UniversidadeVigo
Signal Theory and
Communications
Department
University of Vigo
Spain

Secure Signal Processing for Genomic Privacy Protection

Mina Namazi

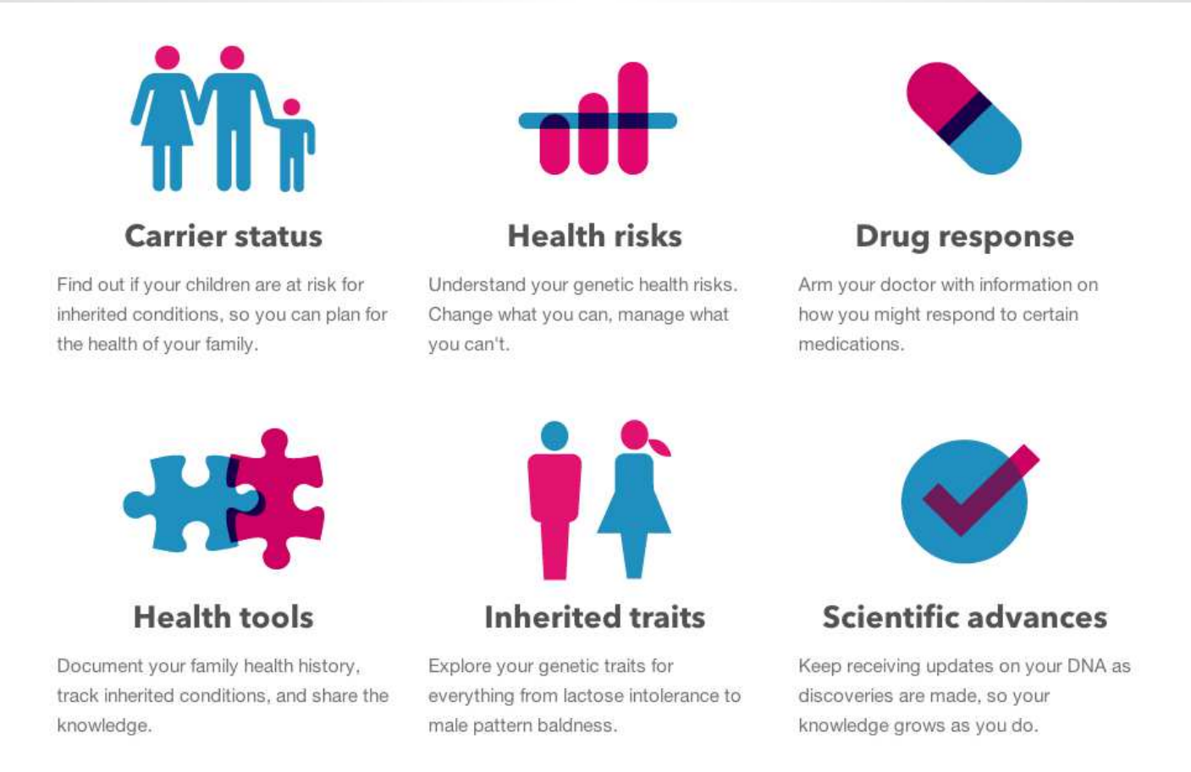
Advisors: Juan Ramón Troncoso-Pastoriza and Fernando Pérez-González

