SparkFHE : Distributed Dataflow Framework with Fully Homomorphic Encryption

Advisor: Dr Peizhao Hu
Student: Aamir Jamal
Three Stages of Data Privacy

- In Transit to Cloud (TLS)
- At rest in Cloud (AES)
- During computation in Cloud?
Why it matters?

• Insider threats
• Outsider threats
• Regulations
• Liability concerns
Homomorphic Encryption

Data owner

- Secret input
  - 20
  - Encrypt
  - $fA4!&s2FDfs4

- Computed result
  - 30
  - Decrypt
  - e3Ad09!B%gD

Untrusted environment

- $fA4!&s2FDfs4
- Compute while encrypted
- -5
- x2
- e3Ad09!B%gD
Proposed Work

A working Application demonstrating the SparkFHE project

• An application which would be taking in encrypted inputs from the user, computing on the basis of a chosen operation, and giving out the encrypted output, by using the SparkFHE APIs

• Handling encryption and decryption of the values on the front end using a generated key
Evaluation

• Successful and clear demonstration of encrypted computation
• Easy to understand and use
• Produce value or be useful
Background & related work

- Apache Spark, a highly efficient data processing framework for the cloud
- FHE, a cryptography technique that enables arithmetic computations on encrypted data without decrypting it first
- SparkFHE integrates Apache Spark with FHE
- Project Source code: https://github.com/SpiRITlab/SparkFHE-Examples/wiki
Milestones

• First milestone is to successfully integrate a basic front end with the Spark API
• Rest milestones are yet to be decided
Thank You