Privacy Risk Assessment: From Art to Science, By Metrics

Isabel Wagner and Eerke Boiten

Cyber Technology Institute,
De Montfort University, Leicester, UK,
{isabel.wagner, eerke.boiten}@dmu.ac.uk

Abstract. Privacy risk assessments aim to analyze and quantify the privacy risks associated with new systems. As such, they are critically important in ensuring that adequate privacy protections for individual users are built in. However, current methods to quantify privacy risk rely heavily on experienced analysts who pick the "correct" risk level on a five-point scale. In this paper, we argue that a more scientific quantification of privacy risk increases accuracy and reliability and can thus make it easier to build privacy-friendly systems. We discuss how the impact and likelihood of privacy violations can be quantified and stress the importance of meaningful units of measurement. Finally, we argue that privacy risk metrics should be expressed as distributions instead of average values.

Keywords: privacy risk metrics, privacy impact assessment

1 Introduction

A privacy impact assessment (PIA) is a process of identifying and mitigating privacy risks in real or planned systems. During a privacy impact assessment, organizations identify possible privacy risks, then quantify and rank these risks, and finally take decisions on whether and how to reduce, remove, transfer, or accept the risks. "PIA" also refers to the document produced in this process, and as such it is generally seen as a *living* document in software development – because privacy risks can change as a consequence of design and implementation.

PIAs are an essential component of Privacy by Design [4], an approach to dealing with privacy in a proactive rather than reactive way. They have been recommended by national data protection authorities for more than 5 years already [8, 5]. In the new European data protection regulation [1], PIAs (called "data protection impact assessments") are mandated for some cases, including surveillance, data sharing, and new technologies. As PIAs include consultation with stakeholders, they are also a useful mechanism for obtaining their buy-in in what might otherwise be seen as "risky" data processing processes.

However, the wider application of PIAs may be limited because privacy risk assessments currently rely heavily on experience, analogy and imagination, that is, risk assessment more closely resembles an *art* than a science. We argue that a

more scientific approach to risk assessment can improve the outcomes of privacy impact assessments by making them more consistent and systematic.

Contributions. In this paper, we investigate how to quantify privacy risk systematically with the aim of moving privacy risk assessment from being an art closer to being a science. We focus on *data* driven privacy (i.e. the impact of data decisions, possibly outside the data sphere) because this is the scope of the GDPR, currently the strongest driver of PIAs. In line with the common decomposition of risk into impact and likelihood, we discuss quantification of impact and likelihood separately and suggest possible metrics for each (Sections 3 and 4). We then discuss how metrics for impact and likelihood can be combined to form privacy risk metrics that can be used directly in privacy impact assessments and privacy requirements engineering (Section 5). Finally, we highlight open issues in the area of privacy risk quantification.

2 Why Quantify Privacy Risk?

Before discussing the benefits and building blocks of a more scientific method for quantifying privacy risk, we briefly describe the state of the art in risk assessment and privacy metrics.

Risk assessment. Risk is commonly calculated as some function of likelihood and impact. Several proposals exist to determine the risk of security threats, for example the NIST guidelines [10] or the OWASP Risk Rating Methodology. These are often cited in the privacy literature because security risks can be quite close to privacy risks. Both methods measure impact and likelihood on Likert scales, e.g. from "very low" to "very high", with no clear guidelines on how to determine the position on this scale. For example, the NIST guidelines [10] list examples of adverse impacts, such as harm to operations, assets, or individuals, and explain how the expected extent of each impact should be mapped to the Likert scale: "significant" financial loss, for example, is a moderate impact, while "major" financial loss is high impact. Likelihood is split into the likelihood that a threat event occurs, and the likelihood that an adverse impact results. The ratings for likelihood and impact are then combined according to a table that indicates the resulting risk rating for each combination of the separate Likert scale ratings. For example, "low" impact and "very high" likelihood result in a "low" overall risk. These impact and likelihood ratings may be rated differently by different people, and the resulting risk ratings may not be accurate or reliable.

Privacy metrics. The privacy metrics that have been proposed in the literature [13] focus mostly on measuring the amount of privacy that a privacy enhancing technology (PET) can provide against some adversary, for example expressed as the adversary's error, uncertainty, or information gain. Some privacy metrics focus on the adversary's success rate and may thus be suitable to quantify the likelihood of a privacy violation (see Section 4). Some privacy metrics measure risk directly, for example, the privacy score in social networks [9] is computed as the sensitivity of profile items multiplied by their visibility.

Benefits of more scientific risk quantification. The accuracy and reliability achievable through a more scientific and systematic way of measuring privacy risk would allow for several important benefits.