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

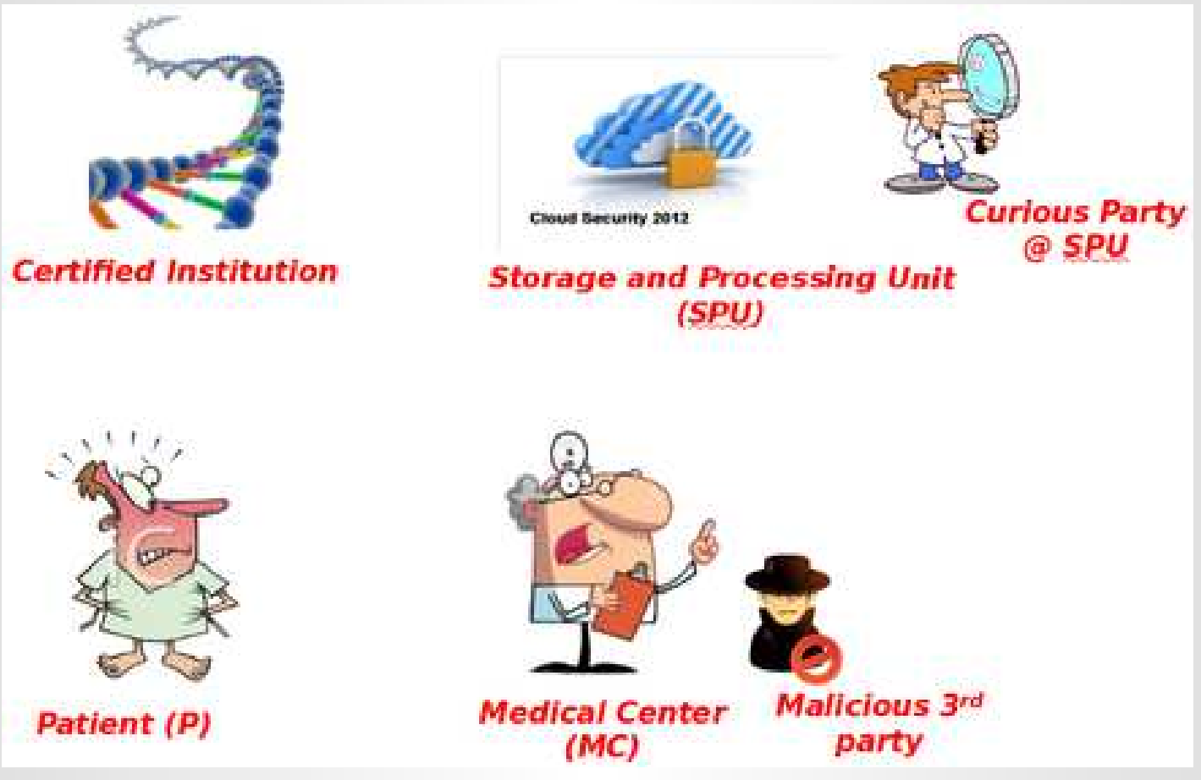
Workshop on Monitoring PhD Student Progress. June 22, 2017

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 and need widespread use of outsourced processing. The benefits of an extensive study of genomic data in advancing medicine research are unquestionable.


Carrier status
Find out if your children are at risk for inherited conditions, so you can plan for the health of your family.
Health risks
Understand your genetic health risks. Change what you can, manage what you can't.
Drug response
Arm your doctor with information on how you might respond to certain medications.
Health tools
Document your family health history, track inherited conditions, and share the knowledge.
Inherited traits
Explore your genetic traits for everything from lactose intolerance to male pattern baldness.
Scientific advances
Keep receiving updates on your DNA as discoveries are made, so your knowledge grows as you do.

The sensitive nature of the genome entails severe privacy risks when the sequences are outsourced to an untrustworthy environment, like a Cloud service and makes them vulnerable to attacks and accesses violating patient's privacy.

