Aadhaar Based Social Credit System for India

Under Mr. Saibal Pal , DRDO

Submitted by
Bishal Kumar Patel

Contents

1 Acknowledgement

I pay my sincere thanks to Mr. Saibal Pal who permitted to work under him and guided me throughout the period of internship. Without his help and guidance it would not have been possible to complete this project.

2 Cryptography

Cryptography is a process in which the data is converted into unintelligible text, transmitted, then converted back into a comprehensible form so it can only be read and processed by whom it's intended.

Cryptography is associated with the process of preparing protected and secure data for communication.

There are three main types of cryptography

- Symmetric Key Cryptography
- Asymmetric Key Cryptography
- Cryptographic Protocols

3 Symmetric Key Cryptography

Symmetric encryption, also referred to as conventional encryption or single-key encryption, was the only type of encryption in use prior to the development of public key encryption in the 1970s.

Here a single key is used for both encryption and decryption. It is used for bulk data transmission.

3.1 Substitution Cipher

In simple substitution ciphers, an alphabet or symbol is substituted for each single or multiple letters. The earliest known, and the simplest, use of a substitution cipher was by Julius Caesar. To express Caesar's cipher mathematically, first replace each letter by an integer from 1 to 26, based on its position in the alphabet. That is 'A' is replaced by 1, 'B' by 2, 'C' by 3 and so on. Caesar cipher is represented by function f that has non-negative integer value 'a', where 'a' less than or equal to 26 [3], [5].

$$f\ (\text{a}) = (\text{a}+3) \bmod 26$$

Thus, when the letter 'a' is substituted, d is used, and when 'b' is to be written, 'e' is used and so on. The letters wrap around at the end of the alphabet. So, if a person wants to encipher 'z', it is written as 'c' by taking modular 26. Similarly, 'y' is written as 'b'. The entire cipher is represented by two rows of letters. These rows are called a lookup table.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

When someone wants to encrypt a word, he or she looks up the original letter in the top row and uses the corresponding cipher text letter in the bottom row. For example, secret message produced from the message "ATTACK THE HILL" using the Caesar cipher is "DWWDFN WKH KMOO".

While the above substitution cipher is easy to remember, it is also easily breakable. To make a substitution cipher more complex, multiple substitutions and sometimes even numbers are added to the cipher.

3.2 Shift Cipher

It is actually a special case of the substitution cipher and has a very elegant mathematical description.

The shift cipher itself is extremely simple: We simply shift every plain text letter by a fixed number of positions in the alphabet. For instance, if we shift by 3 positions, A would be substituted by d, B by e, etc.

The shift cipher also has an elegant description using modular arithmetic. For the mathematical statement of the cipher, the letters of the alphabet are encoded as numbers, as depicted below.

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25

Definition: Shift Cipher
Let x,y,k belong to First 26 Natural numbers.
Encryption: e_k(x) = x + k mod 26.
Decryption: d_k(y) = y - k mod 26.

Example .Let the key be k = 17, and the plain text is:
ATTACK = x1,x2, . . . ,x6 = 0,19,19,0,2,10.

The ciphertext is then computed as
y1,y2, . . . ,y6 = 17,10,10,17,19,1 = rkkrtb

3.2.1 Program to implement Shift Cipher

Encryption:

```
string encrypt(string s, int key){
    string result;
    for(int i=0; i<s.size(); i++){
        //for uppercase
        if(isupper(s[i]))
            result += (s[i]+key-'A')%26+'A';
        //for lowercase
        else
            result += (s[i]+key-'a')%26+'a';
    }
    return result;
}
```

Decryption:

```
string decrypt(string s, int key){
    string result;
    for(int i=0; i<s.length(); i++){
        //for uppercase
        if(isupper(s[i]))
            result += (s[i]-key-'A'+26)%26+'A';
```

```
        //for lowercase
        else
            result += (s[i]-key-'a'+26)%26+'a';
    }
    return result;
}
```

3.2.2 Program for Integration of Dictionary with Shift cipher

```
def encrypt(text,s):
    result = ""
    for i in range(len(text)):
        char = text[i]
        if (char.isupper()):
            result += chr((ord(char) + s-65) % 26 + 65)
        else:
            result += chr((ord(char) + s - 97) % 26 + 97)
    return result

text = "english"
s = 4
print ("Text  : " + text)
print ("Shift : " + str(s))
print ("Cipher: " + encrypt(text,s))
```

Output:
Text : english
Shift : 4
Cipher: irkpmwl

```
def cipher_decrypt(ciphertext, key):
    decrypted = ""
    for c in ciphertext:
        if c.islower():
            c_index = ord(c) - ord('a')
            c_og_pos = (c_index - key) % 26 + ord('a')
            c_og = chr(c_og_pos)
            decrypted += c_og
        else:
            decrypted += c
    return decrypted
cryptic_text = "irkpmwl"
dict = {}
for i in range(0,26):
    plain_text = cipher_decrypt(cryptic_text, i)
    dict[i] = plain_text
    print("For key {}, decrypted text: {}".format(i, plain_text))
```

Output:
For key 0, decrypted text: irkpmwl

For key 1, decrypted text: hqjolvk
For key 2, decrypted text: gpinkuj
For key 3, decrypted text: fohmjti
For key 4, decrypted text: english
For key 5, decrypted text: dmfkhrg
For key 6, decrypted text: clejgqf
For key 7, decrypted text: bkdifpe
For key 8, decrypted text: ajcheod
For key 9, decrypted text: zibgdnc
For key 10, decrypted text: yhafcmb
For key 11, decrypted text: xgzebla
For key 12, decrypted text: wfydakz
For key 13, decrypted text: vexczjy
For key 14, decrypted text: udwbyix
For key 15, decrypted text: tcvaxhw
For key 16, decrypted text: sbuzwgv
For key 17, decrypted text: ratyvfu
For key 18, decrypted text: qzsxuet
For key 19, decrypted text: pyrwtds
For key 20, decrypted text: oxqvscr
For key 21, decrypted text: nwpurbq
For key 22, decrypted text: mvotqap
For key 23, decrypted text: lunspzo
For key 24, decrypted text: ktmroyn
For key 25, decrypted text: jslqnxm

```
from PyDictionary import PyDictionary
dictionary = PyDictionary()
for i in dict:
  if (dictionary.meaning(str(dict.get(i)))==None):
    pass
  else:
    print("Shift Key = " + str(i))
    print("Decrypted word is : " +dict.get(i))
    print("Meaning of decrypted word is :" + str((dictionary.meaning(str(dict.get(i))))))
    break
```

Output:
Shift Key = 4
Decrypted word is : english
Meaning of decrypted word is :'Noun': ['an Indo-European language belonging to the West Germanic branch; a language spoken in Britain and the United States and most of the commonwealth countries', 'the people of England', 'the discipline that studies the English language and literature', '(sports'], 'Adjective': ['of or relating to or characteristic of England or its culture or people', 'of or relating to the English language']

3.3 Affine Cipher

Its the improved version of shift cipher. It is slightly different to the other examples encountered here, since the encryption process is substantially mathematical. The whole process relies

on working modulo m (the length of the alphabet used). By performing a calculation on the plain text letters, we encipher the plain text.

Definition: Affine Cipher
Let x,y,k belong to First 26 Natural numbers.
Encryption: e_k(x) = y = a · x+b mod 26.
Decryption: d_k(y) = x = 1/a · y -b mod 26.
with the key: k = (a,b), which has the restriction: gcd(a,26) = 1.

The restriction gcd(a,26) = 1 stems from the fact that the key parameter a needs to be inverted for decryption. Element a and the modulus must be relatively prime for the inverse of a to exist. Thus, a must be in the set:

$$a \in \{1, 3, 5, 7, 9, 11, 15, 17, 19, 21, 23, 25\}$$

To find a inverse, we can simply compute it by trial and error:
For a given a we simply try all possible values a inverse until we obtain:
a.a inverse = 1 mod 26

Example .Let the key be k = (a,b) = (9,13), and the plaintext be:
ATTACK = x1,x2, . . . ,x6 = 0,19,19,0,2,10.

