Publishing Data While Preserving Privacy

Harry Ros McArthur

Supervisors: Kate Smith-Miles and Peter Taylor (University of Melbourne),
Industry mentors: Joseph Chien and Chris Mann (ABS)

Data-dissemination agencies have been tackling the problem of privacy-preserving data analysis for decades. The central goal is to allow for the utility of published data to be as high as possible, while still ensuring the privacy of individuals contained in the data is respected.

Introduction

Suppose you are in control of a database, consisting of rows of information, where each row represents an individual. Suppose the database contains some non-sensitive information: ‘Name’, ‘Age’, ‘Postcode’ etc. Additionally, suppose it also contains some sensitive information. The classic examples are whether or not they are: a smoker, or if they have the ‘sickle cell' trait? Your goal is then to publish statistics about this database because researchers are interested in the connection between the non-sensitive information and the sensitive information. The more accurate the statistics you provide, the more ‘useful’ this information is to researchers. At the same time, a different party could use this information to their advantage. For example, if an insurance company can determine whether or not a particular individual is a smoker from these statistics, then they might charge this person higher premiums – violating that person’s privacy. From this, we arrive at the central question: “How do we make the data as 'useful’ as possible while ensuring the privacy of individuals is maintained?”

Motivating Example

To understand the philosophy behind the methods used to tackle this problem, consider a simple yes or no question which reveals some illegal or embarrassing activity, for example “Do you take drugs?”. In fear of judgment, people tend to answer untruthfully. To give participants a layer of protection, consider the following technique. Before letting them answer, flip a coin. If it lands on heads, tell them to answer truthfully. If it lands on tails, then flip the coin again. This time write down ‘Yes’ for heads, and ‘No’ for tails, regardless of the true answer, Figure 1. The reason why this method works, is that it provides plausible deniability to participants – their privacy has been preserved to some degree. Additionally, the data is still statistically ‘useful’ – on average, the proportion of truthful Yes and No responses stays the same.

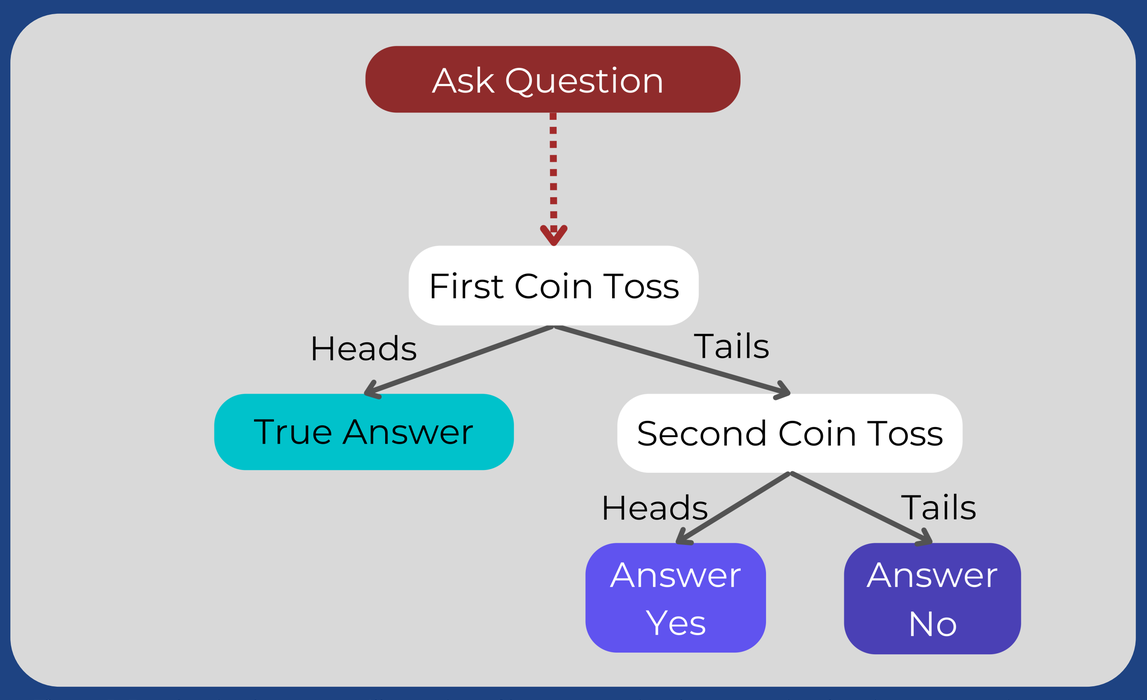

Figure 1: Randomised response. Illustration of how we can implement a simple privacy preserving mechanism to protect participants from revealing potentially embarrassing or illegal activity.