When building new systems, risk metrics could allow to compare the risks associated with different ways of building the system. In particular, for systems that are composed of smaller building blocks, risk could be measured on the level of building blocks, and composition rules would allow to compute the overall risk. In effect, these risk metrics allow to rationalize and substantiate decisions about how systems that affect privacy are built and evaluated.

Risk metrics are also needed in privacy requirements engineering [6], which is a similar process to privacy impact assessment (PIA), but with the goal of deriving formal privacy requirements and identifying suitable protections in the form of privacy-enhancing technologies. The privacy requirements engineering process can identify many risks and thus needs a way to prioritize risks. For example, the LINDDUN method [6] uses risk scores, but does not state specifically how these scores should be determined. In this context, risk metrics can also allow to set thresholds for risks that need to be addressed.

Finally, companies that offer cyber insurance benefit from accurate risk metrics to correctly determine insurance premiums.

Building blocks for a more scientific risk quantification. An important foundation of a more scientific approach is the ability to measure and predict the relevant quantities, i.e. the risk of a privacy violation which may be expressed as some combination of likelihood and impact of privacy violations.

In addition, units of measurement are important to make risk metrics more understandable and manageable. Risks measured using the same unit can be meaningfully aggregated (e.g. computing the total privacy risk from contributing risk factors by adding them) and directly compared (when considering different technical alternatives, or when prioritising risks). When units differ, such operations become more difficult or fundamentally dubious.

In business, financial value may be acceptable as the ultimate unit which is used to quantify even human lives and reputation, but the public sector may prefer units that more closely relate to the concept of privacy risk.

3 Impact Quantification

To make the measurement of privacy impact more systematic, privacy impact metrics should be based on four key components. First, the number of people potentially affected by a privacy violation indicates the scale of the problem. Everything else being equal, a violation that affects one person is less severe than one that affects a hundred. This component is relatively easy to quantify, and is widely reported in the news when privacy breaches become public.

Second, the sensitivity of the affected data indicates the type and extent of possible harm to individuals. The sensitivity of data is not necessarily aligned with the GDPR's categories of *personal data* and *special category data* – credit card data are classified as personal data, but can incur direct financial harm,

whereas trade union membership is classified as special category data, but its exposure would not be seen as harmful in many countries. Importantly, if the privacy of more than one type of data is breached, then the overall sensitivity may be higher than a linear combination of individual sensitivities. For example, spatio-temporal data (e.g. from cell towers) reveal a person's location; CCTV images reveal appearance without necessarily identifying individuals; and the combination of both additionally allows to identify individuals in CCTV images. The sensitivity of data is thus difficult to quantify. In a first approximation, metrics from information theory could be used to measure the amount of information (in bits) revealed by a privacy breach, even though *amount* does not fully coincide with *sensitivity*. Another approach that is useful when users can choose their individual privacy settings is to compute sensitivity from the privacy settings of a large number of users [9].

Third, the expectation individuals have of how their data will be treated, and how much a privacy violation deviates from this expectation, indicates as how "creepy" a privacy violation will be perceived. For example, the leak of electronic health records from a third party server located in a foreign country would be quite unexpected because people may not expect that the storage of health records is outsourced abroad. Depending on social norms, there may also be a reasonable expectation of privacy in public places. An approach to quantify this deviation from expectation may be to first state the expectation in terms of Solove's taxonomy of privacy [11], i.e. to state which aspects of information collection, information processing, information dissemination, or invasion are expected by individuals. Then, a specific privacy violation can be analyzed with respect to the number of aspects that differ from the stated expectation.

Fourth, the harm to affected individuals can be financial, but also harm to an individual's reputation, harm caused by discrimination, or distress and anxiety. These privacy harms are all covered by (European) data protection legislation and individuals can sue for damages [7]. An important contributing factor in this is what has actually happened to the data: has it been exposed, modified, processed non-transparently, or used to make a decision affecting individuals? If exposed, to whom and what harms could and would they cause, given existing and potential future information available to the receivers of the data?

Individuals may have different perceptions of harm, especially non-financial harm, and therefore harm is difficult to quantify. A useful approximation may be to instead measure the amount of damages a court would be likely to grant.

Once suitable metrics have been identified for each of the four components, it is desirable to combine them into a single metric for privacy impact. Ideally, this combination should result in a metric with a meaningful unit, and not just an arbitrary number. Using our above suggestions as an example, we need to combine the number of people affected, the number of aspects that differ from expectation, bits of information revealed, and the monetary equivalent of harm to individuals. It is not immediately clear how this combination should be done, but it is unlikely that it is a simple matter of addition or multiplication.