The inverse of a inverse of a exists and is given by a inverse = 3.
The ciphertext is then computed as
y1,y2, . . . ,y6 = 13,2,2,13,5,25 = nccnfz

4 Asymmetric Key Cryptography

The term public-key cryptography is used interchangeably with asymmetric cryptography; they both denote exactly the same thing and are used synonymously.

Public-key cryptography is an encryption scheme that uses two mathematically related, but not identical, keys - a public key and a private key. Unlike symmetric key algorithms that rely on one key to both encrypt and decrypt, each key performs a unique function. The public key is used to encrypt and the private key is used to decrypt.

It is computationally infeasible to compute the private key based on the public key. Because of this, public keys can be freely shared, allowing users an easy and convenient method for encrypting content and verifying digital signatures, and private keys can be kept secret, ensuring only the owners of the private keys can decrypt content and create digital signatures.

Since public keys need to be shared but are too big to be easily remembered, they are stored on digital certificates for secure transport and sharing. Since private keys are not shared, they are simply stored in the software or operating system you use, or on hardware (e.g., USB token, hardware security module) containing drivers that allow it to be used with your software or operating system.

Terminology Related to Asymmetric Encryption

Asymmetric Keys
Two related keys, a public key and a private key, that are used to perform complementary operations, such as encryption and decryption or signature generation and signature verification.
Public Key Certificate
A digital document issued and digitally signed by the private key of a Certification Authority that binds the name of a subscriber to a public key. The certificate indicates that the subscriber identified in the certificate has sole control and access to the corresponding private key.
Public Key (Asymmetric) Cryptographic Algorithm
A cryptographic algorithm that uses two related keys, a public key and a private key. The two keys have the property that deriving the private key from the public key is computationally infeasible.
Public Key Infrastructure (PKI)
A set of policies, processes, server platforms, software and workstations used for the purpose of administering certificates and public-private key pairs, including the ability to issue, maintain, and revoke public key certificates.

4.1 Public-Key Cryptosystems

Asymmetric algorithms rely on one key for encryption and a different but related key for decryption. These algorithms have the following important characteristic.

- It is computationally infeasible to determine the decryption key given only knowledge of the cryptographic algorithm and the encryption key.

In addition, some algorithms, such as RSA, also exhibit the following characteristic.

- Either of the two related keys can be used for encryption, with the other used for decryption.

A public-key encryption scheme has six ingredients

- **Plaintext:** This is the readable message or data that is fed into the algorithm as input.
- **Encryption algorithm:** The encryption algorithm performs various transformations on the plaintext.
- **Public and private keys:** This is a pair of keys that have been selected so that if one is used for encryption, the other is used for decryption. The exact transformations performed by the algorithm depend on the public or private key that is provided as input.
- **Ciphertext:** This is the encrypted message produced as output. It depends on the plaintext and the key. For a given message, two different keys will produce two different ciphertexts.
- **Decryption algorithm:** This algorithm accepts the ciphertext and the matching key and produces the original plaintext.

5 RSA algorithm

The RSA algorithm is the basis of a cryptosystem – a suite of cryptographic algorithms that are used for specific security services or purposes – which enables public key encryption and is widely used to secure sensitive data, particularly when it is being sent over an insecure network such as the internet.

RSA was first publicly described in 1977 by Ron Rivest, Adi Shamir and Leonard Adleman of the Massachusetts Institute of Technology, though the 1973 creation of a public key algorithm by British mathematician Clifford Cocks was kept classified by the U.K.'s GCHQ until 1997.

In RSA cryptography, both the public and the private keys can encrypt a message; the opposite key from the one used to encrypt a message is used to decrypt it.This attribute is one reason why RSA has become the most widely used asymmetric algorithm: It provides a method to assure the confidentiality, integrity, authenticity, and non-repudiation of electronic communications and data storage.

5.1 Working of RSA

The idea of RSA is based on the fact that it is difficult to factorize a large integer. The public key consists of two numbers where one number is multiplication of two large prime numbers. And private key is also derived from the same two prime numbers. So if somebody can factorize the large number, the private key is compromised. Therefore encryption strength totally lies on the key size and if we double or triple the key size, the strength of encryption increases exponentially. RSA keys can be typically 1024 or 2048 bits long.

5.1.1 Generating Public Key

Select two prime no's. Suppose P = 53 and Q = 59.
Now First part of the Public key : n = P*Q = 3127.
We also need a small exponent say e :
But e Must be

- An integer.
- Not be a factor of n.
- 1 less than e less than $\Phi(n)$

5.1.2 Generating Private Key

Calculation of $\Phi(n)$:
Such that $\Phi(n)$= (P-1)(Q-1)
so, $\Phi(n)$ = 3016

Private Key, d :
d = (k*$\Phi(n)$ + 1) / e for some integer k
For k = 2, value of d is 2011.

Now, Public Key (n = 3127 and e = 3) and Private Key(d = 2011)

Example. Encrypt **"HI"** using RSA:
Converting letters to numbers : H = 8 and I = 9

Thus Encrypted Data c = 89^e mod n.
Thus Encrypted Data = 1394

Now, for decryption of 1394 :

Decrypted Data = c^d mod n.
Thus our Encrypted Data comes out to be 89
8 = H and I = 9 i.e. "HI".

5.1.3 Program to implement RSA

```
// Returns gcd of a and b
int gcd(int a, int h)
{
    int temp;
    while (1)
    {
        temp = a%h;
        if (temp == 0)
          return h;
        a = h;
        h = temp;
    }
}
// Code to demonstrate RSA algorithm
int main()
{
    // Two random prime numbers
    double p = 3;
    double q = 7;

    // First part of public key:
    double n = p*q;

    // Finding other part of public key.
    // e stands for encrypt
    double e = 2;
    double phi = (p-1)*(q-1);
    while (e < phi)
    {
        // e must be co-prime to phi and
        // smaller than phi.
        if (gcd(e, phi)==1)
            break;
```

```
        else
            e++;
    }

    // Private key (d stands for decrypt)
    // choosing d such that it satisfies
    // d*e = 1 + k * totient
    int k = 2;  // A constant value
    double d = (1 + (k*phi))/e;

    // Message to be encrypted
    double msg = 20;

    // Encryption c = (msg ^ e) % n
    double c = pow(msg, e);
    c = fmod(c, n);

    // Decryption m = (c ^ d) % n
    double m = pow(c, d);
    m = fmod(m, n);

    return 0;
}
```

6 Homomorphic Encryption

Homomorphic Encryption (HE) enables a user to perform meaningful computations on sensitive data while ensuring the privacy of the data. HE ensures that performing operations on encrypted data and decrypting the result is equivalent to performing analogous operations without any encryption.

Like SMPC, we can use HE to achieve input privacy but with only one party needed to encrypt and decrypt the data.

Types of Homomorphic Encryption

- Partially Homomorphic Encryption
- Somewhat homomorphic encryption
- Fully Homomorphic Encryption

6.1 Partially Homomorphic Encryption

Partially homomorphic encryption (PHE) helps sensitive data remain confidential by only allowing select mathematical functions to be performed on encrypted values. This means that one operation can be performed an unlimited number of times on the ciphertext. Partially homomorphic encryption (with regard to multiplicative operations) is the foundation for RSA encryption, which is commonly used in establishing secure connections through SSL/TLS. Some

examples of PHE include ElGamal encryption (a multiplication scheme) and Paillier encryption (an addition scheme).

6.2 Somewhat homomorphic encryption

A somewhat homomorphic encryption (SHE) scheme is one that supports limited operations (for example, either addition or multiplication) up to a certain complexity, but these operations can only be performed a set number of times. This is the precursor to fully homomorphic encryption, which we'll discuss more in depth momentarily.

