RELIABLE MACHINE LEARNING (ML) MODEL DEPLOYMENT ON EDGE DEVICES

Mingyi Zhou, Professor John Grundy,
Professor Li Li, Dr Chunyang Chen and Dr Xiao Chen

At a glance

Goal

Build a reliable machine learning (ML) deployment framework to make the model 'non-reverse-engineerable'.

Background

Malicious users can access internal information of the ML models deployed on devices as the models are directly hosted on devices. They can then fool the smart Apps/devices easily.

Strategies

Used traditional techniques in software engineering like code obfuscation and program analysis to rephrase model information so it is difficult to reverse engineer.

Project partners

This project involved researchers from the CSIRO's Data61

Key outcomes

Identified methods of reverse engineering

We summarised potential ways to reverse engineer the ML model and further threats.

Developed ways to make ML files 'unreadable'

We developed a method to make the information of deployed ML models 'unreadable' for humans and reverse engineering tools with no impact on performance.

Protective custom program generation

We created a custom ML program generation approach that removes explicit ML-related information from the edge device to reduce the risk of reverse engineering.

More information

The approaches developed in this project:

- do not impact prediction accuracy
- will not affect the overhead of model inference
- can achieve greater security than conventional deployment strategies on real-world smart apps
- are able to work with existing methods of protection to safeguard ML models.

Access our open-source prototype tools

Learn more

To discover more about this project,
contact Mingyi Zhou or scan the QR code.

Acknowledgements

Mingyi Zhou is supported by a Monash Graduate scholarship.
Professor John Grundy is supported by the Australian Research Council Laureate Fellowship FL190100035.

www.monash.edu/it/humanise-lab

Australian Government
Australian Research Council