Detecting Privacy Violations with Machine Learning

Elliot BP - Supervised by: Kostas Markantonakis and Raja Naeem Akram
Information Security Group, Smart Card and IoT Security Centre

Objectives

- Identifying and understanding data regulations relating to user data and privacy (for example GDPR)
- Identifying and understanding consumer/user data requirements
- Learning and experimenting with machine learning models (deep learning) to analyse large scale enterprise data activities to assess whether a violation regarding consumer/user data occurred
- Predicting with probability weightage whether a squence of events leads to a data regulation violation

Introduction

This research project seeks to analyse the logs of a large-scale corporate network, to detect whether a violation of a user's data has occurred. The project uses the European Union's new General Data Protection Regulation as a regulatory framework to define and classify privacy violations.

A vast corpus of log activity is recorded, as every time data is accessed, manipulated or transferred is registered. This data is used firstly to construct a network, creating an 'exposure' model. After the exposure model is generated, a neural network will analyse the logged activity to detect violations.

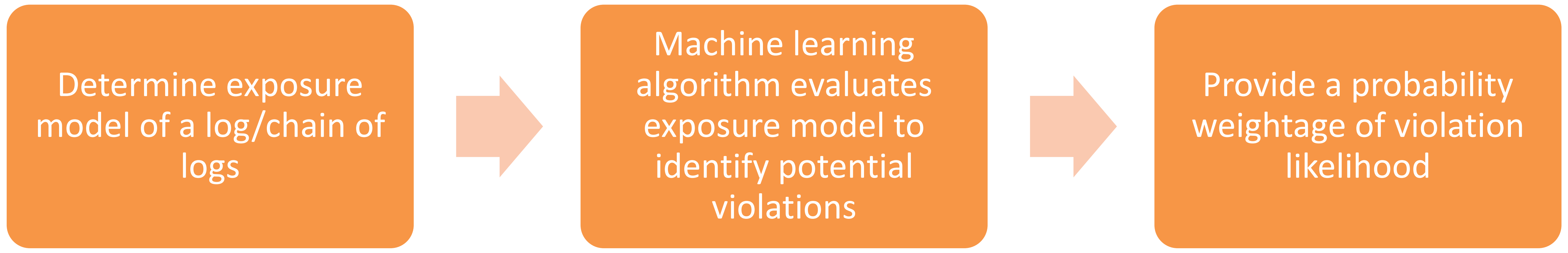
Materials

The following materials and resources are scheduled for use at various points in this project:

- Twitter API (provides dummy data to develop the exposure model generator)
- TensorFlow (for machine learning)
- NetworkX (for network graph modelling)

To Do

- Adapt the exposure model generator to process log data
- Create a neural network software package to analyse an exposure model and detect privacy violations



Determine exposure model of a log/chain of logs → Machine learning algorithm evaluates exposure model to identify potential violations → Provide a probability weightage of violation likelihood

Figure 1: Process of evaluating a log to detect privacy violations

Progress So Far

Twitter has been used as a placeholder to develop a dynamic exposure model generator; this is because the Twitter ecosystem closely mirrors a large-scale network of known agents. Tweets are analogous to user data, and the actions that Twitter users make - read, write or 'expose' by retweeting – mimics how data is processed by data processors.

At the time of publication, a dynamic network/graph generator that builds exposure models based on user data and user actions has been created. Currently using Twitter data, the model is easily adapted to logs and system event data.

Exposure Models

For the purposes of this project, an exposure model is defined to mean a subgraph of an overall network, given the context of a particular action.
More specifically, an exposure model is a network of agents (computational or human data processors) that have been 'exposed' to a particular instance of user data.

Agents are classified within an exposure model as one of the following types:

- Definitely exposed
- Possibly exposed
- Possibly exposed after another action*

*This classification of agents are known to not have been exposed to user data by the activity in question, but could be exposed in future by any of the agents known to have been definitely/possibly exposed.

Machine Learning

Later in this project, a (semi-) supervised machine learning algorithm will be developed to evaluate the effectiveness of the technology in this application. Labeled data will 'train' the algorithm to classify the activity within an exposure model.

Previous research into the applications of machine learning, particularly in the areas of Intrusion Detection System (IDS), human activity recognition and data mining are of interest and will form the foundation to the algorithm's development.

References

M. Mondal, P. Druschel, K. P. Gummadi and A. Mislove.
Beyond Access Control: Managing Online Privacy via Exposure.
Proceedings of the Workshop on Useable Security. 2014.

J. Saxe and K. Berlin.
Deep Neural Network Based Malware Detection Using Two Dimensional Binary Program Features.
Malicious and Unwanted Software (MALWARE), 2015 10th International Conference on. IEEE, 2015.

Contact Information

- Web: https://scc.rhul.ac.uk/
- Email: Elliot.Burke-Perrin.2017@live.rhul.ac.uk