6.3 Fully Homomorphic Encryption

Fully homomorphic encryption (FHE), while still in the development stage, has a lot of potential for making functionality consistent with privacy by helping to keep information secure and yet still accessible. Derived from homomorphic encryption scheme, this is capable of using any efficiently computable functions (such as addition and multiplication, not just one or the other) any number of times and makes secure multi-party computation more efficient. Unlike other forms of homomorphic encryption, it can handle arbitrary computations on your ciphertexts.

The goal behind fully homomorphic encryption is to allow anyone to use encrypted data to perform useful operations without access to the encryption key. In particular, this concept has applications for improving cloud computing security.

6.4 Implementation of Homomorphic Encryption

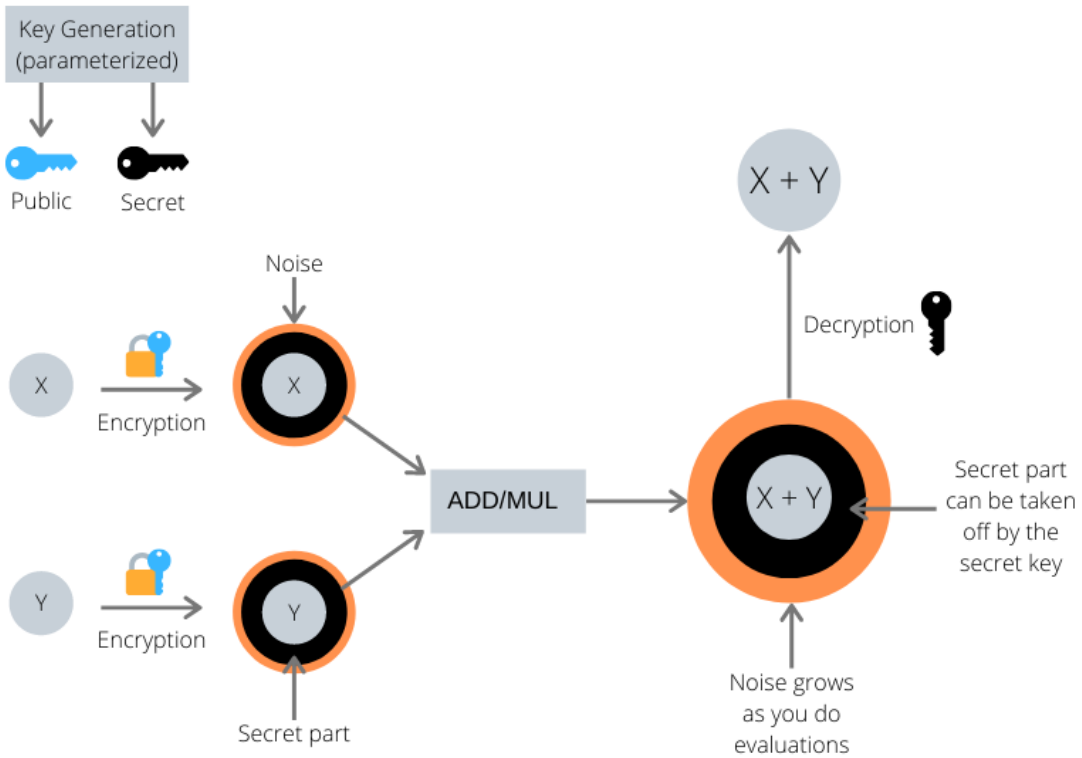

Figure 1: HE Implementation

```
import numpy as np
from numpy.polynomial import polynomial as poly

def polymul(x, y, modulus, poly_mod):
    """Add two polynoms
    Args:
        x, y: two polynoms to be added.
        modulus: coefficient modulus.
        poly_mod: polynomial modulus.
    Returns:
        A polynomial in Z_modulus[X]/(poly_mod).
    """
    return np.int64(
        np.round(poly.polydiv(poly.polymul(x, y) % modulus, poly_mod)[1] % modulus)
    )


def polyadd(x, y, modulus, poly_mod):
    """Multiply two polynoms
    Args:
        x, y: two polynoms to be multiplied.
        modulus: coefficient modulus.
        poly_mod: polynomial modulus.
    Returns:
        A polynomial in Z_modulus[X]/(poly_mod).
    """
    return np.int64(
        np.round(poly.polydiv(poly.polyadd(x, y) % modulus, poly_mod)[1] % modulus)
    )
```

Key Generation

```
def gen_binary_poly(size):
    """Generates a polynomial with coeffecients in [0, 1]
    Args:
        size: number of coeffcients, size-1 being the degree of the
            polynomial.
    Returns:
        array of coefficients with the coeff[i] being
        the coeff of x ^ i.
    """
    return np.random.randint(0, 2, size, dtype=np.int64)


def gen_uniform_poly(size, modulus):
    """Generates a polynomial with coeffecients being integers in Z_modulus
    Args:
        size: number of coeffcients, size-1 being the degree of the
```

```
            polynomial.
    Returns:
        array of coefficients with the coeff[i] being
        the coeff of x ^ i.
    """
    return np.random.randint(0, modulus, size, dtype=np.int64)


def gen_normal_poly(size):
    """Generates a polynomial with coeffecients in a normal distribution
    of mean 0 and a standard deviation of 2, then discretize it.
    Args:
        size: number of coeffcients, size-1 being the degree of the
            polynomial.
    Returns:
        array of coefficients with the coeff[i] being
        the coeff of x ^ i.
    """
    return np.int64(np.random.normal(0, 2, size=size))
```

Key generator function

```
def keygen(size, modulus, poly_mod):
    """Generate a public and secret keys
    Args:
        size: size of the polynoms for the public and secret keys.
        modulus: coefficient modulus.
        poly_mod: polynomial modulus.
    Returns:
        Public and secret key.
    """
    sk = gen_binary_poly(size)
    a = gen_uniform_poly(size, modulus)
    e = gen_normal_poly(size)
    b = polyadd(polymul(-a, sk, modulus, poly_mod), -e, modulus, poly_mod)
    return (b, a), sk
```

The public-key (b, a) can then be used for encryption, and the secret-key sk for decryption.

Encryption

```
def encrypt(pk, size, q, t, poly_mod, pt):
    """Encrypt an integer.
    Args:
        pk: public-key.
        size: size of polynomials.
        q: ciphertext modulus.
        t: plaintext modulus.
        poly_mod: polynomial modulus.
        pt: integer to be encrypted.
```

```
    Returns:
        Tuple representing a ciphertext.
    """
    # encode the integer into a plaintext polynomial
    m = np.array([pt] + [0] * (size - 1), dtype=np.int64) % t
    delta = q // t
    scaled_m = delta * m  % q
    e1 = gen_normal_poly(size)
    e2 = gen_normal_poly(size)
    u = gen_binary_poly(size)
    ct0 = polyadd(
            polyadd(
                polymul(pk[0], u, q, poly_mod),
                e1, q, poly_mod),
            scaled_m, q, poly_mod
        )
    ct1 = polyadd(
            polymul(pk[1], u, q, poly_mod),
            e2, q, poly_mod
        )
    return (ct0, ct1)
```

Decryption

```
def decrypt(sk, size, q, t, poly_mod, ct):
    """Decrypt a ciphertext
    Args:
        sk: secret-key.
        size: size of polynomials.
        q: ciphertext modulus.
        t: plaintext modulus.
        poly_mod: polynomial modulus.
        ct: ciphertext.
    Returns:
        Integer representing the plaintext.
    """
    scaled_pt = polyadd(
            polymul(ct[1], sk, q, poly_mod),
            ct[0], q, poly_mod
        )
    decrypted_poly = np.round(scaled_pt * t / q) % t
    return int(decrypted_poly[0])
```

