Privacy Risk and Harms Assessment with the PRIAM Methodology

UiO :

Group Assignment 4B
UiO - Department of Informatics

Andreas Nygård Ljøterud, Hemel Howlander, Malte Hansen

October 31th, 2022

Contents

List of Figures

1 Introduction

2 The PRIAM methodology

In this section we aim to introduce how the PRIAM methodology works on a high level. In order to achieve this we will first provide a bird's-eye view of the steps involved and then briefly explain what each step entails.

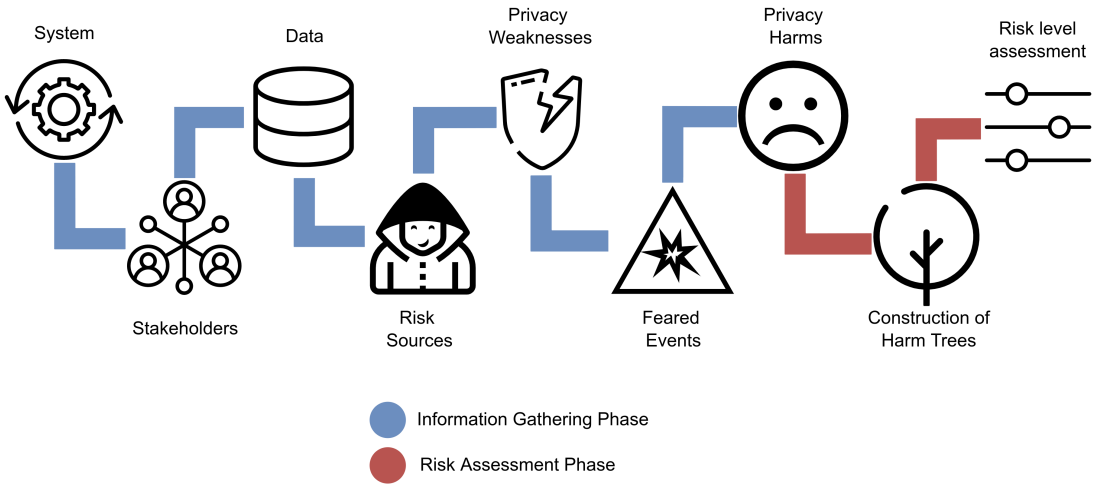

Figure 1:
The PRIAM methodology process[1].

The PRIAM methodology process (see Figure 1) is split into two phases, the first of which being the *Information Gathering Phase*. The main objective in this phase is to gather the relevant information needed in order to accurately carry out the risk assessment activities in the *Risk Assessment Phase*.

2.1 Information Gathering Phase

As the foundation for the upcoming risk assessment, it is essential that the information collected in this phase is complete. To reduce the likelihood of an analyst missing aspects that may have an impact on privacy risks, the framework defines seven *components*. These components, namely the *system*, *stakeholders*, *data*, *risk sources*, *privacy weaknesses*, *feared events* and *privacy harms* are each associated with a set of *categories* and *attributes*. A category is essentially a list of elements relevant to a specific component, while attributes describes the characteristics of a component that may affect privacy risks[1].

In Figure 2 an example is given to illustrate this concept further. The potential or capability of the risk source, in this case the *Hacker* (which is the category), would be determined by the assigned values of its attributes. Similarly for *Health data*, which is a category of the *data* component, the attributes could indicate a potential privacy weakness. The categories and their corresponding attributes are however not only used for indicative purposes, but for constructing harm trees and calculating risk levels in the next phase.

[1]All icons used are from https://thenounproject.com

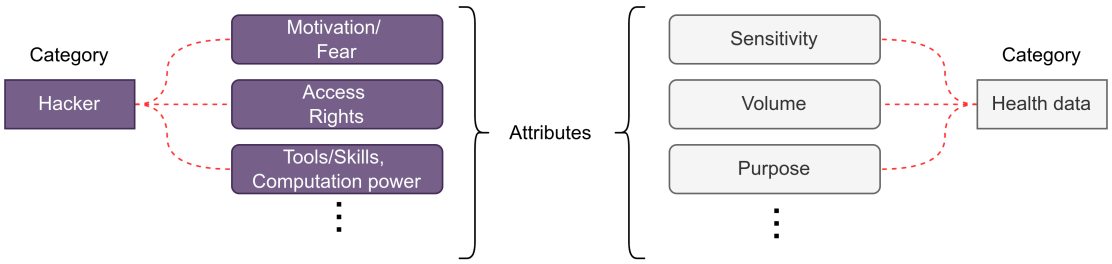

Figure 2:
Categories with their associated attributes.

