Exploiting power hunger of machine learning

Ilia Shumailov, Yiren Zhao, Daniel Bates, Nicolas Papernot, Robert Mullins, Ross Anderson
What is machine learning?

Machine-learning algorithms find and apply patterns in data. And they pretty much run the world.

by Karen Hao

November 17, 2018
Machine Learning

- Operation is data-driven
- It is suddenly hard to define Security
Computer Security in context of Machine Learning

- Adversarial examples exist for all models
- Attacks are scalable because of transferability
- A lot of different attacks are possible...
Availability

Ensuring **timely** and **reliable** access to and use of information
(NIST Special Publication 800-12)
Availability

Benign Data

Sponge Examples

Increased latency

Over-heating and over-consumption of energy
ML models process sentences as humans do

- Read words **sub-word at a time**

- Progressively **change their opinion** what sentence means as it reads

- It takes **more time to handle unfamiliar words**
Why does this happen?

Benign with 4 sub-words for input of size 16:
Athazagoraphobia => ath, az, agor, aphobia

1 error with 7 sub-words for input of size 16:
Athazagoraphpbia => ath, az, agor, aph, p, bi, a

Malicious with 16 sub-words for input of size 16:
A/h/z/g/r/p/p/i/ => A, h, z, g, r, p, p, i
Do sponges exist in practice? Yup
Conclusions

- Caused massive performance degradation with VERY easy tasks
  - Our computers x200 slower
  - Big ML-as-a-Service provider x6000 times slower

- Turned our hardware off through temperature

- Really hard to tell how to stop this attack ...

- May have underestimated impact of ML on climate
Thank you very much for listening!

Please do not hesitate to reach out in case there are any questions at ilia.shumailov@cl.cam.ac.uk