The objectives of my research are in the following directions:

- **To prevent query access pattern in searchable encryption schemes**

  A significant amount of proposed searchable encryption schemes leak some information during search operations due to access and query pattern. Consequently, this potential privacy concern introduces attacks by an adversary that controls the cloud service provider, and hence can observe patterns. Moreover, using some prior knowledge about data and query distribution, the adversary can run simple inference attack. There are known lightweight countermeasures to hide the access pattern by padding-based techniques, but the same is not true for the query pattern. To this purpose, we addressed this issue by hiding query access pattern for scalable databases. Our approach makes significant contributions as the follows:

  - Provides an algorithm that incurs a constant overhead per real query independent of the database size
  - Adopting a technique to make the query distribution uniform
  - Prevents frequency analysis attack knowing prior knowledge about the query access distribution

- **To design an efficient access control scheme for shared data**

  The problem of cloud data sharing has many practical applications in healthcare where patients are willing to share their personal health records with a group of care practitioners to improve availability and coverage of remote assistance. A straightforward solution is to encrypt data before outsourcing it to the cloud to ensure confidentiality. However, several cryptographic techniques do not support access control where different users have different access rights to the shared data. Regarding this concern, we proposed a solution that adopts attribute based encryption, which enables a user to selectively share its data with other users whose access rights match their attribute sets. We addressed the limitations of
attribute based encryption that potentially jeopardize users privacy and does not allow updating attribute sets without re-encryption. Our scheme makes the following contributions:

- Provides a transformation key for which a ciphertext associate with attribute vector can translate to anew ciphertext associated with a different attribute vector
- Provides a policy update mechanism with a performance suitable for many practical applications

- To construct a private set intersection protocol

The problem of privacy-preserving set operations, in particular, private set intersection (PSI), is important whenever one or more parties (who do not fully trust each other) are willing to perform joint operations on their respective sets without revealing additional information beyond the result of the operation itself. This problem has been used in several real-world applications and can be classified based on parties set sizes: Symmetric (balance) when the sets have roughly the same size, and Asymmetric (unbalance) when one of the sets is substantially smaller than the other. To address this problem, we considered a simple PSI set up including a client and a server whose aim is to compute the set intersection of their respective sets while revealing minimal information. Our method makes several contributions:

- Evaluates and compares existing private set intersection protocols in terms of efficiency (computation and communication overhead) and security
- Adopts Hashing and batching method to obtain better performance
- Protects data confidentiality using differential privacy mechanism

- To design a cryptographic cloud-based technique for air cargo cancellation problem

One of the main problems in air cargo revenue management is the modeling of over booking and cancellation in the operation of the service chain. In fact, Airline Cargo companies do not impose any fees for last-minute cancellations of shipments. As a result, customers can book the same shipment on several cargo companies. Cargo companies try to balance cancellations by a corresponding volume of over booking. Towards this concern, we developed a cryptographic technique, enabling the computation on private information of the airline customers and companies data to improve the overall service chain. Our technique is based on proxy re-encryption method and makes the following contributions:

- Mitigates the airline cargo cancellation problem while preserving the privacy of customers' data
- Provides a Game theoretic model enabling a more accurate estimate of the cancellation rate

- To analyze novel annotations in the gene ontology for boosting selection of negative examples

The protein function prediction (PFP), which involves sophisticated computational techniques to accurately predict the annotations of new protein and proteomes, is an emerging and challenging problem in the post-genomic era. Hence, an important issue in PFP is the selection of reliable negative examples for learning accurate predictors. Usually, public repositories for genome and proteome annotations, such as the Gene Ontology (GO), rarely stores negative annotations, i.e. proteins not possessing a given function. If a protein is not currently annotated with a GO term, it could be either that the protein is a negative example for that term or a positive example which has not been detected yet due to insufficient investigations. Therefore, our technique makes the following contribution:

- Provides a technique for the selection of negative examples to train accurate functional classifiers
Risultati della ricerca:

Here, I briefly explain our technique proposed to achieve our mentioned beforehand objectives:

**Objective 1:** Towards this objective, we proposed a low-cost hiding of the query pattern for schemes supporting point and range queries. Our countermeasure is to use fake queries and cache, trying to achieve uniform distribution. We used observed query distribution to generate fake queries, and once we learned the distribution, we make the distribution uniform. We introduced a protection algorithm where the query pattern follows the common Zipf distribution. Our first approach was applied for point queries, and then we improved our proposed idea for schemes supporting range queries as well. Compared to the previous works, our approach does not reveal the query distribution and, therefore, the adversary cannot recover search keys after observing $N^2$ queries when the database is dense.