2.2 Risk Assessment Phase

With basis in the information collected in the *Information Gathering Phase*, we are set to perform the risk assessment. In order to calculate the risk levels associated with a given scenario, so-called *harm trees* are constructed.

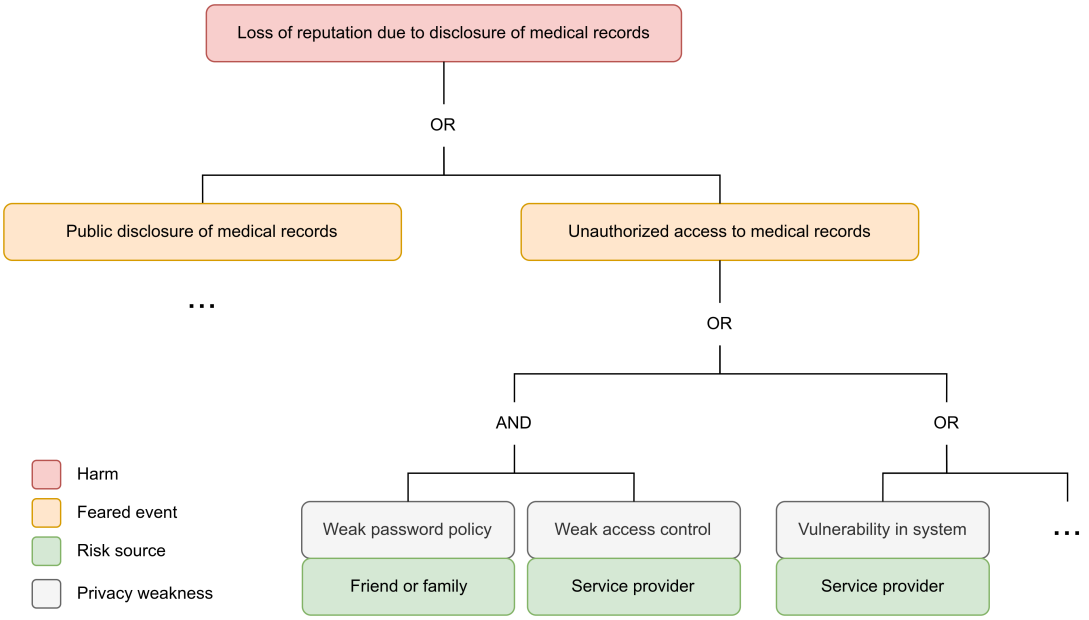

Figure 3:
Harm tree example.

Harm trees (see Figure 3), similarily to attack trees commonly used in computer security[1], are used to describe the relationship between a privacy weakness, a feared event and a harm (REF all these). The root node of such a tree will denote a harm, with branches leading to feared events that may lead to the harm in question. The leaf nodes denote privacy weaknesses and are represented by a pair consisting of the most relevant risk source coupled with a privacy weakness. In order to specify whether multiple events are required for an event to take place, the keywords **AND** and **OR** are used as decorators on the connecting paths between the nodes.

The risk levels for a given scenario, in this case the scenario previously given, is then calculated in the following way[1]:

1. **Find the value of *exploitability* for each leaf node in the harm tree.**

   The *exploitability* value is defined in the previous phase, and denotes how easy it would be to exploit a particular privacy weakness based on relevant system attributes. For simplicity and illustrative purposes, we have assigned the value medium to all three in this case.

2. **For each exploitation, choose the values of the relevant attributes of the risk source most likely to exploit the privacy weakness leading to the harm.**

   In order to calculate the likelihood of an exploit occurring, the characteristics of the risk source should be accounted for. In other words, does the profile[2] of the risk source favor an exploitation or not. In the case of *Friend or family* the majority of attributes does not favor an exploitation, while we consider the opposite to be true for both cases involving the *Hacker*.