4 Likelihood Quantification

Quantifying the likelihood of a privacy violation is somewhat more tangible than quantification of the impact. The NIST guide on privacy engineering [2] focuses on the likelihood of "problematic data actions." However, to make it clear how and why privacy violations happen, a systematic quantification should focus on the probability that a specific privacy violation occurs against an adversary with specific aims and capabilities that corresponds with a realistic attack model. Considering possible adversaries explicitly is necessary to make the likelihood quantification meaningful and highlights the assumptions made during the privacy risk assessment. An adversary is any party that is interested in private data, whether within the organization that holds the data, a connected organization such as a service provider, or an external third party. The issue that privacy risks exist even in the absence of "attacks" is acknowledged through the modelling of human error and accident as a non-malicious insider adversary. There is a wide variety of adversary models considered in the literature (see [13] for an overview). For adversaries that aim to breach privacy it is especially important to consider inference algorithms that allow the adversary to learn private information from public observations as well as the adversary's prior knowledge because combining data types can increase both likelihood and impact of a privacy breach.

The result of modeling possible adversaries is a set of probability distributions that indicate how likely it is for each adversary to succeed in breaching privacy.

5 Outcome of Privacy Risk Quantification

A privacy risk metric can be defined as a combination of metrics for impact and likelihood of privacy violations. The metric should allow not only to determine the average privacy risk, but also extreme values that may occur infrequently. In other words, privacy risk metrics should show the *distribution* of privacy risk instead of a single value.

Likert scores for impact and likelihood are often multiplied to determine the privacy risk score. Depending on the metrics used to quantify impact and likelihood, the multiplication operation may or may not be a good choice. Instead, addition or operations on probability distributions may make more sense and should be carefully considered. For example, privacy risk could be measured as the expected privacy impact, computing the expectation using the likelihood distribution and the impact metric, and taking the unit from the impact metric.

6 Discussion and Conclusion

Given that many different metrics may be used to measure impact and likelihood, and that they may be combined in different ways, it is possible to define different privacy risk metrics. However, depending on the specific "ingredients", a metric may not meaningfully measure privacy risk. It is therefore important to evaluate how strong privacy risk metrics are. One criterion to evaluate the

strength of metrics is monotonicity, i.e. that metrics should indicate lower privacy for stronger adversaries [12]. In addition, it may be helpful to calibrate new privacy risk metrics against a database of cases with known privacy risk, for example past cases where the impact is not speculative anymore, in particular with regard to privacy expectation and non-financial harm.

Finally, as with all rigorous methods supporting systems development, we should also take an economical aspect into account. In privacy risk measurement, we should avoid the false economy of accuracy, as according to Calder and Watkins [3] "the time cost of accuracy quite often outweighs the benefits for the organization".

Acknowledgment

This work was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) grant EP/P006752/1.

References

1. EU General Data Protection Regulation (2016), http://www.eugdpr.org/
2. Brooks, S., Garcia, M., Lefkovitz, N., Lightman, S., Nadeau, E.: An introduction to privacy engineering and risk management in federal systems. Tech. Rep. NIST IR 8062, National Institute of Standards and Technology, Gaithersburg, MD (Jan 2017)
3. Calder, A., Watkins, S.: IT Governance: An International Guide to Data Security and ISO27001/ISO27002. Kogan Page (2015)
4. Cavoukian, A.: Privacy by design: The 7 foundational principles (2011), https://www.ipc.on.ca/wp-content/uploads/Resources/7foundationalprinciples.pdf
5. Commission Nationale de l'Informatique et des Libertés: PIA manual (2015), https://www.cnil.fr/fr/node/15798
6. Deng, M., Wuyts, K., Scandariato, R., Preneel, B., Joosen, W.: A privacy threat analysis framework: Supporting the elicitation and fulfillment of privacy requirements. Requirements Engineering 16(1), 3–32 (Mar 2011)
7. Evans, K.: Vidal-Hall and Risk Management for Privacy Breaches. IEEE Security Privacy 13(5), 80–84 (Sep 2015)
8. Information Commissioner's Office: Conducting privacy impact assessments – code of practice (2014), https://ico.org.uk/media/for-organisations/documents/1595/pia-code-of-practice.pdf
9. Liu, K., Terzi, E.: A Framework for Computing the Privacy Scores of Users in Online Social Networks. ACM Trans. Knowl. Discov. Data 5(1), 6:1–6:30 (Dec 2010)
10. National Institute of Standards and Technology (NIST): Guide for Conducting Risk Assessments. NIST special publication 800-30 r1 (Sep 2012)
11. Solove, D.J.: A Taxonomy of Privacy. University of Pennsylvania Law Review 154(3), 477–564 (Jan 2006)
12. Wagner, I.: Evaluating the Strength of Genomic Privacy Metrics. ACM Trans. Priv. Secur. 20(1), 2:1–2:34 (Jan 2017)
13. Wagner, I., Eckhoff, D.: Technical Privacy Metrics: A Systematic Survey. arXiv:1512.00327 [cs, math] (Dec 2015)