

Artem Dinh Furkan Sarikaya Tom Lu

What & Where & Who & When

All disciplinary incidents reported by school districts in New Mexico to the state's Public Education Department.

Covers the 2010-11 to 2021-22 school years.

Source: ProPublica & New Mexico Public Education Department

Incoming data is noised before entering database

Attack Venue

Batches of raw data

Noising at Education Department server



Batches of noised data



Noised dataset

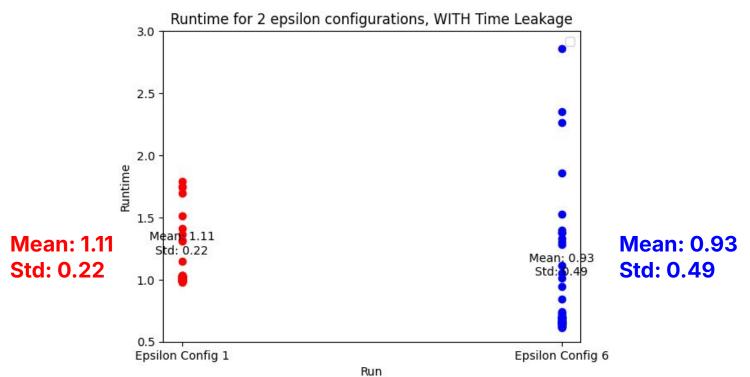


Malicious untrusted party can access Edu Dept. system and exploit side channel leakages



Side Channel Attack - Time

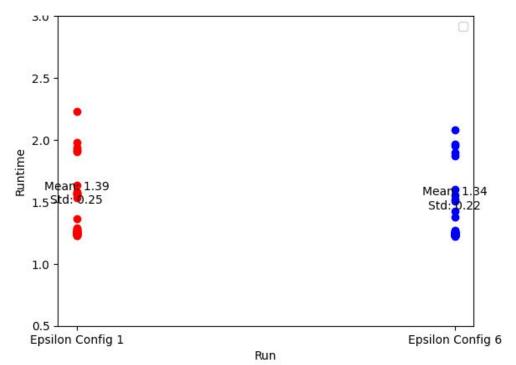
Graph of leakage of 50 runs



Mitigation Method

```
# Check if we should flip the value
              if random number <= flip probability:
                  # Flip the value by selecting a random value from known values other than 'x'
                  v = random.choice([value for value in X if value != x])
 Before
                  return v
              else:
                  # Keep the value x
                  return x
             # Check if we should flip the value
             if random number <= flip probability:
                 # Flip the value by selecting a random value from known_values other than 'x'
                 y = random.choice([value for value in X if value != x])
Improved
                 return y
              else:
                 y = random.choice([value for value in X if value != x]) # side channel prevention
                 # Keep the value x
                  return x
```

Mitigation Result of 50 runs



Mean: 1.39

Std: 0.25

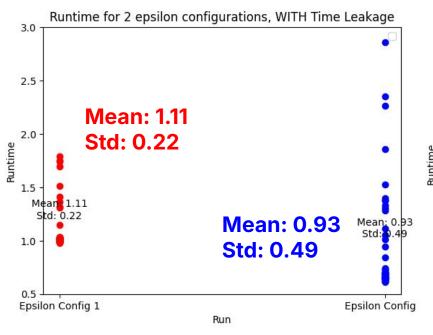
Mean: 1.34

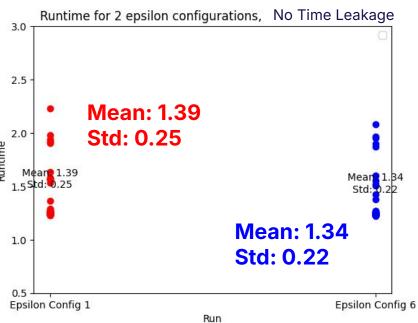
Std: 0.22

Before

&

Improved







Thank you! Enjoy your summer.

https://ieeexplore.ieee.org/document/9519418

https://arxiv.org/abs/1909.09630

https://arxiv.org/pdf/2111.11534.pdf



Attack Venue 2

Batches of raw data

Noising at Education Department server



Batches of noised data

Append

Skewed Noised dataset

Malicious untrusted party can access <u>School system</u> and sends large batch of data for poisoning attack



Significantly skew distribution of noised dataset, making deduction of true pre-noise distribution possible(???)