Secure Signal Processing for Genomic Privacy Protection

Mina Namazi

Advisors: Juan Ramón Troncoso-Pastoriza and Fernando Pérez-González

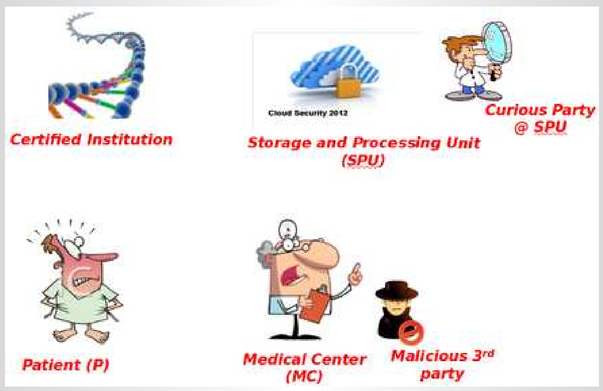
{mnamazi,troncoso,fperez}@gts.uvigo.es

Universidad de Vigo
Signal Theory and Communications Department
University of Vigo
Spain

Workshop on Monitoring PhD Student Progress. June 13, 2016

1. Motivation of the Work

Privacy-aware signal processing applications such as genomic research have considerably grown in the recent years due to the unprecedented advances brought about by Next Generation Sequencing (NGS) and the need increasingly widespread use of outsourced processing. The benefits of an extensive use and study of genomic data in advancing medicine research are unquestionable, but the inherently sensitive and identifiable nature of the genome entails severe privacy risks which are aggravated when the sequences are outsourced to an untrustworthy environment, like a Cloud service, for their processing. Outsourcing genomic data exposes them to the service and infrastructure providers and makes them vulnerable to attacks and accesses violating patient's privacy.

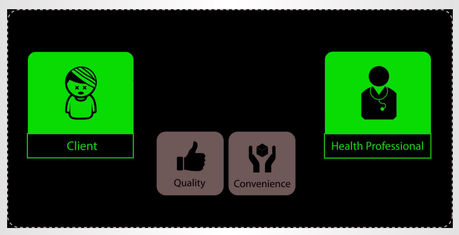


A proper protection of the genomic data by solely applying anonymization techniques is proven to be infeasible, unless the data are rendered useless. Moreover, genomic information can be linked to ancestors and relatives of an individual, so its leakage also affects their privacy. Therefore, a combination of anonymization techniques and encryption techniques under the paradigm of **Secure Signal Processing** (SSP) is a crucial aspect for protecting the individuals' privacy when processing genomic information in outsourced environments.

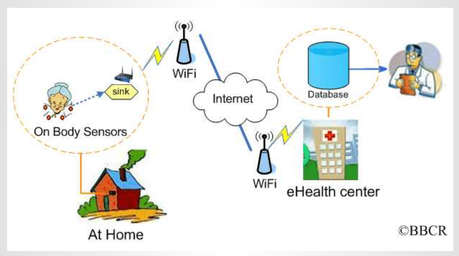
2. Thesis Objectives

The main objective during the development of this PhD Thesis is to **advance the state of the art in secure signal processing cryptographic methods for secure outsourcing of privacy aware applications in the e-Health area**.
Specifically, the three main objectives are the following:

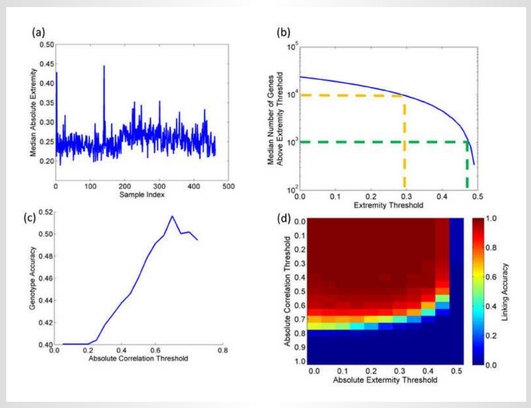
A. Analysing the proposed schemes and techniques for secure signal processing e-Health applications from a privacy and security point of view.



B. Designing novel secure signal processing methods for privacy preserving e-health applications enhancing efficiency, privacy level and reducing interactivity.

