UNIVERSITÀ DI PISA

DIPARTIMENTO DI INFORMATICA

Corso di Laurea Magistrale in Informatica

TESI DI LAUREA MAGISTRALE

Data aggregation using Homomorphic Encryption in Mobile CrowdSensing context

Prof. Stefano Chessa
Dott. Michele Girolami

Chiara Boni
600159

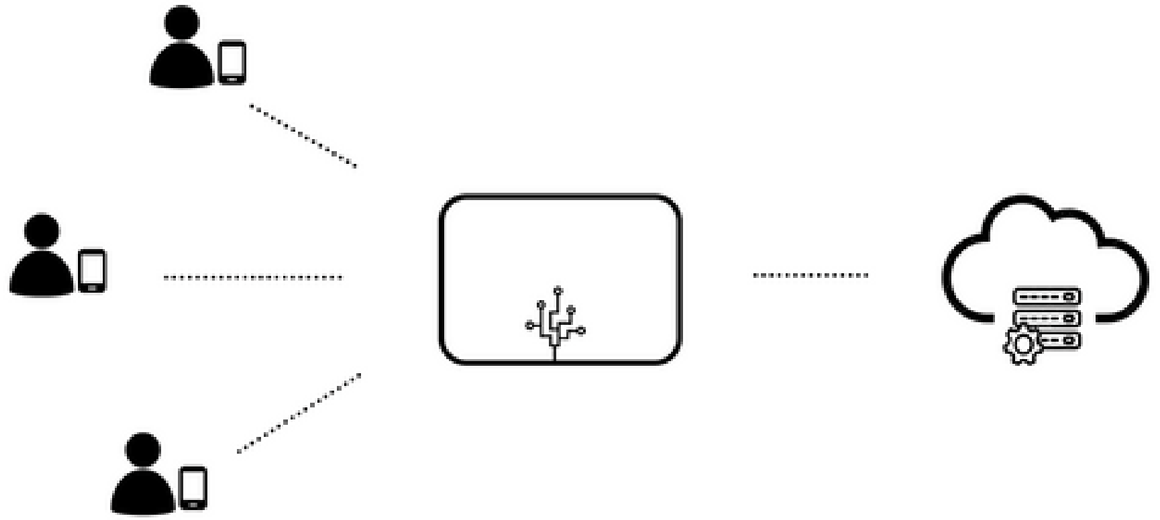
Mobile CrowdSensing

The study that relies on a network of mobile sensors, which are able to collect information on the environment

Tasks

User Profiling

Confidentiality



HOMOMORPHIC ENCRYPTION

Arithmetic operations directly on encrypted data, without the need to decrypt them

$$E(a \circ b) = E(a) \diamondsuit E(b)$$

CONFIDENTIALITY

REDUNDANT RESIDUE NUMBER SYSTEM

Redundancy is provided by encoding more residues than are necessary for representation

ROBUSTNESS

MOBILE CROWDSENSING

METHODOLOGY

THESIS OBJECTIVES

TECHNOLOGIES EMPLOYED

01

Study a MCS model, focusing on the privacy-preserving issues

02

Investigate homomorphic encryption into the MCS model

03

Simulate such research in a real scenario

04

Strengthen the data aggregation model by using redundancy with RRNS

03

Simulate such research in a real scenario

To test homomorphic encryption in an MCS context, the chosen example relies on modelling GPS coordinates to calculate the total distance covered