- Results:
  - Low-Cost Hiding of the Query Pattern, (to be submitted on July 2019)

**Objective 2:** Towards this objective, we proposed a proxy-based technique using a well-known inner-product encryption (IPE) primitive. Using IPE, it is possible to define fine-grained access policies over encrypted data and the decryption key depend, is equal to zero. Since the current IPE schemes do not support efficient access policy changes, therefore, we proposed an efficient inner-product proxy re-encryption (E-IPPRe) scheme that provides the proxy server with a transformation key, with which a ciphertext associated with an attribute vector can be transformed to a new ciphertext associated with a different attribute vector, providing a policy update mechanism with a performance suitable for many practical applications. We experimentally assessed the efficiency of our protocol and showed that it is selective attribute-secure against chosen-plaintext attacks in the standard model under the Asymmetric Decisonal Bilinear Diffie-Hellman assumption.

- Results:

**Objective 3:** Towards this objective, we considered the problem of private set intersection (PSI) in two party computation model. We reviewed the progress made in PSI techniques in terms of performance (communication and computation overhead) adopting a fully homomorphic encryption scheme. We achieved an optimized solution that uses batching and hashing to reduce the size of transmitted data. The proposed technique allows the server to operate $n$ items from the client simultaneously. The work was implemented with C++ using HElib library and the results were compared to the state of the art.

- Results:
  - Differentially Private Two-Party Set Operations (to be submitted by May 15th, 2019)

**Objective 4:** Towards this objective, we proposed a proxy-based technique to mitigate the problem of air cargo cancellation and over booking. Therefore, we proposed a proxy re-encryption technique so that the customer company could agree on the constraints that it will not book the same shipment over more cargo companies and the cargo company can enforce the agreement by querying the privacy preserving information system at every cancellation, or, more realistically, for a sub-sample of cancellations. We described such system based on proxy re-encryption and showed that how the relevant information can be be extracted, still preserving the privacy of customers' data.
- Results:

Objective 5: Towards this objective, we designed a technique for the selection of negative examples to train accurate functional classifiers. Therefore, we carried out the study of annotation evolution through different release of the Gene Ontology (GO). We provided a basilar knowledge for any future approach to detect negatives in the GO. Furthermore, we designed a novel method to select negative examples for GO terms, and we assessed its effectiveness in an experimental comparison involving thousands of GO terms and the state-of-the-art methodologies proposed for the same task. Our results revealed that novel annotations for a given protein tend to appear on terms with high semantic similarly with the terms the protein was already annotated with in the previous GO release

- Results:
  - M. Sepehri, and M. Frasca. Analysis of Novel Annotations in the Gene Ontology for Boosting Selection of Negative Examples. In 9th International Conference on Biomedical Engineering and Technology, ICBET 2019, Tokyo, Japan. (Published)

Attività svolte:

Prodotti della ricerca conseguiti: (in termini di pubblicazioni, brevetti, ...)

International Peer-Reviewed Conferences

Accepted


- M. Sepehri, and M. Frasca. Analysis of Novel Annotations in the Gene Ontology for Boosting Selection of Negative Examples. In 9th International Conference on Biomedical Engineering and Technology, ICBET 2019, Tokyo, Japan. (published)

Accepted without Proceeding

- M. Frasca, M. Sepehri, A. Petrini, and G. Grossi. Committee-based Active Learning to Select Negative. 15th International Conference on Computational Intelligence Methods for Bioinformatics & Biostatistics, Caparica, Portugal, 2018

Ongoing papers

- Low-Cost Hiding of the Query Pattern (to be submitted around June 30th)

- Differentially Private Two-Party Set Operations (to be submitted on May 15th)

- A Proxy-Based Scheme for Fast Equality Queries over Encrypted Data (Not decided yet)
Descrizione dell'attività di ricerca svolta all’estero (eventuale: specificare: periodo, luogo, affiliazione):

- Visiting researcher in the David R. Cheriton School of Computer Science at the University of Waterloo, Canada to collaborate with Prof. Florian Kerschbaum working on Privacy-Respecting Information Sharing (fall 2018)

A part of my visiting research application at Waterloo University was dedicated to constructing a private set intersection protocol base on differential privacy mechanism and reconstruction attacks on searchable encryption schemes.

I also gave two talks in the Cryptography, Security and Privacy (CrySP) group with the following titles:

- Reconstruction attacks on searchable encryption techniques supporting range queries
- A possible countermeasure for hiding query pattern for point and range queries

La presente relazione, non contiene dati sensibili e dati giudiziari di cui all’art. 4, comma 1, lettere d) ed e) del D.Lgs. 30.6.2003 n. 196.
Si autorizza la pubblicazione della relazione annuale sul sito web del Dipartimento.

Firmato (In Stampatello) NOME MARYAM COGNOME SEPEHRI

Data 09.05.2019

Il Responsabile Scientifico

[Signature]

L’ Assegnista di Ricerca

[Signature]