organisation.

Data items, that can indicate arrivals, include

- Location data from phones in proximity of the site
- The amount of traffic issued from the organisation's network to sign-in services
- The number of passengers on collective transport, which serves a station near the site
- The number of vehicles in the site's parking
- The number of transactions at nearby shops, which may be visited by employees before or after work
- The number of transactions at vending machines at the site

We note that the extraction of single identities is not strictly necessary for the task. The adversary may conclude the desired information by studying the variance of one of the data items and increase the certainty of the result by studying two or more.

It is also important to note that, in this special case, the adversary can attain his goal without resorting to the use of personal data. This suggests that the privacy of utilities and devices may be of importance to security considerations even when a direct connection to single persons' behaviour can be excluded.

Time of deployment. A military unit in peacetime wants to keep confidential their schedule of deployment to manoeuvres or exercises. Between exercises, personnel are authorised to leave the unit's site during leisure and during the night. The impact on the organisation, if the schedule is disclosed, includes a bad performance in the exercise or disturbance by observers from the general public.

The adversary can predict deployments in the short term by inferring them from a change in the behaviour of the members of the unit. For example, an exercise starting in the early morning would prompt members of the unit to either stay on site over the night or to rise extra early. Both behaviour patterns are indirectly observable but require the creation of unique identities of the organisation's members in order to distinguish instances of one pattern from the other. For each digital identity the adversary determines the following data items:

- *home location*, which can be acquired from location data or address data or utility providers
- duration of the person's *time en-route* between *home* and the organisation's site, which can be approximated by plotting the route between the two locations
- *regular home departure time*, i.e. the regular time of the day when the person prepares to travel from home, that may be mapped by an interval of activity followed by the cessation of events at the home location by observing utility use, network traffic. It can be determined more directly, from sensor data (e.g. increased use of appliances) or, if available, from location updates.
- *home arrival time*, i.e. the point in time when a person arrives at the home location. That point in time can be determined by the same means as the home departure time.

In an instance where the regular time for travel changes for one person, we can infer that, the person is due to arrive at the organisation's site after the person's typical *en-route time*. An accumulation of instances of this kind would predict an imminent deployment to the adversary. This pattern may be preceded by the observation that some of the

organisation's members failed to arrive at home the day before, in contrast to their regular behaviour. If timely prediction is important to the adversary, these latter instances may be exploited as first signs of an imminent event. Prediction accuracy might be improved by accounting for regular visits to *away* locations, regular variations in the persons' respective schedules

We note that in isolation, the data items correlated by the adversary are innocuous but become instrumental when combined with a correct assumption about a mechanism (in this case: commuting between home and the organisation's site).

Identities of research personnel.

A research company wishes to keep confidential the exact identity of its research employees, in order to avoid their recruiting by competitors. The company files applications for public research grants when offers of grants are published. Before application deadlines, research personnel works longer hours in order to complete the proposal documents.

The adversary formulates the conjecture that those persons whose working hours change in proximity of a deadline are the targets of recruiting by competitors. Their identification makes use of data items and correlations already described in the previous two scenario examples. The difference in this scenario is the need for the extraction of actual identities.

Using the publicly available time-frames of grant applications, the adversary can select a set of candidates, as in the "Time of deployment" scenario, e.g. by observing commute times. The adversary needs to attempt to de-anonymise each of the persons in the candidate set.

Discussion

The scenarios described in the previous section have sketched the use of seemingly harmless personal and non-personal data produced by persons within the organisation to deduct confidential information items of the organisation.

Properties

The *abductive attack* has a number of interesting properties. It is

1. performed within the boundaries of the law. The data exchange and processing employed are legal, and the methods for inferring the desired information points are similar or identical to those used in academic research, or in the prosaic activity of a private detective.
2. indirect, as the organisation itself is never targeted directly.
3. undetectable during initiation and execution and virtually untraceable post mortem. Even knowledge of the adversary's actions does not immediately enable the organisation to conclude that it is being targeted.
4. executable by the adversary remotely, from another legal domain, provided data trading and processing by third parties is legal and available at the adversary's location.
5. executable without special authority or knowledge by using the *broker* and *processor* roles. The adversary needs neither authority for collecting data nor expertise in finding and correlating it.