DATASET MANAGEMENT

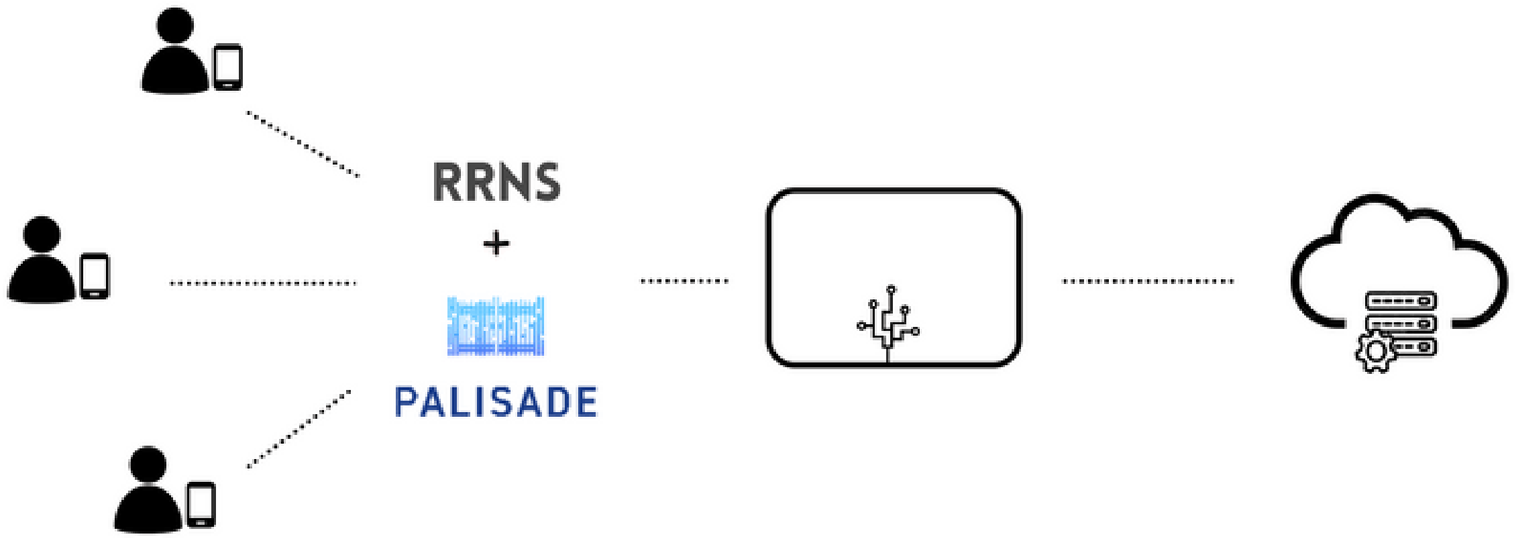
pandas GeoPy

HOMOMORPHIC ENCRYPTION

PALISADE

OVERVIEW

SCHEME IMPLEMENTATION



OVERVIEW

SCHEME IMPLEMENTATION

1

SETUP

Devices collect and encrypt data

2

SENDING

Devices calculate the RRNS representation of the ciphers and send it to the aggregator