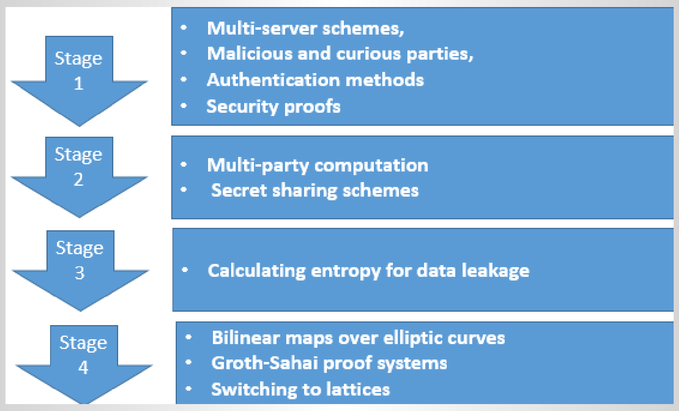


C. Devising new information-theoretic metrics to quantify the information leakage on genomic data when it is partially protected or when the results of several subsequent processes are disclosed.

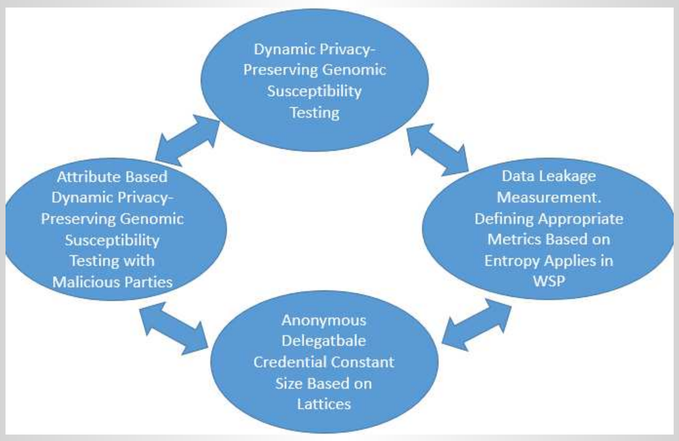


3. Research Plan

The research plan for the next year is focused on **Dynamic Privacy-Preserving Genomic Susceptibility Testing**:

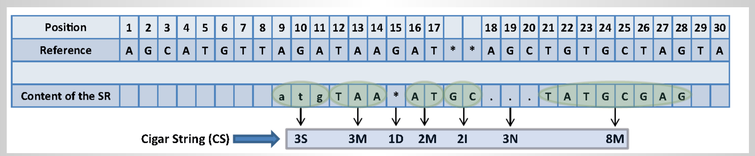


And the **Methodology** to achieving this goals is:



4. Results and Discussions

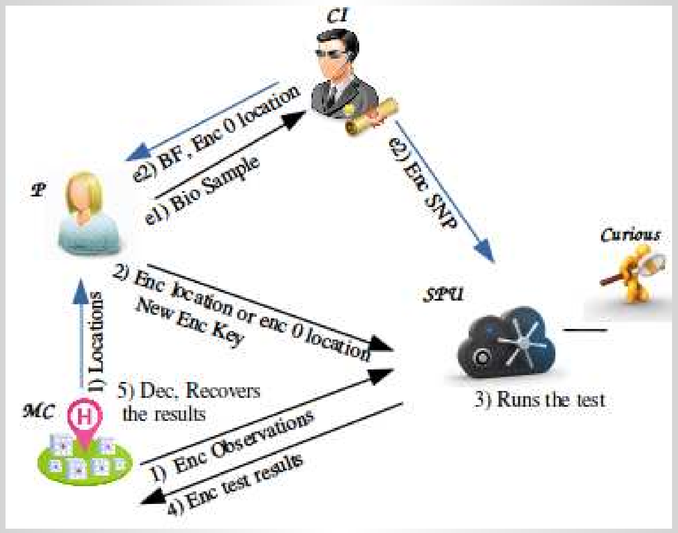
We developed a method for **Dynamic Privacy-Preserving Genomic Susceptibility Testing**
SNP (Single Nucleotid Polymorphisms) represent locations in an individual genomic string where simple variants with respect to the reference DNA sequence occur.



SNPs are suitable for running susceptibility tests, as certain SNPs hold a high correlation with determined diseases

$$S_{E_P}^{P,x} = \frac{1}{\sum_{i\in\Omega_x} c_{E_P}^{x,i}} \times \left\{ \Sigma_{i\in\Omega_x} c_{E_P}^{x,i} \{ \frac{pr_0^{x,i}{}_{E_P}}{0-1} [\mathrm{SNP}_{E_P}^{P,i} - 1] + \frac{pr_1^{x,i}{}_{E_P}}{1-0} [\mathrm{SNP}_{E_P}^{P,i} - 0] \} \right\}.$$

We proposed a novel protocol where the $\mathcal{SPU}$ calculates the susceptibility test function without having access to the clear-text genomic data of the patient.