Limitations

Not every kind of secret can be acquired by the means illustrated in the scenarios. To be effective, the approach must tie in to the behaviour of persons observable in their personal and utility domains.

The method relies on the correct formulation of a theory by the adversary, which allows interpretation of the available data by abduction. Even the theories in the examples in Section "Example scenarios", simple and plausible as they seem, may not hold for every production site, military unit or research lab. Therefore, a targeted attack seems difficult to automate fully, as it requires the judgement of a human mind. It might be possible, however, to increase the level of automation by recording patterns of judgement.

Finally, the accuracy of the information gained cannot be ascertained without at least verifying several instances of the same case. Obviously, this property originates in the abductive nature of the process.

Countermeasures

The abductive attack seems impossible if the data produced by all persons associated with an organisation are insufficient to derive with any useful probability a given confidential item of information that the organisation wishes to safeguard. Since an assurance of this state is implausible, the organisation and its members should aim to limit the data that is useful to the adversary.

Differential privacy

Similar problems have been studied in the context of statistical databases, to address combinations of queries that, while harmless alone, would allow the inference of privileged information if their results were combined. Dwork introduces the concept of *differential privacy* to address this issue in [14, 16]. This line of study is valuable and is being

applied in privacy audit systems (e.g. [19]). The introduction of noise in the data has been shown [15] to suppress the exposure of private information.

The structural dissimilarity of databases with the global data market seems to obviate the introduction of such measures in the cases described in this paper. While privacy is being pursued by the operator of the database for the benefit of the persons registered therein, there is neither a single operator for the data market, nor is there an incentive for the operators that do exist, to act.

Policy

Suppressing abductive attacks by policy seems difficult: policy makers would have to differentiate between inference as used in any research and that performed with a malicious intent. Scholars have proposed granting rights to the so-called *ad hoc groups* [20] that are assembled through analysis and classification of individuals' behaviour. The members of an ad hoc group may correspond to the members of an organisation.

Self-defense Organisations may wish to guard themselves against instances of abductive attacks that may constitute industrial espionage. However, interfering with the lives of their members beyond the professional domain of our model may prove difficult for all but very few organisations. Incentives for the protection of personal data could be given by

- subsidising paid, anonymous services for the benefit of its members—to avoid them using the "free" ones,
- guiding in the choice of devices and services and
- raising awareness for the security risk to the organisation.

Conclusion

Analysis of personal data, like any conceptual tool, may be used for society's benefit, for example in medical and pharmaceutical research or reputation systems. However, a rich data market paired with automated data analysis render possible the inference of confidential information of an organisation from the personal and utility data produced by its members.

Corporations, i.e. commercial organisations may wish to guard confidential information. Especially during strategic restructuring (merger, acquisition) even small items of information being disclosed may have a large impact. The protections afforded individual persons by law do not extend to organisations. Hence, the victim organisations are third parties to an exchange between their members and the data industry.

It is important to note that today the process of acquisition of confidential data can be automated to a high degree thus lowering both the financial and the time cost for the procedure. If the reward for the adversary justifies the cost of obtaining and processing the necessary data, the subversion of an organisation's confidentiality becomes possible without substantial legal risk or risk of detection.

Acknowledgments Open Access funding provided by Projekt DEAL.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

References

1. Comments of the Electronic Privacy Information Center to the Federal Trade Commission on privacy and security implications of the Internet of Things. Published comments, Federal Trade Commission (FTC), 2013.
2. Internet of Things—privacy & security in a connected world. FTC Staff Report, Federal Trade Commission, January 2015.
3. Internet of Things—status and implications of an increasingly connected world. Report to congressional requesters, United States Government Accountability Office Center for Science, Technology, and Engineering, May 2017.
4. Information technology–security techniques–information security management systems–overview and vocabulary. International Standard ISO/IEC 27000:2018, International Organization for Standardization, 2018.
5. Acquisti A, John LK, George L. What is privacy worth? J Legal Stud. 2013;42(2):249–74.
6. Beresford A, Kübler D, Preibusch S. Unwillingness to pay for privacy: a field experiment. Econ Lett. 2012;117:25–7.
7. Boeckl K, Boeckl K, Fagan M, Fisher W, Lefkovitz N, Megas KN, Nadeau E, O'Rourke DG, Piccarreta B, Scarfone K. Considerations for managing Internet of Things (IoT) cybersecurity and privacy risks. Draft NISTIR 8228, National Institute of Standards and Technology (NIST), 2018
8. Cadwalladr C, Graham-Harrison E. 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. The Guardian. London, United Kingdom: Guardian Media Group; 2018. https://www.theguardian.com/news/2018/mar/17/cambridgeanalytica-facebook-influence-us-election
9. Cohen Julie E. Configuring the networked self: law, code, and the play of everyday practice. New Haven: Yale University Press; 2012.
10. Cvrcek D, Kumpost M, Matyas V, Danezis G. A study on the value of location privacy. 2006; 109–118, 01