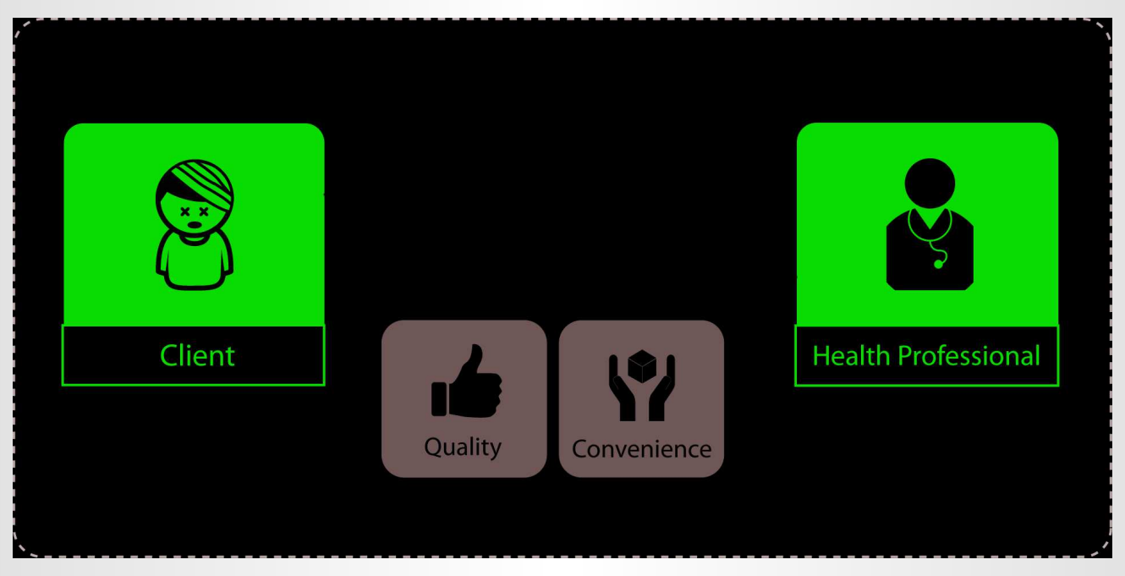


A proper protection of the genomic data by solely applying anonymization techniques is proven to be infeasible, unless the data are rendered useless. Moreover, genomic information can be linked to ancestors and relatives of an individual, so its leakage also affects their privacy. Therefore, a combination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

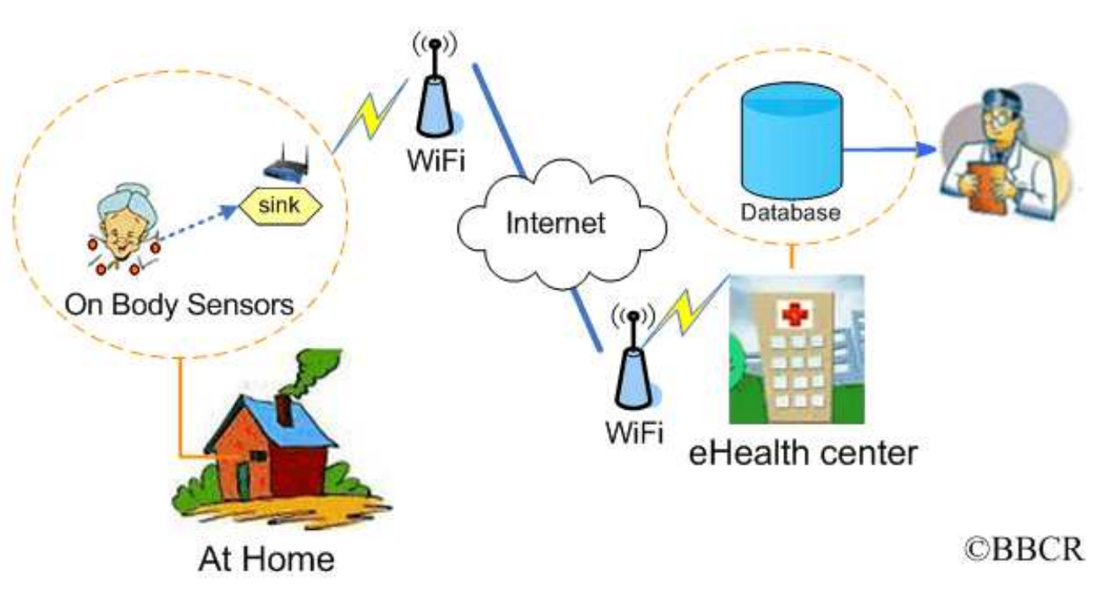
2. Thesis Objectives

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:

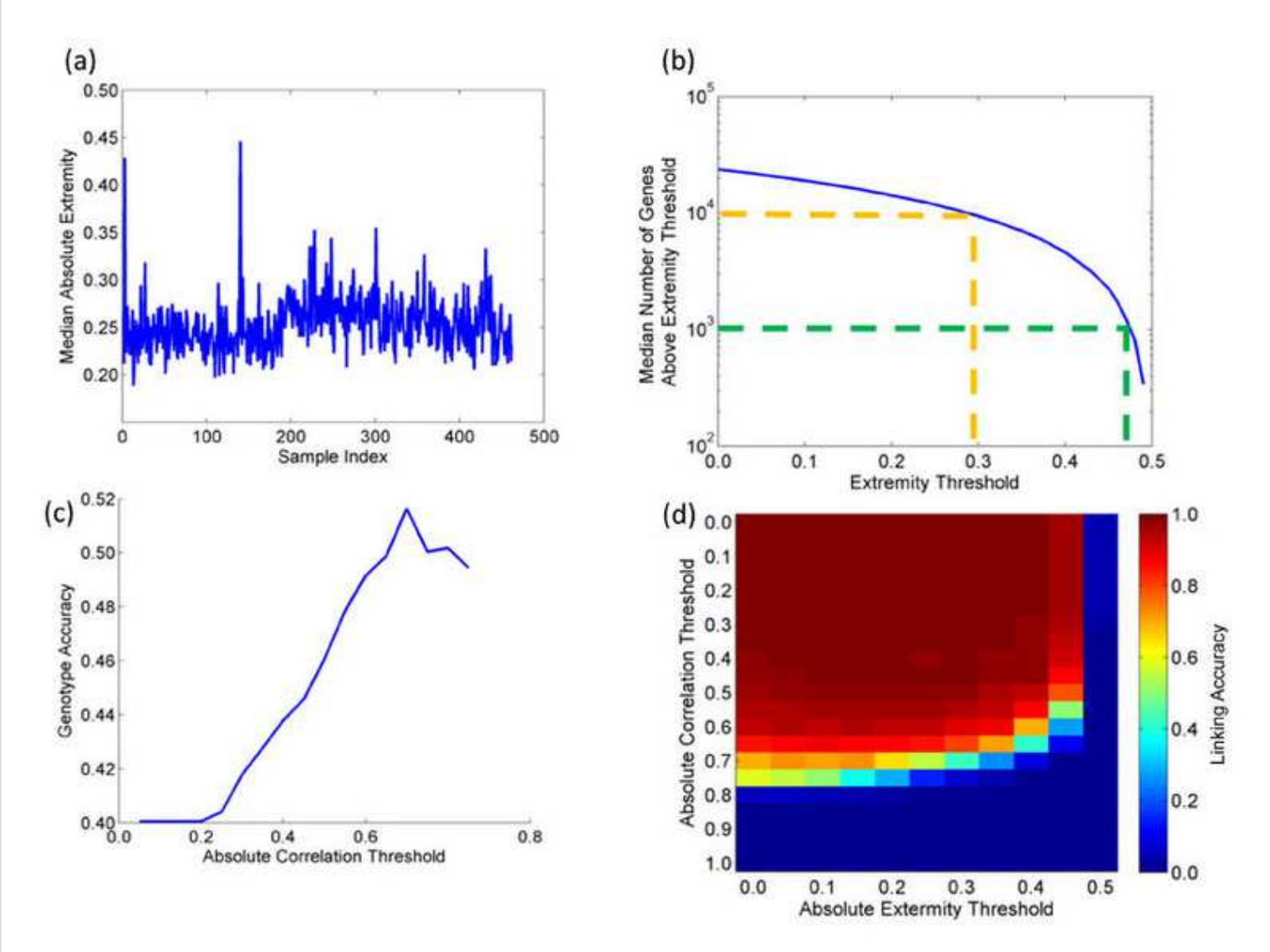
A. Analyzing existing schemes and techniques for secure signal processing e-Health applications from a privacy and security point of view.