Addition

```
def add_plain(ct, pt, q, t, poly_mod):
    """Add a ciphertext and a plaintext.
    Args:
        ct: ciphertext.
        pt: integer to add.
        q: ciphertext modulus.
```

```
        t: plaintext modulus.
        poly_mod: polynomial modulus.
    Returns:
        Tuple representing a ciphertext.
    """
    size = len(poly_mod) - 1
    # encode the integer into a plaintext polynomial
    m = np.array([pt] + [0] * (size - 1), dtype=np.int64) % t
    delta = q // t
    scaled_m = delta * m  % q
    new_ct0 = polyadd(ct[0], scaled_m, q, poly_mod)
    return (new_ct0, ct[1])
```

Multiplication

```
def mul_plain(ct, pt, q, t, poly_mod):
    """Multiply a ciphertext and a plaintext.
    Args:
        ct: ciphertext.
        pt: integer to multiply.
        q: ciphertext modulus.
        t: plaintext modulus.
        poly_mod: polynomial modulus.
    Returns:
        Tuple representing a ciphertext.
    """
    size = len(poly_mod) - 1
    # encode the integer into a plaintext polynomial
    m = np.array([pt] + [0] * (size - 1), dtype=np.int64) % t
    new_c0 = polymul(ct[0], m, q, poly_mod)
    new_c1 = polymul(ct[1], m, q, poly_mod)
    return (new_c0, new_c1)
```

Final function

```
# Scheme's parameters
# polynomial modulus degree
n = 2**4
# ciphertext modulus
q = 2**15
# plaintext modulus
t = 2**8
# polynomial modulus
poly_mod = np.array([1] + [0] * (n - 1) + [1])
# Keygen
pk, sk = keygen(n, q, poly_mod)
# Encryption
pt1, pt2 = 73, 20
cst1, cst2 = 7, 5
ct1 = encrypt(pk, n, q, t, poly_mod, pt1)
ct2 = encrypt(pk, n, q, t, poly_mod, pt2)
```

```
print("[+] Ciphertext ct1({}):".format(pt1))
print("")
print("\t ct1_0:", ct1[0])
print("\t ct1_1:", ct1[1])
print("")
print("[+] Ciphertext ct2({}):".format(pt2))
print("")
print("\t ct1_0:", ct2[0])
print("\t ct1_1:", ct2[1])
print("")

# Evaluation
ct3 = add_plain(ct1, cst1, q, t, poly_mod)
ct4 = mul_plain(ct2, cst2, q, t, poly_mod)

# Decryption
decrypted_ct3 = decrypt(sk, n, q, t, poly_mod, ct3)
decrypted_ct4 = decrypt(sk, n, q, t, poly_mod, ct4)

print("[+] Decrypted ct3(ct1 + {}): {}".format(cst1, decrypted_ct3))
print("[+] Decrypted ct4(ct2 * {}): {}".format(cst2, decrypted_ct4))
```

7 Case Studies

7.1 Credit Rating

A credit rating is a quantified assessment of the creditworthiness of a borrower in general terms or with respect to a particular debt or financial obligation.

Credit ratings determine not only whether or not a borrower will be approved for a loan or debt issue but also the interest rate at which the loan will need to be repaid.

A credit rating or score can be assigned to any entity that seeks to borrow money—an individual, a corporation, a state or provincial authority, or a sovereign government.

Individual credit is rated on a numeric scale based on the FICO calculation; bonds issued by businesses and governments are rated by credit agencies on a letter-based system.

7.2 Social credit system

The premise behind a nationwide social rating system itself is relatively simple: every citizen receives a certain score to start, and certain actions either lower or increase your score. For example, donating to charity would increase your score, while buying cigarettes would lower it.

People can then either be rewarded or punished based on their rating.

The government, for example, could restrict a person's travel, prevent them from entering the best universities, or even take their dog away if their score drops low enough.

Two of the biggest unanswered questions are what the system will actually look like in practice and how the system operates from a technological standpoint. There are several possible

reasons for this.

First, the developers of the system understand that the more they reveal about the system, the more they open it up to vulnerabilities (more on that later).

Second, it's entirely possible that the creators themselves do not fully understand how this system works. Many rating systems used today are built on large volumes of historical data and machine learning models that predict the future behavior and successes of system participants. For example, Microsoft uses such models to rank the skill level of players in online games, banks evaluate the reliability of potential borrowers when they submit applications for loans, and several companies have even tried to automate the process of reviewing resumes for open vacancies. In these situations, developers put their trust in the algorithms.

AI is based on machine learning algorithms, which, despite being generally lauded by the technological industry as a "cure-all", are far from perfect. Like any other computer system, they can be prone to errors. For example, in cybersecurity, machine-learning algorithms are used to rapidly detect previously unknown malware. However, there is a problem: the higher the detection rate, the higher the chance you'll run into "false positives"—i.e., the system determines a non-malicious file is malicious. This happens due to the very nature of how machine learning works: the ML-based system doesn't look into the details of an object but compares the way it "looks" to other known objects. In some cases, "clean" objects may "look" a lot like malicious ones and a system that makes decisions based on scoring would most likely confirm the object as a malicious one. When applied to a world where people's behavior is evaluated by a fully automated system, this particularity of machine learning systems may lead to multiple unpleasant situations where an innocent person is confirmed as the one behind "wrong" actions.

These systems are also susceptible to issues such as developer bias, false correlations, and feedback loops, and, unless specifically included by the developer, the algorithms do not factor in ethical considerations. To simply input massive quantities of information into a machine learning system and then accept the result without any critical assessment could lead to a host of unintended consequences, including choices that ultimately infringe upon the rights of certain citizens.

Lastly, should the control over a particular system be concentrated in the hands of only one social group, their ability to change the rules of how the system works may significantly influence the life of those social groups who are not in a position to influence the scoring rules.

7.3 Social credit system of China

7.3.1 Technologies Used

China's surveillance system:
The Chinese big data department has calculated that 80% of the data available in China is centred in the government, due to this they must enhance the operation of government data to develop the big data system the government desires. So far, the big data department together with big data administration and big data development bureau have established departments in various places across the country, to manage and collaborate big data resources nationwide.

The Chinese big data department holds "the goal of the big data system is to form data assets through the collection and exchange of a wide range of data sources, through data governance and development, and in accordance with the internal and external sharing needs of the government, thereby proving a comprehensive, efficient, and reliable data supply chain"

Rui, a giant tech-company expert, wrote a column on china.org.cn at the beginning of 2020, claiming that AI and blockchain are the future leading technologies of scientific and technological innovation. Furthermore, he presents that China will become the leader of this game-changing and widely influential technology, due to the "Next Generation Artificial Intelligence Development Plan" (henceforth the Plan), which the Council developed in 2017. The column holds that this was a huge event in history of science and technology since it was the first AI plan ever developed. The Plan holds "that it will give full play to the role of AI technology in enhancing social interaction, promoting credible communication, facilitating the integration of blockchain technology and AI, establishing a new social credit system, and minimizing the cost and risk of interpersonal communication"

Blockchain
According to a blockchain expert, the government has mainly been concentrating its resources on private blockchain projects, although it has cautiously and discreetly been engaging in public blockchain projects too. The expert believes that the government aims to track everything with this new technological network.

Fifth Generation cellular network
Xinhua holds that 5G in combination with big data and AI, will take revolutionary steps towards a digital economy and this will make a huge positive impact on China's future economy.

Monitoring behaviour across platform
A United States Human Rights report (Henceforth U.S HR-report) claim that behavioural data can include: "collected information on academic records, traffic violations, social media presence, friendships, adherence to birth control regulations, employment performance, consumption habits, and other topics"

The report criticizes China's expansion of AI, by arguing that the system is already causing damage by monitoring speech and movement, not only through cameras but also through other electronics such as phone apps.

Identifying individuals
The Ministry of Foreign Affairs of the PRC holds that the "Next Generation Artificial Intelligence Development Plan Issued by the Council" has developed a three years Internet + AI plan, enhancing the industrial development, and measures for applications and technology Robotic Development. During the last couple of years, this plan has made China become the world-leading within the technology of voice and visual recognition, but also developed advanced biometric recognition and intelligence monitors, and also evolved practical apps of Chinese information processing.

