**SUSPICIOUS PERSON LIST COMPARISON**

**Process**

Suspicious persons lists are collated periodically using one or many unique identifiers such as driver licence, passport or national ID number. All available identifiers (multiple passport numbers and licences per person) are compared against identity lists held by each participating party.

**Privacy considerations**

**High trust situations** use existing security and governance arrangements to compare data and perform unencrypted operations on secure systems. Data is encrypted during transit and at rest.

**Medium trust situations** require privacy preserving technology to compare data while encrypted. Governance arrangements must be developed between parties.

**Low trust situations** require privacy preserving technology to compare data while encrypted. Governance arrangements must prevent data sharing unless a trusted party approves.

**Use cases**

**Public to private**

|  |  |  |
| --- | --- | --- |
| **Low trust** | **Encryption in transit, operation, rest** | **Results available to public only** |

Compare publicly held suspicious person lists (SPL) to the customer lists of private companies. SPL is push only – the public party pushes to the private parties for matching or the private parties push to the public party. Comparison results are available to the public party only. Private parties should remain unaware of results. Public parties can use the results to request additional information from private parties.

**Privacy**

This use case enables low trust comparison by maintaining data encryption during transit, operation and rest. It maintains privacy by ensuring the results are available to public parties only.

**Technology**

1. Homomorphic encryption (comparison)
2. Symmetric encryption (sharing results).

Binary result is returned for a request. The partner who is doing the request will decide on weather to disclose information on that person. You can use any identifier that is current or past to ID someone. The entity that is doing the matching will give all the identifiers they have available to the entity they would like to match. If there are additional identifiers, they you can make additional match requests iteratively.

Suggestion 1: identifying individuals. There is a technique called Probabilistic Seek (P Seek). Technique that can be used in the Australian scenario when there is no unique identifier. If the identifier is only available to one agency such as Austrac. They would not necessarily have that identifier. There can be access issues.

Suggestion 2: Investigate MPC (secure multiparty computation) this matches one scenario in MPC. Helps match who on your list is in a particular data base. It can be done efficiently. Both parties will be able to know who the customers are who are the intersection between these two sects.

**Public to public – domestic trusted group**

|  |  |  |
| --- | --- | --- |
| **High trust** | **Encryption in transit and at rest** | **Results available to all parties** |

Compare SPL to other SPL held by trusted domestic public parties. SPL is push only – a public party pushes the SPL to one or more domestic public parties for comparison. Results are available to all parties.

**Privacy**

This use case enables high trust comparison by maintaining data encryption during transit and rest, combined with trusted domestic public party security and governance arrangements. It maintains privacy by ensuring the results are available to trusted domestic public parties only.

**Technology**

1. Symmetric encryption (sharing full details)

**Public to public – international trusted group**

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

Compare SPL to other SPL held by trusted international public parties. SPL is push only – a public party pushes the SPL to one or more international public parties for comparison. Results are available in summary only, with full details available on 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)
2. Symmetric encryption (sharing summary and full details)

Possible use case for SMC.

**Public to public – international untrusted group**

|  |  |  |
| --- | --- | --- |
| **Low trust** | **Encryption in transit, operation, rest** | **Results available to data holder only** |

Compare SPL to other SPL held by untrusted international public parties. SPL is push only – a public party pushes the SPL to one or more international public parties for comparison. Results are only available to the data holder. The data holder must choose whether to share the results with the requesting public party and whether they share a summary or full details.

**Privacy**

This use case enables low trust comparison by maintaining data encryption during transit, operation and at rest. It maintains privacy by ensuring results are only available to the data holder which can choose whether to share results with the requesting public party.

**Technology**

1. Homomorphic encryption (comparison)
2. Symmetric encryption (sharing summary or full details)

Countries that you do not necessarily deal with often and have no trust in what they are going to do with your data. The first example of SMC or homomorphic encryption where only one or an intermediary will obtain the results. An international organisation like Interpol is a good example of this use case where it does not necessarily provide the results back to the party who requested the results.

**Private to private – domestic trusted group**

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

Compare SPL to other SPL held by trusted domestic private parties. Results available to all parties in summary, with full details available by mutual agreement.

**Privacy**

This use case enables medium trust comparison by maintaining data encryption during transit, operation, and at rest. It maintains privacy by ensuring only personal information related to suspicious *customers in common* is shared between private parties. Governance arrangements between the parties would need to be developed. Private parties likely could not rely on existing exceptions in privacy legislation related to law enforcement and would probably need to gain the consent of their customers through changes to company privacy policies and terms of service.