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

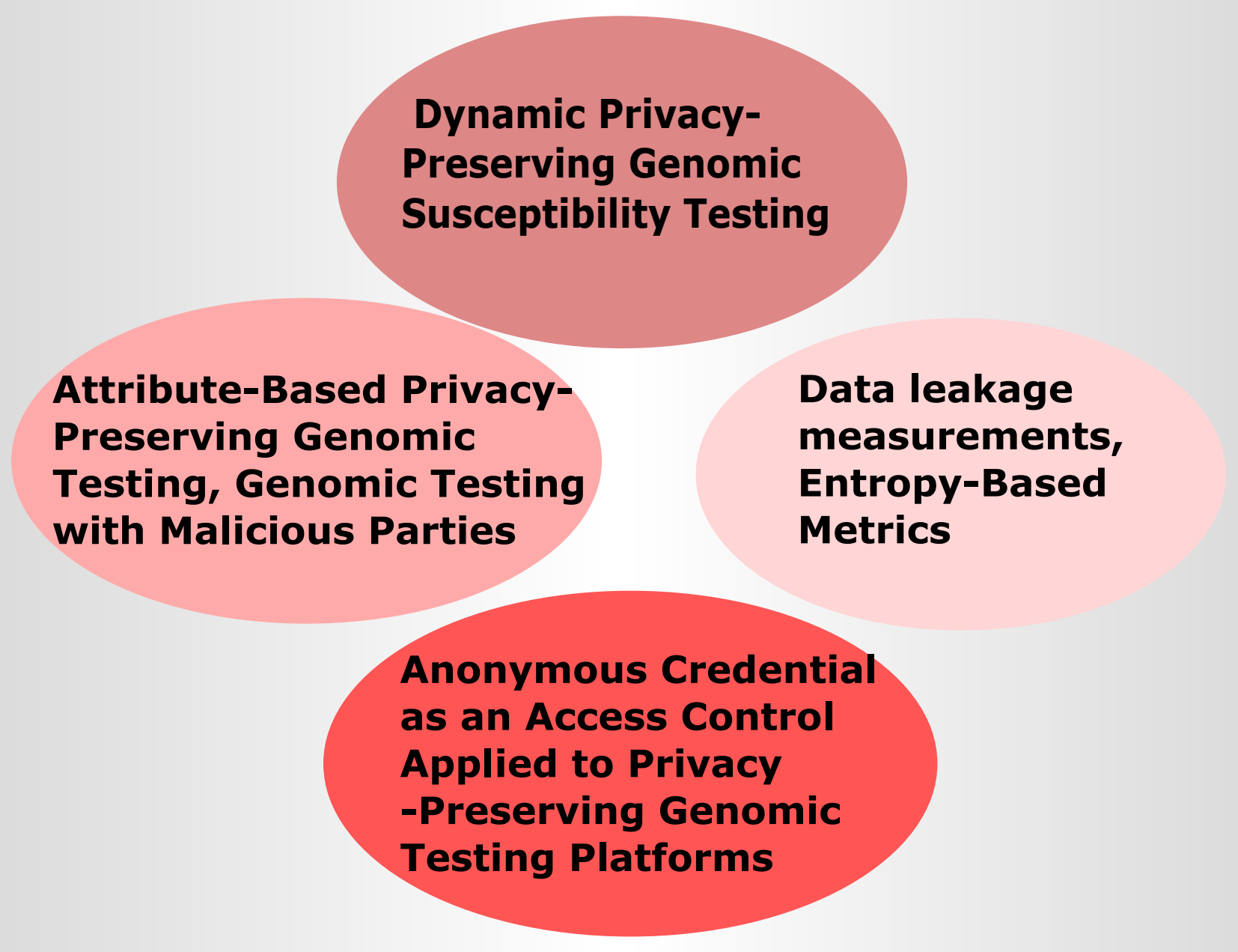


C. Devising new information-theoretic metrics to quantify the information leakage on genomic data when it is partially protected or when the results of several subsequent processes are disclosed.