3

AGGREGATION

Aggregator receives the ciphers and sums them

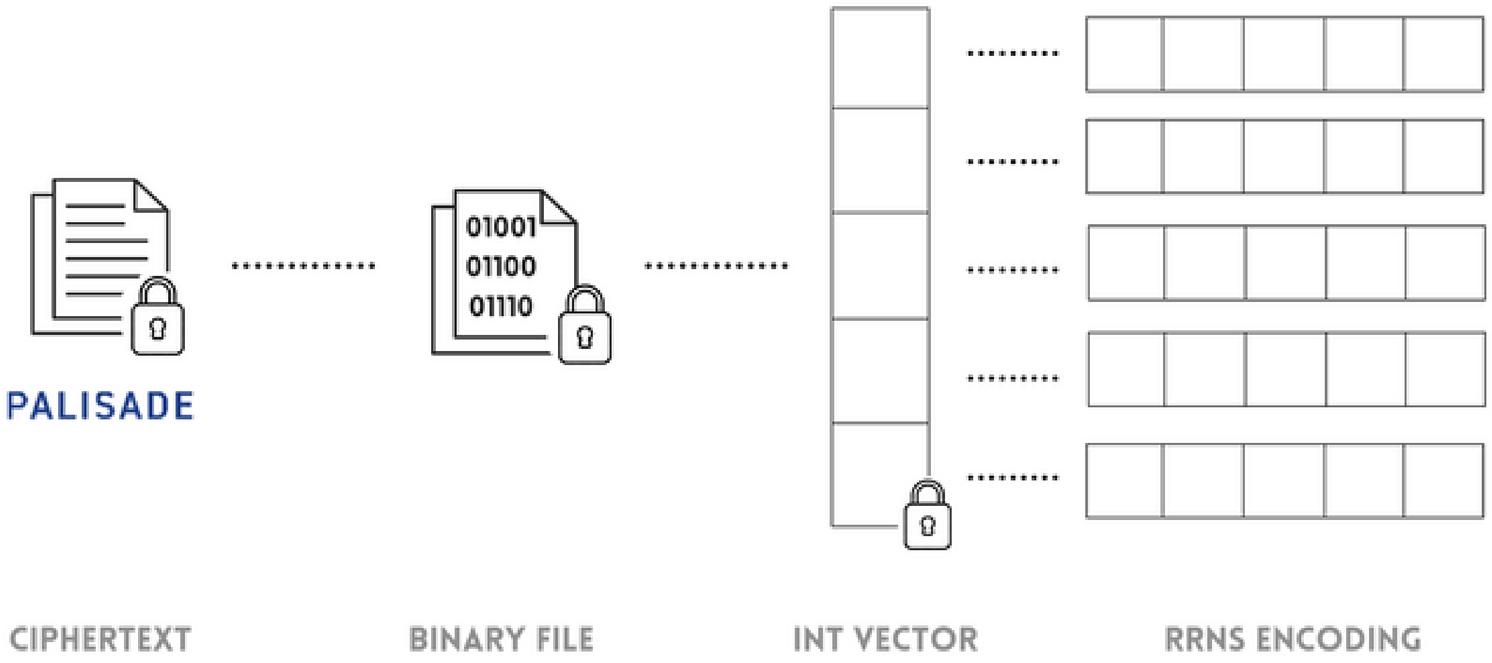
OVERVIEW

SCHEME IMPLEMENTATION

2

SENDING

RRNS ENCODING



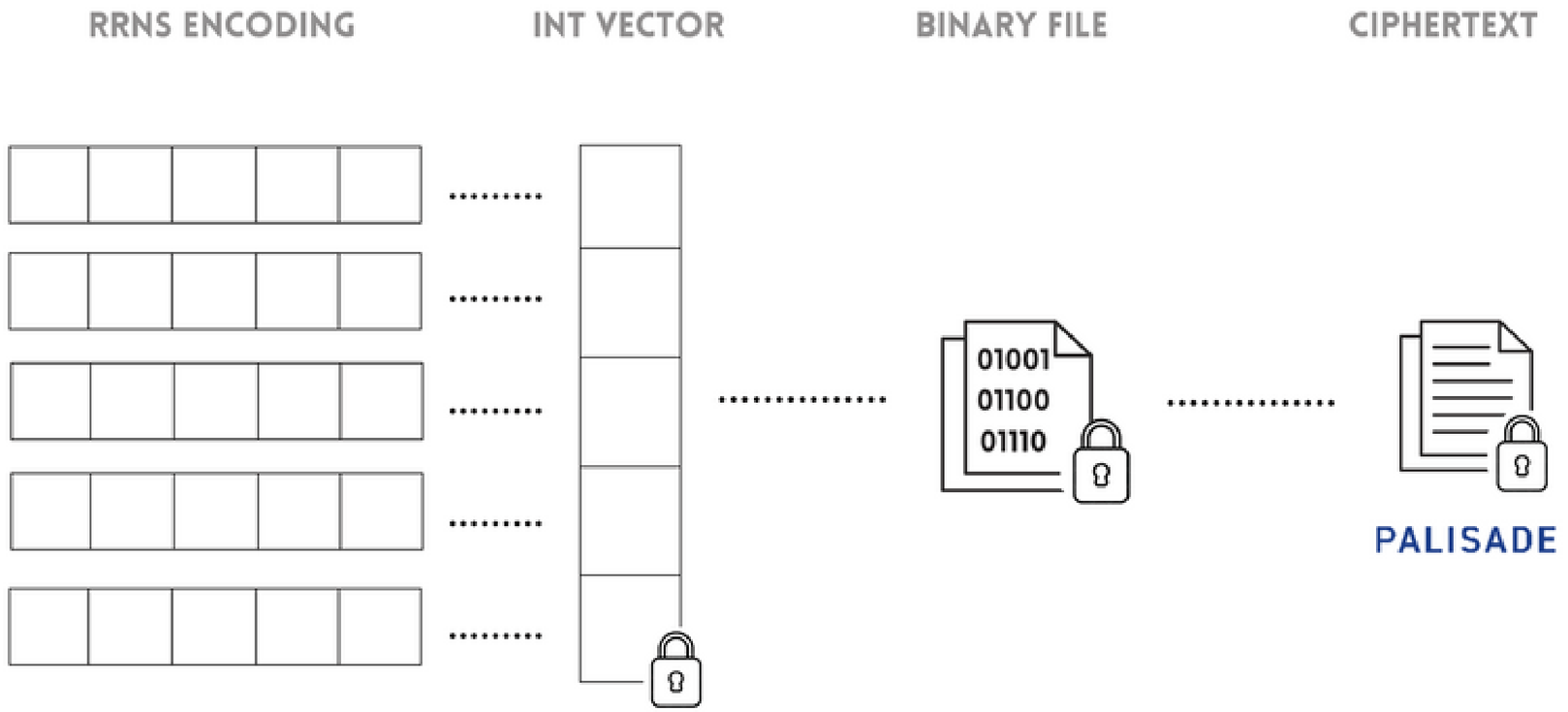
OVERVIEW

SCHEME IMPLEMENTATION

3

AGGREGATION

RRNS DECODING



OVERVIEW

SCHEME IMPLEMENTATION

3

AGGREGATION

SUM


```
// homomorphic addition

auto ciphertextResult = cryptoContext->EvalAdd(ciphertext1, ciphertext2);
```

CASE STUDY

DATASET PROCESSING

EVALUATIONS

The real example case concerns the manipulation of GPS coordinates, in order to calculate the total distance traveled by a group of users

GeoLife GPS Trajectories

This GPS trajectory dataset was collected in (Microsoft Research Asia) Geolife project by 182 users in a period of over three years (from April 2007 to August 2012)....