11. Danciu V. Individual privacy supporting organisational security. In: Dang TK, Küng J, Takizawa M, Bui SH, editors. Future data and security engineering. Cham: Springer International Publishing; 2019. p. 3–14.
12. Danezis G, Lewis S, Anderson R. How much is location privacy worth? In: Proceedings of the Workshop on the Economics of Information Security Series (WEIS), 2005
13. de Montjoye Y-A, Hidalgo CA, Verleysen M, Blondel VD. Unique in the crowd: the privacy bounds of human mobility. Sci Rep. 2016;3:1376.
14. Dwork C. Differential privacy. In: 33rd International Colloquium on Automata, Languages and Programming, part II (ICALP 2006), volume 4052 of Lecture Notes in Computer Science, pp 1–12. Springer Verlag, July 2006
15. Dwork C, McSherry F, Nissim K, Smith A. Calibrating noise to sensitivity in private data analysis. J Privacy Conf. 2017;7(3):17–51.
16. Dwork C, Naor M. On the difficulties of disclosure prevention in statistical databases or the case for differential privacy. J Privacy Conf. 2010;2(1):93–107.
17. Grossklags J, Acquisti A. When 25 cents is too much: an experiment on willingness-to-sell and willingness-to-protect personal information. In: Workshop on the Economics of Information Security, 2007
18. Huberman BA, Adar E, Fine R. Valuating privacy. In: IEEE Security and Privacy, 2005
19. Lu H, Li Y, Vaidya J, Atluri V. An efficient online auditing approach to limit private data disclosure. In: 12th International Conference on Extending Database Technology (EDBT). Research Collection School Of Information Systems, 2009.
20. Mittelstadt Brent. From individual to group privacy in big data analytics. Philos Technol. 2017;30:475–94.
21. Narayanan Shmatikov Vi. Robust de-anonymization of large sparse datasets. In: Proceedings of the 2008 IEEE Symposium on Security and Privacy, pp 111–125, Washington, DC, USA. IEEE Computer Society. 2008
22. Renaud K, Galvez-Cruz D. Privacy: aspects, definitions and a multi-faceted privacy preservation approach. In: Proceedings of the 2010 Information Security for South Africa Conference, ISSA 2010, 2010; pp 1–8, 09
23. Daniel JS. Privacy self-management and the consent dilemma. Harv Law Rev. 2013;126(7):1880–903.
24. Su J, Shukla A, Goel S, Narayanan A. De-anonymizing web browsing data with social networks. In: Proceedings of the 26th International Conference on World Wide Web, WWW '17, pp 1261–1269, Republic and Canton of Geneva, Switzerland. International World Wide Web Conferences Steering Committee. 2017
25. Tan J, Sharif M, Bhagavatula S, Beckerle M, Mazurek M, Bauer L. Comparing hypothetical and realistic privacy valuations. In: WPES'18: Proceedings of the 2018 Workshop on Privacy in the Electronic Society; ACM. pp. 168–182.
26. Vinocur N. How one country blocks the world on data privacy. Politico. 2019. https://www.politico.com/story/2019/04/24/ireland-data-privacy-1270123
27. Warren SD, Brandeis LD. The right to privacy. Harv Law Rev. 1890;4(5):193–220.
28. Zhou Bin, Pei Jian, Luk WoShun. A brief survey on anonymization techniques for privacy preserving publishing of social network data. SIGKDD Explor Newslett. 2008;10(2):12–22.