3. Research Plan

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

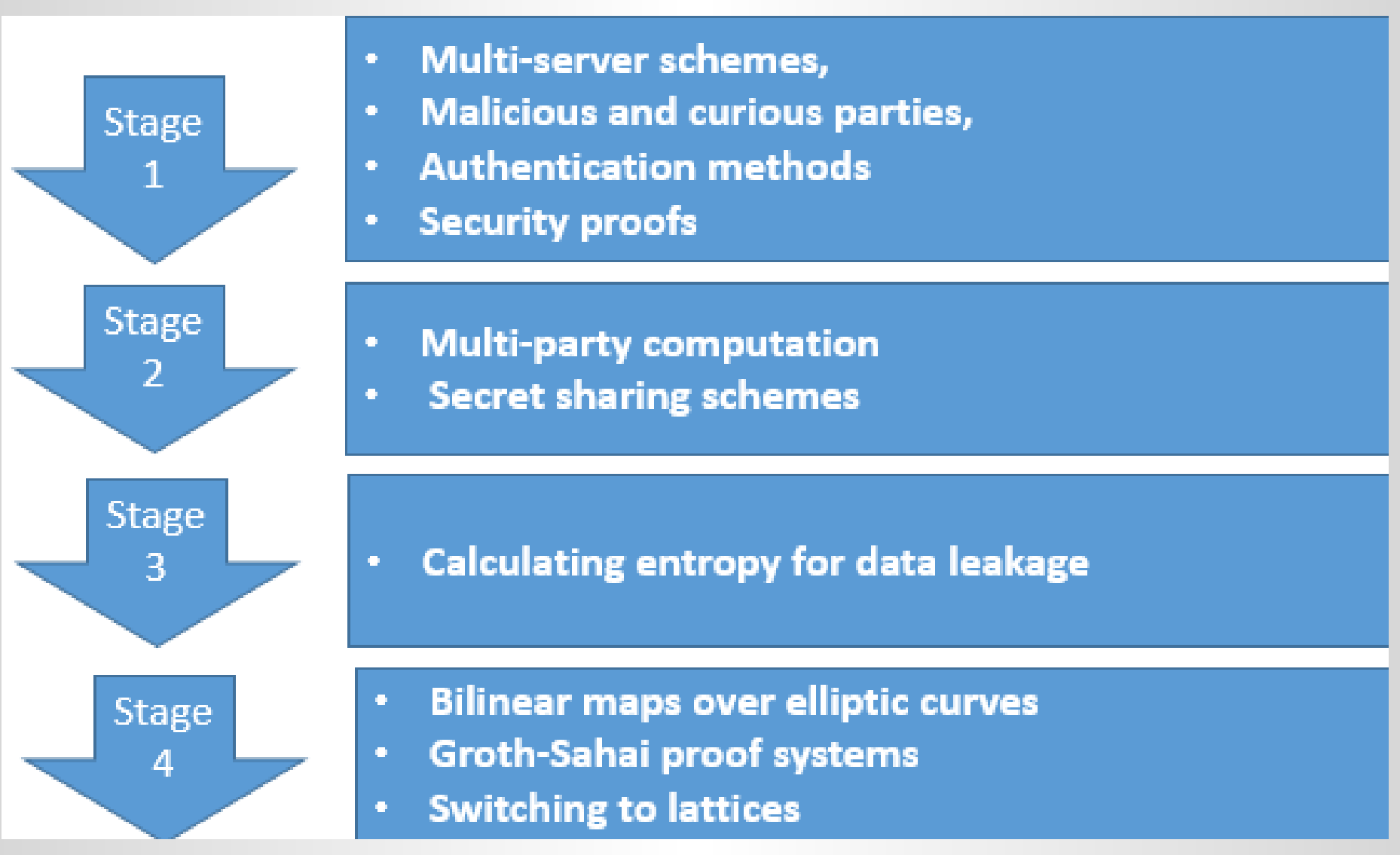

Dynamic Privacy-Preserving Genomic Susceptibility Testing

Attribute-Based Privacy-Preserving Genomic Testing, Genomic Testing with Malicious Parties

Data leakage measurements, Entropy-Based Metrics

Anonymous Credential as an Access Control Applied to Privacy -Preserving Genomic Testing Platforms

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


Stage 1
- Multi-server schemes,
- Malicious and curious parties,
- Authentication methods
- Security proofs

Stage 2
- Multi-party computation
- Secret sharing schemes

Stage 3
- Calculating entropy for data leakage

Stage 4
- Bilinear maps over elliptic curves
- Groth-Sahai proof systems
- Switching to lattices