Microsoft Download Center

lat	lon	tid	uid	date_time

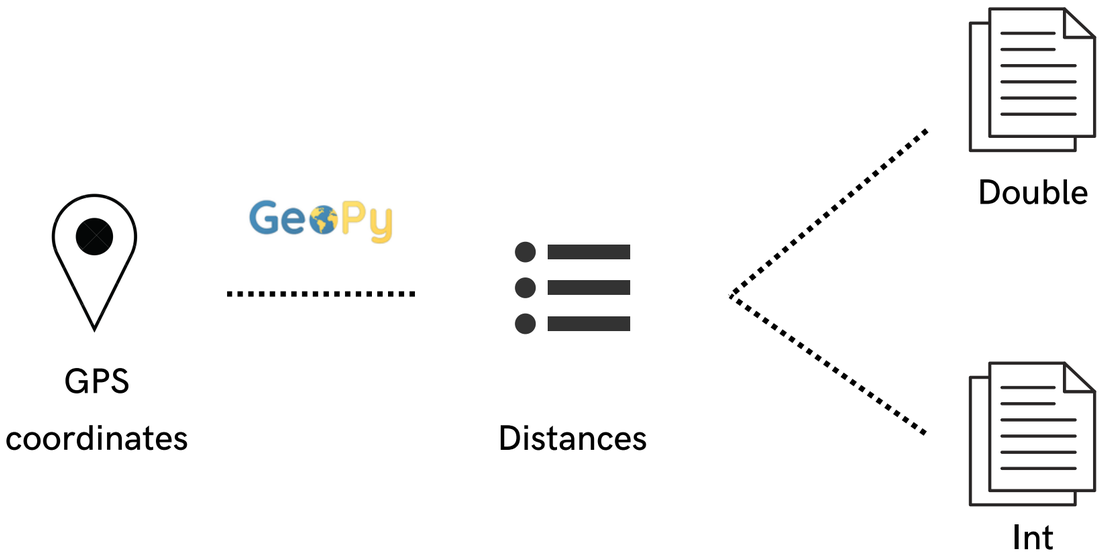
Dataset Columns

pandas

CASE STUDY

DATASET PROCESSING

EVALUATIONS

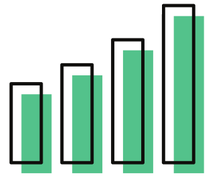


Taking a trajectory, distance calculation is performed between pairs of its consecutive points

765.041 distances

Correctness

Despite the introduction of a over-structure, such as the one given by RRNS encoding, the system is still able to produce correct results



Efficiency

The analysis which concerns the correlation between the choice of parameters and the execution time

CASE STUDY

DATASET PROCESSING

EVALUATIONS

*average of 100 runs

				BGV Integer Scheme							
				Sending Time*				Aggregator Time*			
Multipl. Depth	Chunk Size	Plaintext Modulus**	# Modules RNS Base	Wall Time (s)	Conf. Interv.	CPU Time (s)	Conf. Interv.	Wall Time (s)	Conf. Interv.	CPU Time (s)	Conf. Interv.
1	5.000	65.537	12	3,15699	±0,004	8,44010	±0,045	3,50408	±0,014	3,61340	±0,008
2	5.000	65.537	12	7,56084	±0,021	25,47307	±1,012	7,08278	±0,019	7,32784	±0,020
0	5.000	65.537	12	0,89403	±0,008	3,00979	±0,058	0,93209	±0,003	0,98195	±0,022
1	10.000	65.537	12	1,63603	±0,013	4,63063	±0,066	1,76003	±0,005	1,84178	±0,021
1	50.000	65.537	12	0,40069	±0,002	1,20935	±0,030	0,37179	±0,004	0,42848	±0,036

CASE STUDY

DATASET PROCESSING

EVALUATIONS

*average of 100 runs

Multipl. Depth	Chunk Size	Plaintext Modulus**	# Modules RNS Base	Sending Time*				Aggregator Time*			
				Wall Time (s)	Conf. Interv.	CPU Time (s)	Conf. Interv.	Wall Time (s)	Conf. Interv.	CPU Time (s)	Conf. Interv.
1	5.000	49.153	12	3,22643	±0,012	8,83064	±0,084	3,51076	±0,008	3,68469	±0,028
1	5.000	73.729	12	3,23130	±0,011	8,81470	±0,076	3,51627	±0,008	3,63185	±0,012
1	5.000	65.537	8	2,83338	±0,019	8,36826	±0,064	3,38839	±0,012	3,60736	±0,036
1	5.000	65.537	14	3,56004	±0,065	9,83674	±0,281	3,52046	±0,009	3,62886	±0,012
0	50.000	65.537	8	0,11994	±0,0004	0,45632	±0,006	0,09620	±0,001	0,09477	±0,001

Best Run

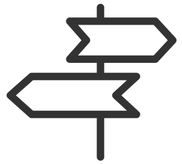
CONCLUSIONS

The objective was to analyze a MCS architecture and focus on the confidentiality limit, addressing it via homomorphic encryption

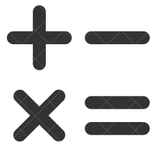
Increasing robustness by inserting redundancy when sending data, from the devices to the aggregator, using RRNS encoding

This type of encryption is expensive: it requires powerful sensors such as smartphones

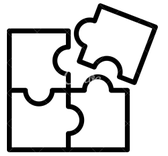
FUTURE WORK



Analyze how different libraries encrypt data, construct operations and behave computationally



Extend the use scenario to other homomorphic operations and contexts



Combining HE and RRNS to aggregate on the encrypted fragments instead of having to reconstruct the encrypted data

REFERENCES

ChiaraBn/**Master-Thesis**

1 Contributor · 0 Issues · 0 Stars · 0 Forks

ChiaraBn/Master-Thesis

Contribute to ChiaraBn/Master-Thesis development by creating an account on GitHub.