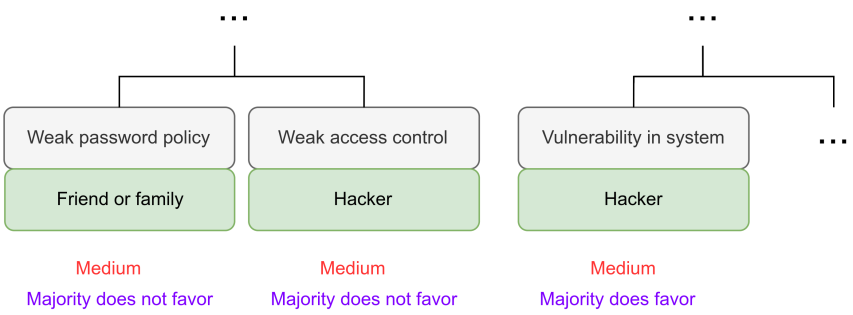

Figure 4:
Exploitability and risk source favoring.

3. **Find out the likelihood of each of these exploitations from the above exploitability value and values of the relevant risk source attributes**

Likelihood of exploitation	Exploitability of privacy weaknesses	Relevant risk source attributes
Negligible	Low	Majority do not favour
Limited		Majority favour
Intermediate		All Favour
Limited	Medium	Majority do not favour
Significant		Majority favour
Maximum		All favour
Intermediate	High	Majority do not favour
Maximum		Majority favour
Maximum		All favour

Figure 5:
Likelihood measurement scale.

The measurement rule used in PRIAM to determine the likelihood of exploitation can be seen in Figure 6 above, which means that the likelihood of our chosen exploitations are *Limited*. *Limited* and *Significant* respectively.

[2]Value of exploitation, motivation/fear etc.

4. **Compute the likelihood of each feared event and harm according to the following rules, where P_i is the likelihood of ith child node:**

   - AND node with independent child nodes: $\prod_i P_i$.
   - AND node with dependent child nodes: $Min_i(P_i)$.
   - OR node with independent child nodes: $1 - \prod_i (1 - P_i)$.
   - OR node with dependent child nodes: $\sigma_i P_i$.

   In order to compute the likelihood of the feared event and ultimately the harm, we simply apply the rules above to each level of the tree.

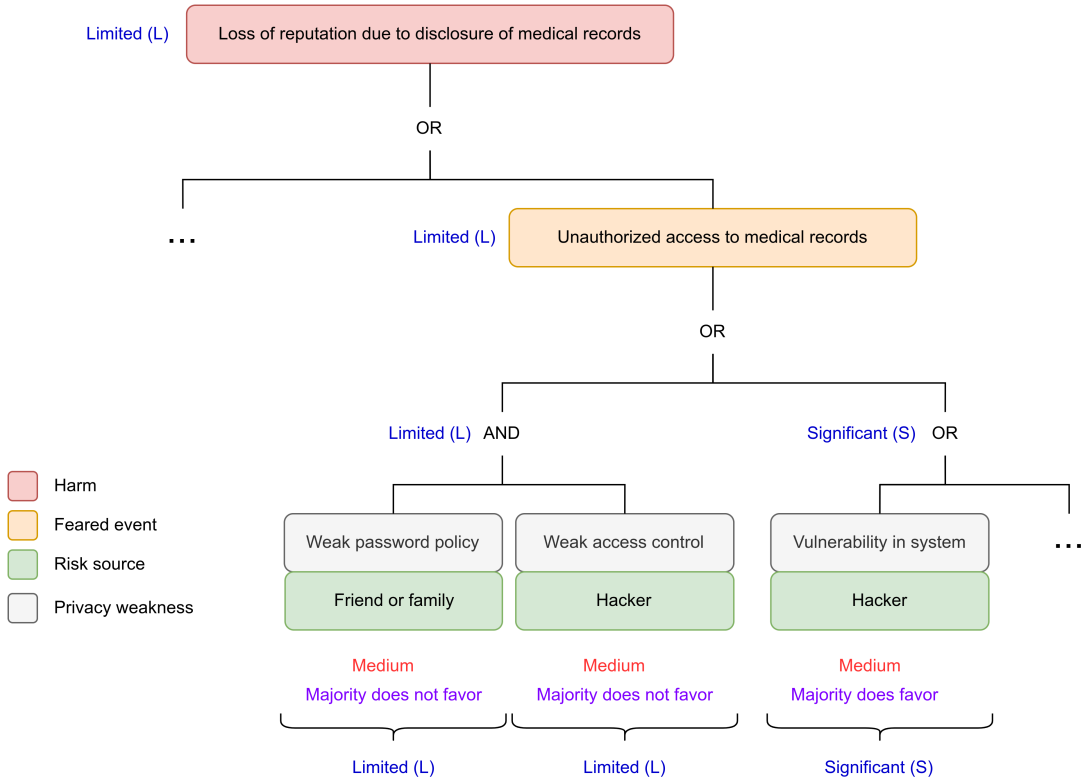

Figure 6:
Harm tree with computed likelihood.