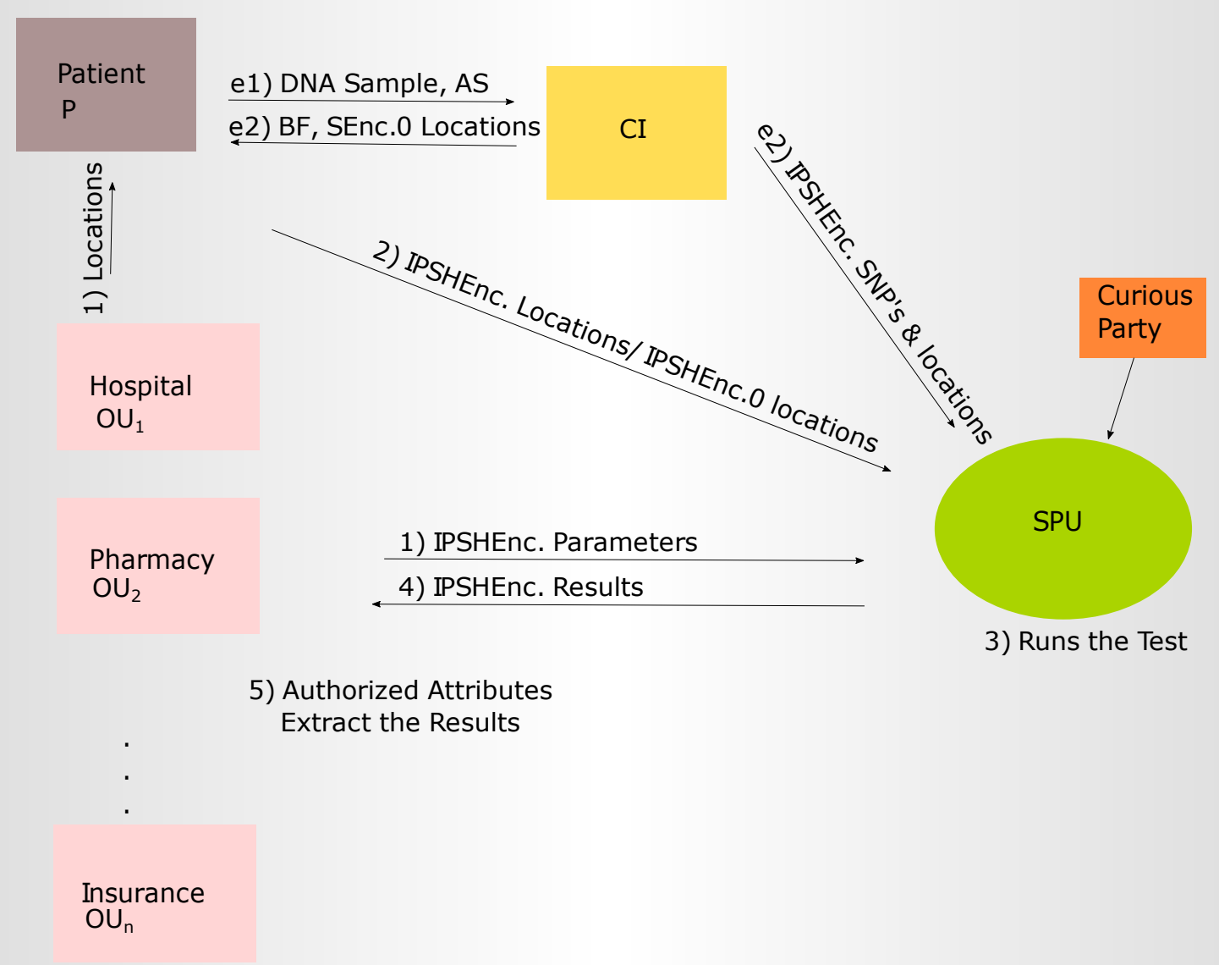
4. Results and Discussions

We developed a method for **Dynamic Privacy-Preserving Genomic Susceptibility Testing**
We proposed a novel protocol where a server calculates the susceptibility test function without having access to the clear-text genomic data of patients.

Our contributions with respect to prior approaches 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 public 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 presented at an international conference [5].
We developed a method for **Attribute-Based Genomic Susceptibility Testing**, 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.