The video made by VICE News on the SCS holds that AI appears everywhere, for instance at various fast food restaurants in China, citizens can order their food at a machine and then pay with their face. This system is called "smile-to-pay" and is a highly technological camera which will scan the person's face and identify him/her. The camera will recognise the paying

customer even though (s)he is wearing a wig or a lot of make-up. The video also states that many giant tech-companies in China are developing video analysis systems.

7.3.2 Possible Reasons for Rewarding and Deducting score

There are multiple social credit systems in China right now. Scholars have conceptualized four different types of systems:

- The Judicial system,
- The Municipal social credit system,
- People's Bank of China financial credit system, and
- Commercial credit-rating system.

These four systems are not interconnected seamlessly, but relevantly independent from each other with their own jurisdictions, rules, and logic.

The social credit system compiles a score for both individuals and companies after collecting, aggregating and analysing data from different sources.

For businesses, in addition to its own operations, companies are asked to submit information on their partners and suppliers to local and national authorities.

- Bad behaviour,
- low trustworthiness and
- ratings from suppliers and customers

will also influence a company's own credit score.

A good rating will lead to rewards, while a poor rating could see an individual or a company punished or sanctioned.

Every citizen starts off with a score of 1,000. NPR reported the ranking as follows: 960 to 1,000 is an A; 850 to 955 points is a B; 840 to 600 is a C; and any score below that is a D, which designates the score-holder as "untrustworthy."

7.3.3 Advantages of China's Social Credit System

Solving the "trust deficit"
According to officials, the SCS will address China's "trust deficit," an epidemic of low-quality goods and fraud and financial scams that are harming the population. For example, the contaminated milk powder scandal in 2008 sickened almost 300, 000 children. By increasing trust and monitoring businesses, the government can better hold people accountable.

Subprime credit
Many Chinese citizens don't have a credit history, making it hard to obtain loans or mortgages.

By using non-economic measures to evaluate credit, these citizens could gain access to resources for economic and social mobility.

Punished for actions, not credit
In an article, "No, China isn't Black Mirror," Ed Jefferson points out that the current government SCS does not punish citizens simply for having low credit (those pilots like Sesame may). Only people who commit fraud or civil offenses are punished. While some petty actions, like walking your dog off-leash, may lower your score, they won't get you blacklisted or banned from flights.

7.3.4 Risks in China's Social Credit System

Rates businesses and individuals
According to a 2014 State Council report, the goal of the SCS is to promote "trustworthiness" through four goals: honesty in government affairs, commercial integrity, societal integrity, and judicial credibility. Three of these four goals (excluding societal integrity) could easily be accomplished by focusing on businesses and state institutions. It's one thing to hold producers and government accountable, but a whole other ball game to hold individuals to such rigid standards.

Algorithms are biased too
Since algorithms are made by humans, they inherently contain human bias. Thus, the decision of what is "good" or "bad" behaviour has serious implications. For example, Sesame's understanding of video gaming as a negative trait is clearly debatable, but holds serious consequences for gamers. Just as racial profiling often surfaces in crime prediction algorithms, I'd be worried about discrimination in the SCS.

Algorithms are "black boxes"
Academics have described algorithms as a "black box," highlighting how little we know about their inner workings. While some blacklists and credit scores may be published, we have no way of knowing how these are combined with consumer and personality attributes to result in an individual rating. Without transparency, there are concerns that the government could misuse this program for its own purposes.

Ignores social and personal context
Algorithms are extremely reductive, yet have the power to change people's lives. For example, there is clearly a difference between one person who misses a payment while in hospital and another who is simply a freeloader. Further, social norms privilege the elite, rather than those in need.

Labelling and stigma
As I learned in an introductory criminology course, labelling acts of "deviance" often results in stigma, decreased life chances, and a higher rate of repeated offending. Stigma requires an audience, and what better way to shame someone than defaming them through a nation-wide reputation system? This may harm an individual's chances of rehabilitation after an offense or mistake.

Human rights violations
Some aspects of SCS and its pilot initiatives may violate basic human rights. For example,

Sesame users with a high score can skip the long lineups at healthcare clinics. This policy gives priority access based on a judgement that is not even medical. Further, SCS adjusts your score based on who your friends and associations are. A friend whose "bad" actions lower their credit score may drag yours down as well, punishing innocent people.

7.3.5 Potential Rewards of a Good Score

- Discounts on energy bills
- Rent things without deposits
- Better interest rates at banks
- Will reportedly get more matches on dating websites.
- Fewer inspections and audits
- Fast-tracked approvals.

Companies with a high score are placed on a "redlist". There is a range of rewards to businesses that do well in this regard, including:

- Streamlined administrative procedures.

For example, companies that are classified as an 'Advanced Certificate Enterprise' may receive faster customs clearance. A-rated tax-payers may have their tax returns processed more quickly.

7.3.6 Potential Negative Effects of a Bad Score

- Banning you from flying or getting on the train.

China has already started punishing people by restricting their travel. Nine million people with low scores have been blocked from buying tickets for domestic flights.

- Throttling internet speeds.

Spreading fake news, specifically about terrorist attacks or airport security, will also be punishable offences.

- Banning from the best schools.

Citizens with low social credit would also be prohibited from enrolling their children at high-paying private schools.

- Stopping from getting the best jobs.

"Trust-breaking" individuals would also be banned from doing management jobs in state-owned firms and big banks.

- Keeping out of the best hotels.

People who refused military service were also banned from some holidays and hotels.

- Being publicly named as a bad citizen.

Naming and shaming is another tactic available. A 2016 government notice encourages companies to consult the blacklist before hiring people or giving them contracts.

7.4 Social Credit System in other countries

7.4.1 Social Credit System in Australia

It is not news that the Australian government is relentless in its pursuit of those on social security payments. The robodebt program is evidence of that. But there are other concerning pointers to an active government agenda of social control.

One of these is the roll out of the Indue card. Trialled in Indigenous communities as a means of controlling spending — limiting it to the 'basics' — the card is being rolled out to other communities also. Ostensibly to stop spending on alcohol, cigarettes, and gambling, it also limits recipients' ability to pay for goods by cash and limits where recipients can spend.

A second is the control over activities by mothers receiving parenting payments through the 'Parents Next' program. The overwhelming majority of recipients are women, and there is a significant proportion of those who are Aboriginal or Torres Strait Islander Australians.To continue to qualify for the payment, recipients must verify weekly that they have undertaken nominated activities with their children. Failure to do so without receiving an exemption results in losing their payment.

A third is the drug testing of social security recipients. Wrapped up as concern about addiction, the proposed program will place anyone testing positive on income management for two years. If they get a second positive result, they will be referred to a doctor for treatment options and will be required to undergo drug treatment as part of their job plan.

A fourth, and latest, pointer is the suggestion by Home Affairs Minister, Peter Dutton, that climate change protesters should have their welfare payments cut. He also suggested mandatory jail sentences for protesters and encouraged members of the public to take photos of protesters and to 'name and shame' them. Employment Minister Michaelia Cash has agreed reportedly saying: 'Protesting is not, and never will be, an exemption from a welfare recipient's mutual obligation to look for a job.'

All of these programs smack of benevolent paternalism that judge morality of social security recipients, and the message is clear. Drug use is personal failure; cigarettes, alcohol and gambling are weakness; single mothers are bad mothers; protesters are lazy and unproductive members of society. The solution? Government is everybody's long-suffering father, here to instruct a wayward public in proper civic behaviour by offering correctives to immoral behaviour.

The programs are also couched in terms of mutual obligation, namely that the recipient has a duty to government to receive its munificence. It is, in fact, the responsibility of government to provide for the people.