**Technology**

1. Homomorphic encryption (comparison)
2. Symmetric encryption (sharing summary or full details)

Can be problematic due to the law in most countries. Privacy legislation operates meaning that you need encryption throughout transit operation and at rest. Better suited for an international context rather than the Australian market. America potentially?

**Feasibility**

Comparing SPL between private parties using homomorphic encryption is feasible where the number of suspects is relatively low for each participating private party. Experiment 02 in trimester 1 2024 showed that a list of ten suspects compared against ten lists of ten suspects using homomorphic encryption was approximately 9000 times slower than cleartext or hashed comparison. However, the security gained through homomorphic encryption could be worth this tradeoff in the right use case.

For example, a group of banks which wanted to compare SPL daily could restrict the suspects on their lists to only those which were under active investigation. This would limit the match operations required. This list could be expanded to include all suspects which were still customers of the bank (had not been debanked) if the time required to complete the SPL matching was less than 24 hours.

If the process is completed once every 24 hours, the requirement for daily matching would be met. If the 24-hour threshold is exceeded, the banks could invest in additional computing resources to reduce the match time or restrict the number of suspects being matched through administrative decisions.

**Private via public to private – domestic untrusted group**

|  |  |  |
| --- | --- | --- |
| **Low trust** | **Encryption in transit, operation, rest** | **Results summary available to public party** |

Compare SPL to other SPL held by untrusted domestic private parties. Results available to the public intermediary only, with summaries or full results made available to private parties at the discretion of the public party.

**Privacy**

This use case enables low trust comparison by maintaining data encryption during transit, operation, and at rest. It maintains privacy by ensuring only the public party receives personal information related to suspicious *customers in common.*

The public party may require changes to their enabling legislation to define their role as an intermediary between private parties, including changes or addition to prevent private parties from further disclosing personal information shared with them by the public party, not dissimilar to existing ‘tipping off’ provisions in Australia’s AML/CTF legislation.

**Technology**

1. Homomorphic encryption (comparison)
2. Symmetric encryption (sharing summary or full details)

Should be suited to the Australian Environment, the results are only available to a public intermediary accessing as a broker who we want to match against all of these different people. Puts the private parties in contact under a governance arrangement.

**SUSPICIOUS TRANSACTION DISCOVERY - MODEL COMPARISON AND SHARING**

**Process**

Models are developed by each party using their own data, then encrypted and shared in a federated learning process, undergoing multiple iterations to improve performance.

**Privacy considerations**

Not applicable.

**Use cases**

**Public/private to public/private**

|  |  |  |
| --- | --- | --- |
| **Trust N/A** | **Encryption in transit** | **Results available to all parties** |

Share models and compare relative effectiveness to test and develop more effective models for detecting suspicious transactions. Results available to all parties.

**Privacy**

Models do not include personal information. Data is encrypted in transit for best practice and protection of commercial interests.

**Technology**

1. Federated learning (comparison and sharing)
2. symmetric encryption (protecting commercial interests).

Sharing private data that the models are based on. Sharing the models themselves that are available to the parties that are contributing. Anyone can try out the model and test the best model that suits the data. This can be a good approach for smaller institutions. Larger institutions have more specialised models which could be less suitable.

**SUSPICIOUS TRANSACTION DISCOVERY - DATASET SHARING**

**Process**

Datasets are shared between parties to allow development of new suspicious transaction detection methods.

**Privacy considerations**

Individual personal information within the dataset must be protected. The dataset should meet privacy standards and successfully resist reconstruction attacks.

**Use cases**

**Public to public/private**

|  |  |  |
| --- | --- | --- |
| **Low trust** | **No encryption** | **Results available to all parties** |

Share transaction datasets to enable public and private parties to test suspicious transaction discovery models. Results available to all parties.

**Privacy**

Dataset must guarantee privacy mathematically using differential privacy or similarly effective techniques. Dataset must successfully resist reconstruction and other attacks designed to identify individual persons within the data. If the dataset is released in its entirety, it does not need to resist timing or floating-point attacks.

**Technology**

1. Differential privacy

Injecting noise into the data set to share the data set and works on plausible deniability so you cannot confirm if someone is in the data set. Making data generally available, generally sensitive data and not a common use case.