Homomorphic Encryption for Arithmetic of Approximate...

We suggest a method to construct a homomorphic encryption scheme for approximate arithmetic. It supports an...

link.springer.com

PALISADE

Files · release-v1.11.2 · PALISADE / PALISADE Development · GitLab

This is the development repository of the PALISADE lattice cryptography library. The current development version is...

GitLab

IEEE Xplore®

Implementation and Performance Evaluation of RNS Variants of the BFV...

Homomorphic encryption is an emerging form of encryption that provides the ability to compute on...

ieee.org

GeoLife GPS Trajectories

This GPS trajectory dataset was collected in (Microsoft Research Asia) Geolife project by 182 users in a period of over three years (from April 2007 to August 2012). Last...

Microsoft Download Center

CiteSeerX =7M

CiteSeerX — Fully homomorphic encryption without bootstrapping

CiteSeerX - Document Details (Isaac Councill, Lee Giles, Pradeep Teregowda): We present a radically new...

ist.psu.edu

(PDF) A reliable and energy efficient IoT data transmission...

PDF | On Jun 1, 2015, Chinmaya Mahapatra and others published A...

ResearchGate

CASE STUDY

DATASET PROCESSING

EVALUATIONS

*average of 100 runs

				CKKS Approximate Scheme							
				Sending Time*				Aggregator Time*			
Multipl. Depth	Chunk Size	Scale Factor	# Modules RNS Base	Wall Time (s)	Conf. Interv.	CPU Time (s)	Conf. Interv.	Wall Time (s)	Conf. Interv.	CPU Time (s)	Conf. Interv.
1	5.000	50	12	3,67258	±0,033	13,53826	±0,023	3,94596	±0,011	8,48766	±0,034
2	5.000	50	12	7,51354	±0,031	30,12514	±0,119	7,88696	±0,016	14,93852	±0,035
1	10.000	50	12	1,93718	±0,010	7,50640	±0,166	1,98382	±0,008	4,24750	±0,059
1	50.000	50	12	0,53840	±0,003	1,98702	±0,010	0,41315	±0,001	0,78828	±0,007