And even if one accepts some degree of accountability for payments, these programs beg the question of just how far government can reach into our lives to control how we live. If a person meets reporting requirements for Newstart, what right does government have to inquire further? So what if recipients spend their time volunteering, lying in bed, or engaging in peaceful protest? What possible value is there to anyone in interfering with the parenting decisions of parents receiving social security — on pain of losing payment?
If government is concerned for citizens' wellbeing, then it should properly resource services — drug and alcohol support, parenting support, subsidised childcare, financial counselling, edu-

cation, and so on. Instead, it is generating a system of social credit: rewarding those who toe the line and punishing those whose 'score' falls below that of the 'good citizen'.

7.4.2 Social Credit System in Canada

Vancouver, British Columbia, Canada: China's Orwellian "social credit system" that records the social and financial behaviour of individuals and corporations across China, using a vast surveillance system, has expanded globally, and is now openly operational at the renowned Haidilao hot pot restaurant, in Western Canada.

Ryan Pan, a manager with Haidilao Hot Pot in Vancouver confirmed that over 60 surveillance cameras have been installed in the restaurant at the request of the Haidilao corporation, as part of the social credit system in China. He said that the Vancouver location has 30 tables with two cameras assigned to each table.

China's CSCS operating within Canadian borders all boils down to the safety of workers, human rights, privacy of citizenry and national security, all of which is governed by legislation at a municipal, provincial, and federal level. The Canadian government is aware that China has implemented the CSCS and has even issued recommendations on how to conduct business inside China now that it has been implemented. But little to nothing has been done by elected officials to prevent China from implementing the CSCS inside Canada as a way to control foreign workers from China and Chinese owned businesses, or anyone for that matter, who is ethnic Chinese with personal or professional ties to China.

At a municipal (city) level, until recently, commercial security cameras used to be powered from a plug in the wall with a video cable going back to a personal video recorder (PVR). The City of Vancouver would normally require an electrical permit for this kind of set-up which would have, at the very least, alerted city officials in this case, given the large volume of cameras installed. However, the Haidilao cameras appear to be more modern cameras powered over Ethernet (POE). Meaning they are plugged into a network that can send video footage live back to China. They can be installed at any point without a permit. Quite simply, there is nothing stopping the Chinese government from insisting that all businesses in Canada who have ties to China either professionally or personally install a surveillance system as part of China's social credit system.

British Columbia purports to have strict privacy laws with the Personal Information Protection Act (PIPA) governing how an organization can collect, use, or disclose information on individuals. When asked if the Office of the Privacy Commissioner was aware of CSCS sending private footage of temporary foreign workers and Canadian citizens back to China, Michelle Mitchell with the Office of the Privacy Commissioner in BC (unrelated to author), stressed that an important component of PIPA is consent and that an organization must have consent before collecting, using, or disclosing personal information, citing three types of consent under PIPA, which is express consent, implied consent, and opt-out consent. Since many of the staff at the Haidilao restaurant are Chinese citizens and work under the Temporary Foreign Worker Program in Canada, and are already listed in the Social Credit system in China as citizens of China, they aren't likely to demand privacy under Canadian law. Even staff members who live in Canada permanently with "permanent resident status" face the same pressure to conform to China's mandatory CSCS program, because China doesn't consider individuals with permanent resident status in Canada unless they renounce their Chinese citizenship. "Consent"

isn't an option in totalitarian regimes. A more insidious aspect to China's CSCS is that it can be used to spy on Canadian citizens under the pretence of being a part of China's social credit program. Canadian authorities may take the approach that CSCS has nothing to do with Canada and our laws. However, the Haidilao hot pot restaurant manager clearly stated that there were two reasons for the surveillance cameras, for both social credit and state security purposes. In all likelihood, next-generation data sources—such as information from facial recognition-driven video feeds, cell phone surveillance and e-com purchase history—are being collected from Canadians while they eat and it wouldn't be the first time that the CCP has filmed Canadians inside Canada.

Over the past couple of years, an ever-expanding collection of surveillance cameras have been added to the Vancouver PRC Consulate, located in the prestigious Shaughnessy district of Vancouver, on the edge of a high traffic road that goes into the City. The surveillance camera has been installed on the edge of the property, one of which jettisons out into Canadian space, then retracts back. The street is a frequent place for pro-Hong Kong activists, Uyghurs protesting genocide, Iranians protesting the cooperation pact with China and members of Falun Gong protesting persecution. Canada is a multicultural country that has opened its doors to countless refugees fleeing persecution and promotes equality and the right to free speech as a core value. Yet Canada's elected officials have done little if anything to ensure that the very people we have invited into Canada are promised a safe haven from the oppressive regimes they fled from in the first place.

The cameras used by the PRC consulate are the same cameras that the US banned the purchase and use of under the National Defense Authorization Act (NDAA) as a cyber security threat.
In 2019, Global Affairs, the department of the Government of Canada that manages Canada's diplomatic and consular relations, also doesn't appear to concerned that the Government of China is spying on Canadians with cameras installed on embassy grounds that jettison out into Canadian space, naively suggesting that addressing the matter is left to the Peoples Republic of China and the "honor" system of having a duty to adhere to Canadian laws.

"Diplomatic and consular representatives have a duty to respect local laws, and there is a similar expectation for the manner in which foreign missions operate. In Canada, there is an expectation that foreign missions comply with federal, provincial, and municipal laws and regulations, including those that governs the use of physical security equipment, such as cameras, fences, and lights."

7.4.3 Social Credit System in the United States

Some media outlets have compared the social credit system to credit scoring systems in the United States. According to Fast Company, "increasing number of societal "privileges" related to transportation, accommodations, communications, and the rates we pay for services (like insurance) are either controlled by technology companies or affected by how we use technology services. And Silicon Valley's rules for being allowed to use their services are getting stricter."

7.4.4 Social Credit System in United Kingdom

In 2018, the New Economics Foundation compared the Chinese citizen score to other rating systems in the United Kingdom. These included using data from a citizen's credit score, phone usage, rent payment, and so on, to filter job applications, determine access to social services, determine advertisements served, etc.

7.5 Privacy enhancement technologies

Privacy-enhancing technologies (PETs) are a broad range of technologies (hardware or software solutions) that are designed to extract data value in order to unleash its full commercial, scientific and social potential, without risking the privacy and security of this information.

7.5.1 Importance of PETs

Like any other data privacy solution, privacy-enhancing technologies are important due to three reasons for businesses:

- Data protection laws such as GDPR and CCPA are forcing organizations to preserve consumer data. Businesses can pay serious fines because of data breaches.
- Data may need to be tested by third-party organizations due to the lack of your business' self-sufficiency in analytics and application testing. PETs enable privacy protection while data sharing.
- Privacy breaches can harm your business' reputation, businesses or customers (depending on your business model) may want to stop interacting with your brand.An example is the share price loss of Facebook after Cambridge Analytica scandal.

7.6 Common privacy-enhancing technology

7.6.1 Cryptographic algorithms

Homomorphic Encryption: Homomorphic encryption is an encryption method that enables computational operations on encrypted data. It generates an encrypted result which, when decrypted, matches the result of the operations as if they had been performed on unencrypted data (i.e. plaintext). This enables encrypted data to be transfered, analyzed and returned to the data owner who can decrypt the information and view the results on the original data. Therefore, companies can share sensitive data with third parties for analysis purposes. It is also useful in applications that hold encrypted data in cloud storage. Some common types of homomorphic encryption are:

- Partial homomorphic encryption: can perform one type of operation on encrypted data, such as only additions or only multiplications but not both.
- Somewhat homomorphic encryption: can perform more than one type of operation (e.g. addition, multiplication) but enables a limited number of operations.
- Fully homomorphic encryption: can perform more than one type of operation and there is no restriction on the number of operations performed.

Secure multi-party computation (SMPC): This is a subfield of homomorphic encryption with one difference: users are able to compute values from multiple encrypted data sources. Therefore, machine learning models can be applied to encrypted data since SMPC is used for a larger volume of data. Differential privacy: Differential privacy protects from sharing any information about individuals. This cryptographic algorithm adds a "statistical noise" layer to the dataset which enables to describe patterns of groups within the dataset while maintaining the privacy of individuals.

