**Suspicious Person List – experiment using HE**

**Use Case:**

**Public to public – international trusted group**

|  |  |  |
| --- | --- | --- |
| **Medium trust** | **Encryption in transit, operation, rest** | **Results summary available to all parties** |

Compare SPL to other SPLs held by trusted international public parties.  SPL is push only – a public party compares the SPL to one or more international public parties. Results are available in summary only, with full details on the application to the public party holding the information.

**Privacy**

This use case enables medium trust comparison by maintaining data encryption during transit, operation, and at rest, combined with trusted international security and governance arrangements. It maintains privacy by ensuring results are only available in summary to trusted international parties and requiring the consent of a data holder before full details are made available.

**Technology**

1. Homomorphic encryption (comparison): Allows encrypted data to be compared without decrypting it first, protecting data privacy during the comparison process.
2. Symmetric encryption (sharing summary and full details): Used to encrypt the summary results and ensure that sensitive information is secure when shared or stored.

**Demonstrating the Experiment**

1. **Data Generation:** Create synthetic SPL data for demonstration purposes. This simulates the SPLs held by different parties.
2. **Encryption:** Use homomorphic encryption to securely encrypt the SPL scores, allowing for secure comparison.
3. **Comparison:** Perform the comparison on encrypted data to identify matches without revealing the actual scores.
4. **Results Summary:** Encrypt the comparison results (summary) to ensure it remains confidential during sharing. Decrypt the results to display them.

This approach demonstrates how privacy-preserving technologies can be applied in practical scenarios where multiple parties need to collaborate while protecting sensitive information.

**Code Explanation:**

**Required Packages:**

* **TenSEAL**: For Homomorphic Encryption.
* **Cryptography**: For Symmetric Encryption (e.g., AES).
* **Pandas**: For handling data.

**Synthetic Data Generation**:

* The generate\_synthetic\_spl() function generates synthetic data for the Suspicious Person List (SPL). The person\_id is a unique identifier, and the suspicious\_score is a numerical value indicating the suspicion level. This data simulates real-world SPL data.

**Homomorphic Encryption Context Initialization**:

* The initialize\_he\_context() function creates a TenSEAL CKKS context for homomorphic encryption. CKKS allows for encrypted computations on floating-point numbers, ideal for operations like comparisons. The context is configured with specific parameters, including the polynomial modulus and global scale. Galois keys are generated to enable operations like rotations.

**Encrypt SPL Data**:

* The encrypt\_spl\_data() function encrypts the suspicious\_score column of the SPL using CKKS encryption. Each score is encrypted separately and stored in a list of encrypted vectors.

**Compare SPL Data**:

* The compare\_spl\_data() function compares two lists of encrypted SPL data. It subtracts corresponding encrypted scores to check for matches. If the decrypted result of the subtraction is close to zero, it indicates a match. The results are stored as a list of boolean values.

**Symmetric Encryption (AES)**:

* The symmetric\_encrypt() function encrypts data using AES encryption in CBC mode. The function generates a random initialization vector (IV) for each encryption operation, ensuring that the same plaintext encrypts to different ciphertexts. The data is padded to ensure compatibility with the AES block size.
* The symmetric\_decrypt() function decrypts the AES-encrypted data, reversing the encryption process and removing the padding to retrieve the original plaintext.

**Main Function**:

* The main() function demonstrates the entire process:
  + Synthetic SPL data is generated for two public parties.
  + The SPL data is encrypted using homomorphic encryption.
  + The encrypted data is compared to find matches.
  + A summary of the results (e.g., the number of matches) is created and encrypted using symmetric encryption.
  + The summary is decrypted for demonstration.