CASE STUDY

DATASET PROCESSING

EVALUATIONS

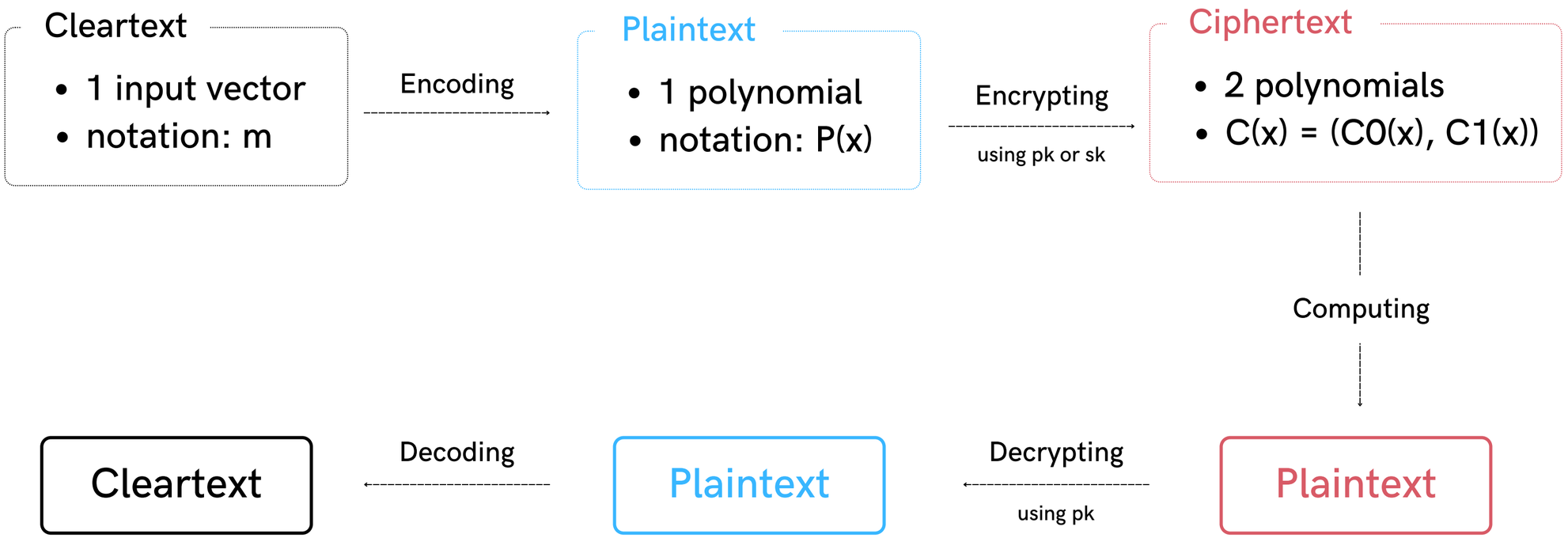
*average of 100 runs

				Sending Time*				Aggregator Time*			
Multipl. Depth	Chunk Size	Scale Factor	# Modules RNS Base	Wall Time (s)	Conf. Interv.	CPU Time (s)	Conf. Interv.	Wall Time (s)	Conf. Interv.	CPU Time (s)	Conf. Interv.
1	5.000	20	12	3,68496	±0,014	13,52364	±0,033	3,99427	±0,034	8,50812	±0,062
1	5.000	50	8	3,29544	±0,024	13,40193	±0,068	3,76236	±0,006	8,37390	±0,051
1	5.000	50	14	3,85562	±0,017	13,95260	±0,118	3,93122	±0,006	8,50410	±0,036
1	50.000	50	8	0,49035	±0,005	1,93060	±0,007	0,39361	±0,001	0,76717	±0,004

Best Run

REDUNDANT
RESIDUE
NUMBER SYSTEM

HOMOMORPHIC ENCRYPTION

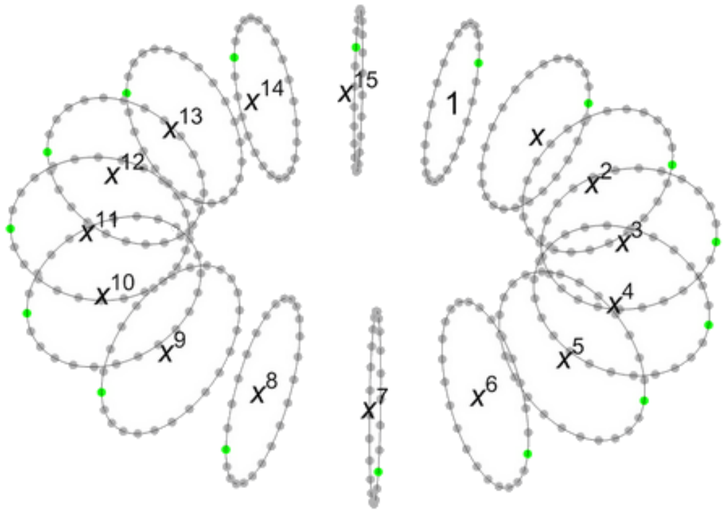


REDUNDANT RESIDUE NUMBER SYSTEM

These models employ complex structures such as polynomial rings, i.e. a ring formed from the set of polynomials in one or more variables, with coefficients in another ring, often a field.
Such ring is defined as $R = Z[X]/(X^n + 1)$.

It can be parameterizable the size of each ring.