3 Case Study: UCLA Health Cyberattack

In this section we will examine a real world example of a medical privacy breach. First, we will quickly introduce the case. Then, we will identify the privacy harms based on PRIAM. As this example provides a specific *feared event* we will not provide harm tree for this case study to keep the scope of the report in line.

3.1 Scenario

UCLA Health, the health system of the University of California, Los Angeles (UCLA), suffered a data breach in 2015, compromising sensitive personal data of approximately 4.5 million patients. In October 2014 UCLA Health started an investigation of suspicious network traffic , which did not appear to have malicious potential. However, in May 2015 a cyberattack involving the compromise of sensitive patient information was confirmed by the officials and affected patients were informed by the organization. The data compromised in the UCLA Health data breach included names, addresses, dates of birth, Social Security and medical record numbers, Medicare or health plan IDs, and medical information.[2]

3.2 Privacy Harms

We will now describe the privacy harms for the presented scenario, based on the PRIAM model[1, Section 3.7].
Considering the scope and circumstances of the UCLA Health breach, the following examples for categories of privacy harms can be observed:
Physical harms: A patient may be subject to stalking due to their address being leaked or receive wrong medical treatment as the consequence of the compromise of their medical data;
Mental or psychological harms: A patient fears misuse of their data, e.g. for identity theft, or is disadvantaged in an application due to a medical condition being disclosed; *Financial harms*: A patient has to pay a higher health insurance premium for their alleged sedentary lifestyle inferred from medical data, or have to defend them self against identity theft ;
Harms to dignity or reputation: Disclosure of intimate personal habits or unhealthy lifestyle may cause a patient embarrassment;
Societal harms: The society losing faith in the healthcare system, leading to a decline in member numbers;
Next, attributes in the form of *victim* and *intensity* can be assigned to the harms described in the categories of harms. Both attributes are measured as either *Low*, *Medium*, or *High*, depending on how many individuals are effected for the *victim* attribute and the consequences, as well as the duration and reversibility of those consequences for the *intensity* attribute. The *severity* of the harm is then evaluated as the multiplication of both attributes. An excerpt of the result for the given example can be seen in Table 1. For the complete description of the scale used, please refer to [1, Section 3.7.1].

Table 1: Examples of harms of the UCLA Health breach and their attributes

Harm	Example of event	Categories	Victims	Intensity	Severity
H.1	Stalking	Physical	Low	Medium	Limited
H.2	Identity theft	Psychological, financial	Low	High	Significant
H.3	Increased health insurance premium	Financial	Medium	High	Maximum
H.4	Disclosure of intimate personal habits	Harm to dignity	Low	High	Significant
H.5	Loss of trust in health care system	Societal	High	Medium	Maximum

4 Discussion

In the last section we took a look at the privacy harms that resulted from a hacking incident at UCLA Health. In this section, we will examine the results more closely and provide a discussion about the information one can gain from the PRIAM analysis.

In our example, we can observe that the *victim* attribtute is most often estimated as *Low*, while the *intensity* is almost always scaled as *High*. An important factor in the scaling was the difficult to impossible reversibility for most harms, as once data is disclosed it can not be undisclosed. This leads to the *severity* of each harm being almost exclusively rated as either *Significant* or *Maximum*.

TO ADD: 2nd paper, impact of each harm category

5 Conclusion

References

[1] Sourya Joyee De and Daniel Le Métayer. *PRIAM: A Privacy Risk Analysis Methodology*. Research Report RR-8876. Inria - Research Centre Grenoble – Rhône-Alpes, Apr. 2016. URL: https://hal.inria.fr/hal-01302541.

[2] Elizabeth Snell. *UCLA Health Data Breach Affects 4.5M Patients*. Ed. by healthitsecurity.com. Accessed on 24-10-22. July 2015. URL: https://healthitsecurity.com/news/ucla-health-data-breach-affects-4.5m-patients.