Zero-knowledge proofs (ZKP): ZKP uses a set of cryptographic algorithms that allow information to be validated without revealing data that proves it.

7.6.2 Data masking techniques

Some privacy enhancing technologies are also data masking techniques that are used by businesses to protect sensitive information in their data sets.

Obfuscation: This one is a general term for data masking that contains multiple methods to replace sensitive information by adding distracting or misleading data to a log or profile.
Pseudonymization: Identifier fields (fields that contain information specific to an individual) are replaced with fictitious data such as characters or other data. Pseudonymization is frequently used by businesses to comply with GDPR.
Data minimisation: Collecting a minimum amount of personal data that enables the business to provide the elements of a service.
Communication anonymizers: Anonymizers replace online identity (IP address, email address) with disposal/one-time untraceable identity.

7.6.3 With the help of AI and ML algorithms

Synthetic data generation: Synthetic data is an artificially created data by using different algorithms including ML algorithms. If you are interested in privacy-enhancing technologies because you need to transform your data into a testing environment where third-party users have access, generating synthetic data that has the same statistical characteristics is a better option.

Federated learning: This is a machine learning technique that trains an algorithm across multiple decentralized edge devices or servers holding local data samples, without exchanging them. With the decentralization of servers, users can also achieve data minimization by reducing the amount of data that must be retained on a centralized server or in cloud storage.

7.7 Uses of PETs

Test data management: Application testing and data analysis are sometimes handled by third-party providers. Even when they are handled in-house, companies should minimize internal access to customer data. Using a suitable PET that doesn't significantly affect test results is important for organizations.

Financial transactions: Financial institutions are responsible for protecting the privacy of the customers due to citizens' freedom to conduct private deals and transactions with other

parties.

7.7.1 In medical domain

Healthcare services: Healthcare industry collects and shares (when needed) electronic health records (EHR) of patients. For example, clinical data can be used for searching for adverse effects of various drug combinations. Healthcare companies ensure the privacy of patients' data in such cases by using PETs.

Data Anonymization and Pseudonymization
The GDPR defines personal data as concerning an identifiable natural person. Therefore, pseudonymized data, where all identifiers have been removed from those data, remain personal data. However, the provisions of the GDPR do not concern anonymized data or data which have been processed so individuals are no longer identifiable. In particular, anonymized data may be used for research or statistical processing without the need to comply with the GDPR.

Data Processing
The GDPR's provisions apply to data controllers, or entities determining the purpose and means of processing personal data. This definition encompasses both health care institutions and research institutions. Data controllers must guarantee personal data processing is lawful, proportionate, and protects the rights of data subjects. In particular, the GDPR provides that encryption should be used as a safeguard when personal data are processed for a purpose other than which they were collected. Although the GDPR does not define encryption, the Article 29 Working Party treats encryption as equivalent to stripping identifiers from personal data. The GDPR also lists encryption as a strategy that can guarantee personal data security. Furthermore, the GDPR emphasizes that data controllers should consider the state of the art, along with the risks associated with processing, when adopting security measures. The GDPR also provides that data processing for scientific purposes should follow the principle of data minimization. This principle requires data processors and controllers to use nonpersonal data unless the research can only be completed with personal data. If personal data are required to complete the research, pseudonymized or aggregate data should be used instead of directly identifying data.

7.7.2 In governance

Governance in the 21st Century report, requires appropriate governance mechanisms, from codes of conduct and ethics to regulation. However, in some cases, technological solutions can help diffuse dilemmas between making use of data and protecting both the individuals and organizations that generate or are subjects within datasets. PETs as a category comprises a broad suite of technologies and approaches—from a piece of tape masking a webcam to advanced cryptographic techniques. While some are focused on protecting private communications, the report explored a subset of five PETs identified during the scoping of the project as being particularly promising to enable privacy-aware data collection, analysis, and dissemination of results.

The key question of this paper is whether, according to the current state of development and the trajectory of technological development, we can utilize PETs in addressing social and ethical tensions in data use, and thereby use them as tools for governing the ways that data

is used. This will involve consideration both of how these technologies can potentially enable governments and others to unlock the value of data, while also recognizing both contingent and in principle limitations on the role of PETs in ensuring well-governed use of data.

How far are these technologies able to underpin the ways that data use is governed in practice? The field of PETs development is moving quickly, and the Royal Society report captures a moment in time where the technologies are maturing and opportunities to use these technologies are beginning to emerge. It may be that some of the technologies surveyed in our report do not achieve their promise in the near term, or that the costs of adoption prove prohibitive, or that other technologies not explored in depth might leapfrog them. However, there are a number of areas where PETs are already in use which were set out in case studies in the report, with examples summarized here.

There are many examples of current uses of these technologies. For example, TEEs are used in mobile phones to process "touch ID" data. They are also an integral part of secure clouds. The following are some specific examples of where organizations have made use of, or promoted, PETs.

For secure MPC, the first real-world application of Sharemind—which uses MPC—was the analysis of key performance indicators for the Estonian Association of Information Technology and Telecommunications (ITL). The ITL proposed collecting certain financial metrics and analyzing them to gain insights into the state of the sector. The member companies expressed concerns over the confidentiality of the metrics, as they would be handing them out to competitors.

In order to share NHS data securely with multiple teams, while maintaining as much as possible the potential usefulness of the data, NHS Digital have been using a de-identification service employing homomorphic encryption. For security reasons, data is de-identified in different "pseudonymization domains" for each different part of an organization. Within one domain, all data with the same base value is replaced with the same "token" (a nonidentifying value). Across domains, the same base value receives different token. Usually, transferring data between domains requires to remove the encryption for the first domain and replace it with the second domain encryption. However, using consistent "tokenization" and partially homomorphic encryption by Privitar Publisher, it is possible to transform data items between any two domains without revealing the base value, even if they have been de-identified by two instances of the de-identification service using different encryption keys.

This methodology allows the de-identification tool set to be deployed to multiple locations across the NHS and makes any data de-identified by any tool from the de-identification tool set potentially linkable with any other data de-identified by any other tool from the tool set.

In an effort to empower consumers, the UK government promoted midata, a Personal Data Store. Launched in 2011, in partnership with multiple organization, the online portal was designed to provide citizens with access and control over data about them. For example, individuals can access the transactional data for their current account, which they can upload to third party price comparison websites to compare and identify the best value.

Facilitating data transfer between multiple parties including intermediaries: For businesses that work as a middle man between two parties, the usage of PETs is crucial since these businesses are responsible for protecting the privacy of both parties' information.

8 Designing Aadhaar based social credit system for India

8.1 Aadhaar

Aadhaar number is a 12-digit random number issued by the UIDAI ("Authority") to the residents of India after satisfying the verification process laid down by the Authority. Any individual, irrespective of age and gender, who is a resident of India, may voluntarily enrol to obtain Aadhaar number. Person willing to enrol has to provide minimal demographic and biometric information during the enrolment process which is totally free of cost. An individual needs to enrol for Aadhaar only once and after de-duplication only one Aadhaar shall be generated, as the uniqueness is achieved through the process of demographic and biometric de-duplication.

8.2 Services that have been linked with Aadhaar

Aadhar card is a mandatory requirement at a number of places and for a number of applications. Some of these are mentioned below:

- Iron Ore or Limestone workers need Aadhar Card for availing house subsidy.
- This card is mandatory to avail supplementary nutrition program.
- Aadhaar is mandatory for all the farmers who want to take crop insurance benefits and for people who are eligible for subsidized food grains.
- Aadhaar is also mandatory to undertake training under Integrated Child Development Services in the Ministry of Women and Child Development.
- It has also been made mandatory to earn benefits under Grih Kalyan Kendra scheme.
- Aadhaar has also been made mandatory for financial support under the National Mission for Empowerment of Women.
- This is also mandatory for e-panchayat training benefits and for all the students who wish to avail central scholarships at college level.
- Soil Health Card scheme and Soil Health management scheme now requires Aadhaar as a mandatory document.
- If you wish to take supplementary meals at creches, you require Aadhar card mandatorily.
- All the maternity benefit program as well as Integrated Child Protection Scheme requires Aadhar card.
- For women candidates looking to avail vocational training, loans and other schemes, Aadhaar has been made mandatory.
- Aadhaar is also mandatory for disabled children between the age group 6-14 who are eligible for benefits under the Sarva Shiksha Abhiyan.
- Linking Aadhaar with bank accounts, Money can be transferred using aadhar number.