Scaling up

Despite the coin flipping example producing relatively useful data, we recognise that its utility has been harmed. This identifies a a key trade-off between privacy and utility. We can only improve one, at the cost of the other.

Based on the ideas presented in the motivating example, the current state of the art methods used to publish statistics about a database achieve their goal by injecting noise at the output level. This means that for a particular query, for example: “How many people over the age of X, living in suburb Y with income above Z, have some particular attribute?” The published noisy count is close to the true answer, but not exactly the same. For certain queries which are deemed more sensitive, this could mean the result is further from the true answer, Figure 2.

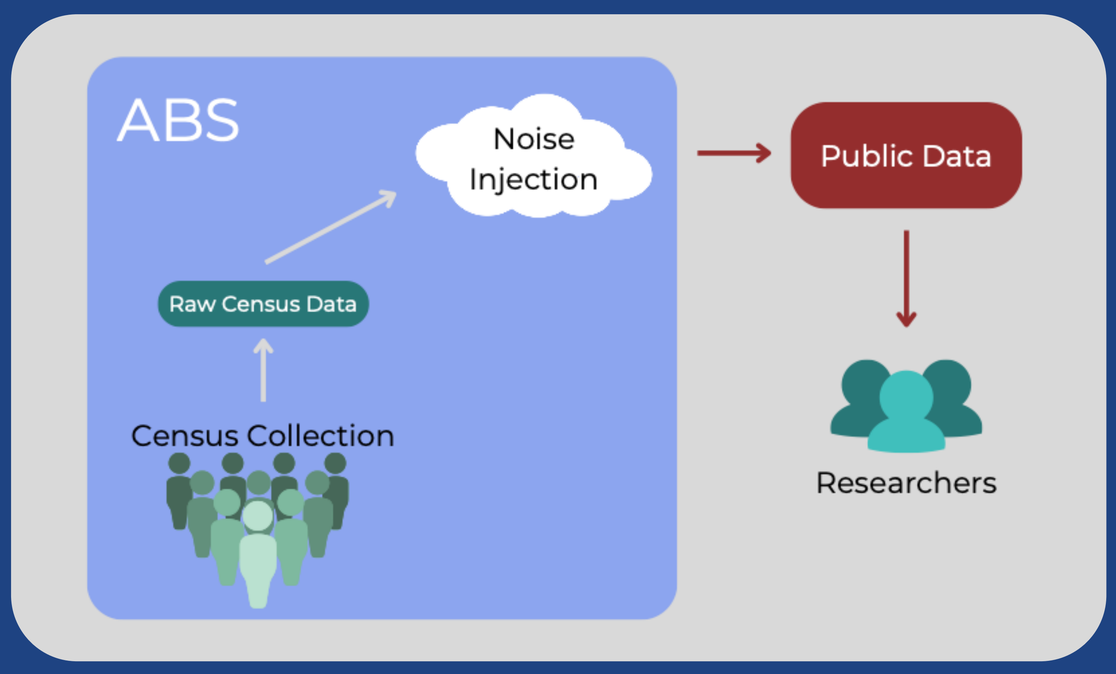

Figure 2: The Australian Bureau of Statistics injects noise into the data before releasing it to the public. This prevents the exposure of sensitive information to anyone analysing the results.

The Attacker’s Perspective

In order to quantify the level of leakage of a given privacy preserving technique, we can consider what an attacker can learn from the released information. Anything that can be learned about an individual is a potential violation of privacy. Broadly, there are two main types of attacks: Tracing and Reconstruction. Tracing attempts to determine whether or not a particular individual is present in the database, whereas Reconstruction attempts to physically recreate the underlying database, consistent with the published statistics. If we can construct a sequence of queries to deduce that there is only one feasible database consistent with those statistics, then we can say we have reconstructed the database exactly. Using optimisation methods we can reconstruct small databases consisting of few individuals and attributes exactly, provided the bound on the noise is relatively small.

Using optimisation methods to design attacks, and then exploring how the success of this attack strategy depends on certain parameters – like the noise distribution or bounds and the number of queries permitted – is a central focus of our research.

FOR FURTHER INFORMATION

Harry McArthur
e hmcarthur@student.unimelb.edu.au
w www.optima.org.au

REFERENCES

Dwork, Cynthia and Aaron Roth (2013). *“The Algorithmic Foundations of Differential Privacy”.* In: Foundations and Trends® in Theoretical Computer Science 9.3, pp. 211–407.

Garfinkel, Simson L., John M. Abowd, and Christian Martindale (2018). *“Understanding Database Reconstruction Attacks on Public Data”.*

ACKNOWLEDGEMENTS

This research was partially funded by the Australian Government through the Australian Research Council Industrial Transformation Training Centre in Optimisation Technologies, Integrated Methodologies, and Applications (OPTIMA), Project ID IC200100009.