HOMOMORPHIC ENCRYPTION



Plaintext example with n = 16

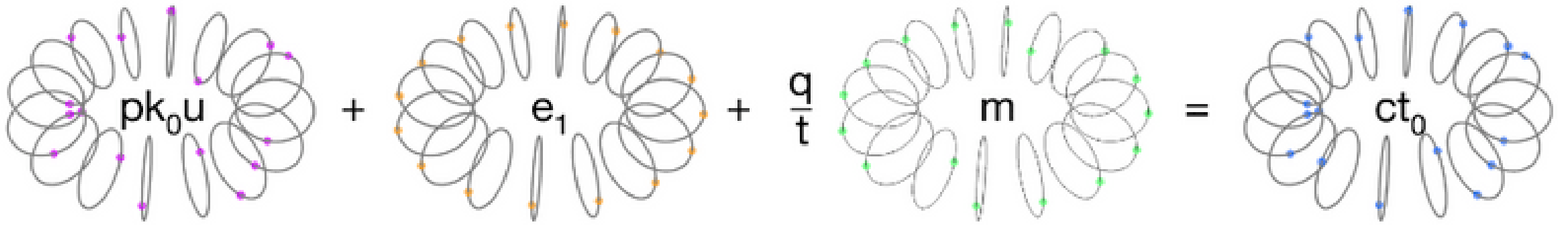
https://humanata.com/blog/illustrated_primer/

REDUNDANT
RESIDUE
NUMBER SYSTEM

HOMOMORPHIC ENCRYPTION

The ciphertext is represented by two polynomials calculated as:

$$ct = ([pk0\ u + e1 + qm/t]q\ ,\ [pk1\ u + e2\ ]q\ )$$



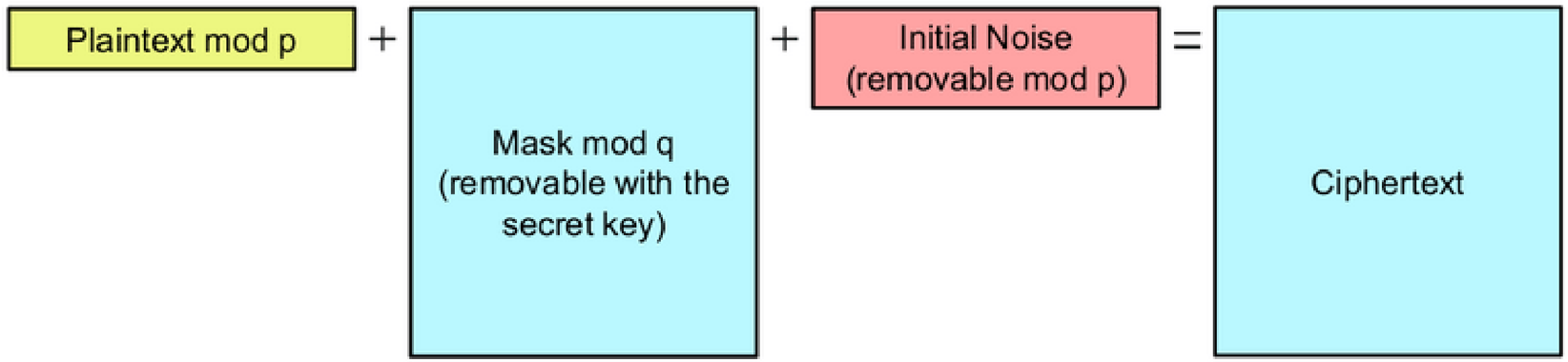
First part of the cipher

https://humanata.com/blog/illustrated_primer/

REDUNDANT
RESIDUE
NUMBER SYSTEM

HOMOMORPHIC ENCRYPTION

FRESH ENCRYPTION



- Horizontal: each coefficient in a polynomial or in a vector.
- Vertical: size of coefficients.
- Initial noise is small in terms of coefficients' size.

REDUNDANT
RESIDUE
NUMBER SYSTEM

HOMOMORPHIC
ENCRYPTION

AFTER SOME COMPUTATIONS