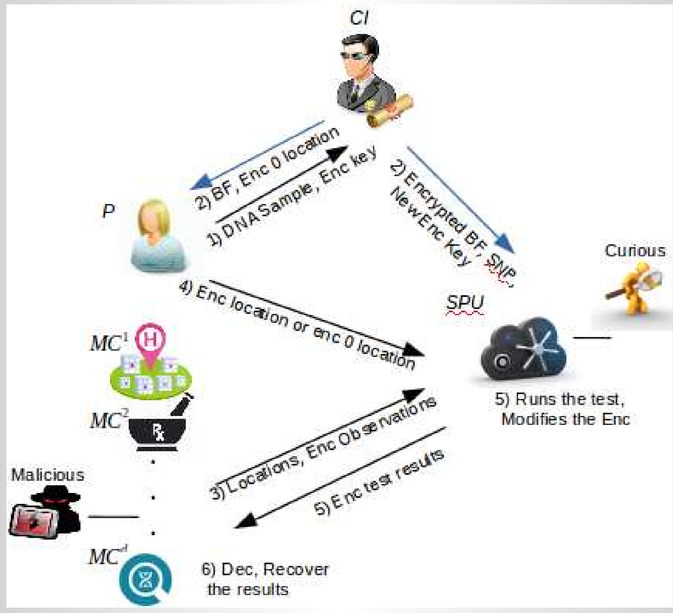


Our contributions with respect to prior works are:

- Reducing the overhead of Patient and medical centers
- Moving the bulk of the computation workload to $\mathcal{SPU}$
- Applying Somewhat Homomorphic Encryption enables multiplications by know values and encrypted values
- Achieving higher efficiency, less round complexity, and more privacy
- We also sketched the method to work with different medical units with different access roles to the data

This work has been accepted at an international conference [5].

As future work, we produced a sketch of the scheme working with an attribute based homomorphic encryption scheme, which preserves homomorphic properties and features an "inherent" access control through attributes, enforcing the patient's access policy referred to the different medical centers' attributes.



5. Temporal Planning for the Next Year

There are three lines of work which will be followed during the next year:

Attribute-based homomorphic encryption applied to flexible and dynamically adapted genomic privacy-preserving processing

The following points will be addressed:

- Multi-server storing and processing unit
- Malicious and curious parties
- Flexible authentication methods
- Security proofs for different trust relations and attacker models

Quantification of data leakage in privacy-preserving genomic processing

The following points will be addressed:

- Advancing state of the art in information-theoretic leakage metrics in e-health
- Analyzing of the effect of generalization and noise addition techniques to genomic information
- Defining proper data leakage measurement for genomic information
- Combining cryptographic and anonymization techniques to fully protect genomic data

Developing Delegatable Anonymous Credentials for genomic access control

The following points will be addressed:

- Research on delegatable lattice encryption schemes
- Optimizing security-efficiency trade-off in homomorphic delegatable anonymous credentials
- Reducing the round-complexity achieving constant size

6. References

[1] M. Naveed, E. Ayday, E. W. Clayton, J. Fellay, C. A. Gunter, J.-P. Hubaux, B. A. Malin, and X. Wang. Privacy in the genomic era. ACM Computing Surveys (CSUR), 48(1):6, 2015.

[2] Z. Brakerski, C. Gentry, and V. Vaikuntanathan. (leveled) fully homomorphic encryption without bootstrapping. In 3rd Innovations in Theoretical Computer Science Conference, pages 309-325. ACM, 2012.

[3] M. Canim, M. Kantarcioglu, and B. Malin. Secure management of biomedical data with cryptographic hardware. IEEE Trans. on Information Technology in Biomedicine, 16(1):166-175, 2012.

[4] C. Gentry, A. Sahai, and B. Waters. Homomorphic encryption from learning with errors: Conceptually-simpler, asymptotically-faster, attribute-based. In Advances in Cryptology-CRYPTO 2013, pages 75-92. Springer, 2013.

[5] M. Namazi, J. Troncoso-Pastoriza, F Pérez-González. Dynamic Privacy Preserving Susceptibilty Testing. In Information Hiding & Multimedia Security, 2016