Open Thesis / Project

Application-specific Privacy for Shared Data

(Embedded Information Processing Team)

Motivation

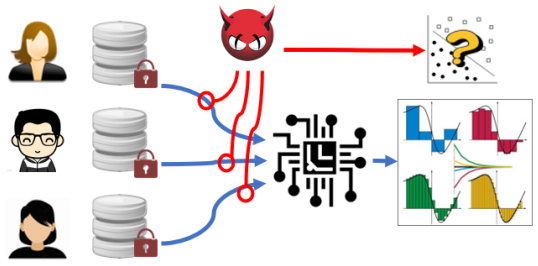
Volumes of data available these days is a major driving factor for innovation. This data, however, often resides in data silos guarded and managed by various parties. Data sharing is crucial to gain useful insights, improve services and ensure transparency of decision making, yet there are legitimate privacy concerns. In this thesis, you will implement and validate a method which enables privacy-preserving data sharing. You will implement a deep model architecture which ensures the data can be used for application-specific discriminative tasks but does not leak information about the defined sensitive classes. In contrast to existing works on preserving privacy in shared data, your method will allow to explicitly control the privacy-utility trade-off. You will test the method on several provided data sets and compare its performance to alternative methods. **Interested? Contact us for more details!**

Target Group

Students in ICE and Computer Science.

Thesis Type

Master Thesis (6 months).



An attacker gets access to shared data, but should not be able to discriminate sensitive attributes. Image source: Google.

Goals and Tasks

You will be provided with a solid starting point: a few ideas on how to achieve application-specific privacy for shared data, papers to read and data sets to work with. Your ideas regarding an efficient implementation of the deep model architecture are welcome. We would also like to have a thorough evaluation of the obtained model and compare it to alternative methods. If we are successful, we aim to publish the achieved results (paper writing is not part of the thesis and you will get all our support in this). The project includes the following tasks:

- Literature research on privacy in shared data;
- Discuss, design, implement and test a deep model architecture capable of preserving privacy of shared data;
- Test the obtained model on provided data sets. Compare the obtained results to at least one non-trivial benchmark. Explore the privacy-utility trade-off in different settings;
- Summarize your results in a written report.

Requirements / Skills:

- Good understanding of deep learning modelling, experience training a model on a GPU;
- Interest in data privacy and in working on a previously unsolved problem;
- Programming skills in Python, experience with PyTorch.

Contact Person

- Franz Papst (papst@tugraz.at)
- Dr. Olga Saukh (saukh@tugraz.at)

Institute for Technical Informatics
Networked Embedded Systems Group

Networked Embedded Systems Group