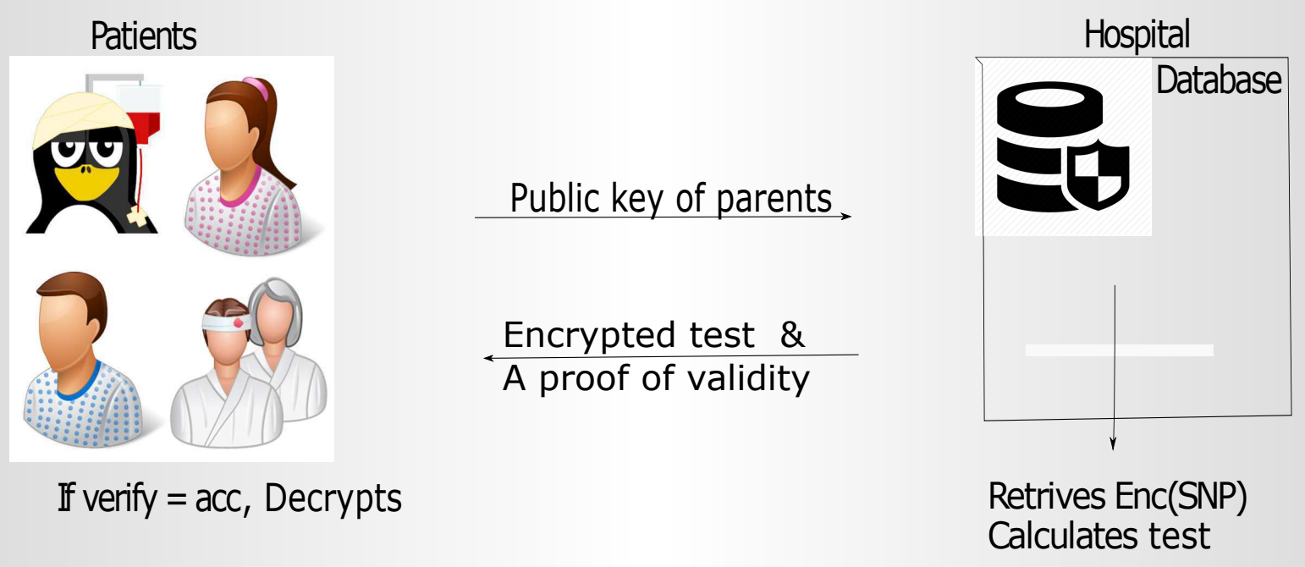


Our contributions with respect to prior approaches are:

- Reviewing possible cryptographic constructions to enhance an attribute-based homomorphic encryption scheme.
- We address not only the outsourced secure test analysis by increasing the privacy level, but also embed access control in the protocol itself.
- Although attributes increased the computation cost, this cost remain quasi-linear regarding number of operations; thanks to lattices, working in parallel with n plaintext values encrypted in one ciphertext is allowed in comparison to other schemes which can only deal with one scalar plaintext.
- Achieving more security due to applying highly secure cryptographic constructions as a core protocol.

We are developing an "entropy-based" analysis for **information-theoretic leakage metrics in e-health**.

We proposed a protocol for **Privacy-Preserving Genomic Testing with Malicious Parties** which applies a verifiable homomorphic encryption scheme and a proof system to resist against malicious parties running active attacks.


Patients
Public key of parents
Encrypted test & A proof of validity
Hospital
Database
If verify = acc, Decrypts
Retrives Enc(SNP) Calculates test

Our main contributions are as follows:

- Combining lattice properties with verifiable computation to enhance a homomorphic verifiable protocol.
- Enhancing a privacy-preserving test protocol which resists against active attacks attempted by users inside the protocol (server/patients).
- Data privacy while outsourcing and working on encrypted data, query privacy by securing data transactions against active attacks are guaranteed.
- Final results are verifiable.

We developed a **Delegatable Anonymous Credential with Constant Size** and implement it in a genomic testing platform as an access control method.
Our main contributions are listed as it follows:

- Obtaining a constant size protocol while delegating the credential to the other authorized parties in the protocol.
- Applying the developed protocol to control the accesses in genomic testing platforms.
- Approaching high privacy by bringing anonymity of the users to the genomic testing protocols.

5. Temporal Planning for the Next Year

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

Flexible and dynamically adapted genomic privacy-preserving processing protocols

The following points will be addressed:

- Multi-server storing and processing unit
- 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, ACM, 2016.