- Victim of the Bhopal Gas Tragedy need Aadhar card to apply for compensation from the government.
- The most recent announcement was of the Aadhar card being made mandatory for salaried professional to file Income Tax return.
- Every citizen who wants to apply for a fresh PAN card now must hold a valid Aadhar card as a pre-requisite.

8.3 Privacy Concerns with Aadhaar

The debate engendered by the Aadhaar project has propelled India from being a predominantly pre-privacy society to one in which privacy protection in digital databases has emerged as a major national concern. The welcome and scholarly Supreme Court judgment has upheld privacy as a fundamental right, and informational self-determination and the autonomy of an individual in controlling usage of personal data have emerged as central themes across the judgment. The main privacy concerns with Aadhaar are:

- **Identity theft.** Aadhaar is vulnerable to illegal harvesting of biometrics and identity frauds because biometrics are not secret information.4,11 Moreover, possible leakage of biometric and demographic data, either from the central Aadhaar repository or from a point-of-sale or an enrollment device, adds to the risk.
- **Identification without consent using Aadhaar data.** There may be unauthorized use of biometrics to identify people illegally. Such violations may include identifying people by inappropriate matching of fingerprint or iris scans, or facial photographs stored in the Aadhaar database, or using the demographic data to identify people without their consent and beyond legal provisions.
- **Correlation of identities across domains.** It may become possible to track an individual's activities across multiple domains of service using their global Aadhaar IDs, which are valid across these domains. This would lead to identification without consent.
- **Illegal tracking of individuals.** Individuals may be tracked or put under surveillance without proper authorization or legal sanction using the authentication and identification records and trails in the Aadhaar database, or in one or more authentication-requesting-agencies' databases. Such records may reveal information on location, time, and context of authentication and the services availed.

8.4 Steps Taken by the Government

The Supreme Court has upheld the constitutional validity of Aadhaar albeit with some riders. The five-judge constitution bench, led by Chief Justice Dipak Misra, struck down Section 57 of the Aadhaar Act which allowed not only the government but also any "body corporate or person" or private entity to demand Aadhaar.

Aadhaar will no longer be required for availing services such as:

Bank Account: The Supreme Court today ruled that Aadhaar is not needed for opening a bank account. Banks have been chasing customers and asking them to update Aadhaar numbers with their bank accounts. However, following today's verdict, banks will not ask you to

link your Aadhaar to your account.

SIM card: With SC striking down Section 57 of the Aadhaar Act, private companies can no longer ask for your Aadhaar. Telecom operators were earlier asked by Department of Telecommunications (DoT) to conduct Aadhaar-based verification of mobile phone connections.

Appearing for competitive examinations and school admissions: In a big relief to students, the Supreme Court on Wednesday ruled that it is not mandatory to provide Aadhaar to register or appear for National Eligibility cum Entrance Test (NEET), University Grants Commission (UGC) and Central Board of Secondary Education (CBSE) exams.

Services which needs to be linked with Aadhaar:
PAN card: Linking of Aadhaar with PAN card continues to be mandatory under section 139AA of the Income Tax Act. Tax evaders used to create multiple PAN cards in order to avoid taxes. The government had on various occasions extended the deadline of linking PAN with Aadhaar

Filing income tax returns (ITR): As Aadhaar-PAN linking continues to be necessary, you'll need the same for filing income tax returns.

Welfare schemes: Aadhaar continues to be mandatory for availing benefits under various government-run social welfare schemes and subsidies. Justice Bhushan said the Central government had given sufficient reasons to uphold Section 7 of Aadhaar Act, which deals with grant of subsidies and welfare benefits.

8.5 Which data needs to be collected and from where

- For individuals

Data	Source
Personal Information	Government Database
Credit score/ rating	Banks
Behavioural nature	Workplace
Crime records	Government Database

- For companies

Data	Source
Company Information	Government Database
Business credit score	Banks
Trustworthiness	Employees and Partners
Ratings from suppliers and customers	Company Database

8.6 Activities on which social credit system of India should be based

- Behaviour

Human Behaviour is a very important factor and it determines how the mind of the person behaves.

- Trustworthiness

Loyalty is rarer than anything else so trust factor is an important deciding factor

- Ratings from suppliers and customers

Every company has a rating and in layman terms, that rating represents the value of the company.

8.7 How should people be rewarded

- Discounts on energy bills
- Rent things without deposits
- Better interest rates at banks
- Will reportedly get more matches on dating websites.
- Streamlined administrative procedures.
- Fewer inspections and audits
- Fast-tracked approvals.

8.8 How should people be penalized

- Banning you from flying or getting on the train.
- Throttling internet speeds.
- Banning from the best schools.
- Stopping from getting the best jobs.
- Keeping out of the best hotels.
- Being publicly named as a bad citizen.

8.9 Technologies used

- **Big data systems**

The goal of the big data system is to form data assets through the collection and exchange of a wide range of data sources, through data governance and development, and in accordance with the internal and external sharing needs of the government, thereby proving a comprehensive, efficient, and reliable data supply chain.

- **Fifth Generation cellular network**

Xinhua holds that 5G in combination with big data and AI, will take revolutionary steps towards a digital economy and this will make a huge positive impact on India's future economy.

- **Data analytics**

Behavioural data can include: "collected information on academic records, traffic violations, social media presence, friendships, adherence to birth control regulations, employment performance, consumption habits, and other topics"

- **Artificial Intelligence**

It would play a huge role in identifying people and detecting their expression along with predicting future behaviour.

8.10 Algorithms to be used

- **C4.5 Algorithm**

C4.5 is used to generate a classifier in the form of a decision tree from a set of data that has already been classified. Classifier here refers to a data mining tool that takes data that we need to classify and tries to predict the class of new data. Every data point will have its own attributes. The decision tree created by C4.5 poses a question about the value of an attribute and depending on those values, the new data gets classified. The training dataset is labelled with lasses making C4.5 a supervised learning algorithm

- **K-mean Algorithm**

One of the most common clustering algorithms, k-means works by creating a k number of groups from a set of objects based on the similarity between objects. It may not be guaranteed that group members will be exactly similar, but group members will be more similar as compared to non-group members. As per standard implementations, k-means is an unsupervised learning algorithm as it learns the cluster on its own without any external information.

- **Apriori Algorithm**

Apriori algorithm works by learning association rules. Association rules are a data mining technique that is used for learning correlations between variables in a database. Once the association rules are learned, it is applied to a database containing a large number of transactions. Apriori algorithm is used for discovering interesting patterns and mutual relationships and hence is treated as an unsupervised learning approach. Thought the algorithm is highly efficient, it consumes a lot of memory, utilizes a lot of disk space and takes a lot of time.

- **Naive Bayes Algorithm**

Naive Bayes is not a single algorithm though it can be seen working efficiently as a single algorithm. Naive Bayes is a bunch of classification algorithms put together. The assumption used by the family of algorithms is that every feature of the data being classified is independent of all other features that are given in the class. Naive Bayes is provided with a labelled training dataset to construct the tables. So it is treated as a supervised learning algorithm.

- **CART Algorithm**

CART stands for classification and regression trees. It is a decision tree learning algorithm that gives either regression or classification trees as an output. In CART, the decision tree nodes will have precisely 2 branches. Just like C4.5, CART is also a classifier. The regression or classification tree model is constructed by using labelled training dataset provided by the user. Hence it is treated as a supervised learning technique

8.11 Block diagram

- **For individuals**

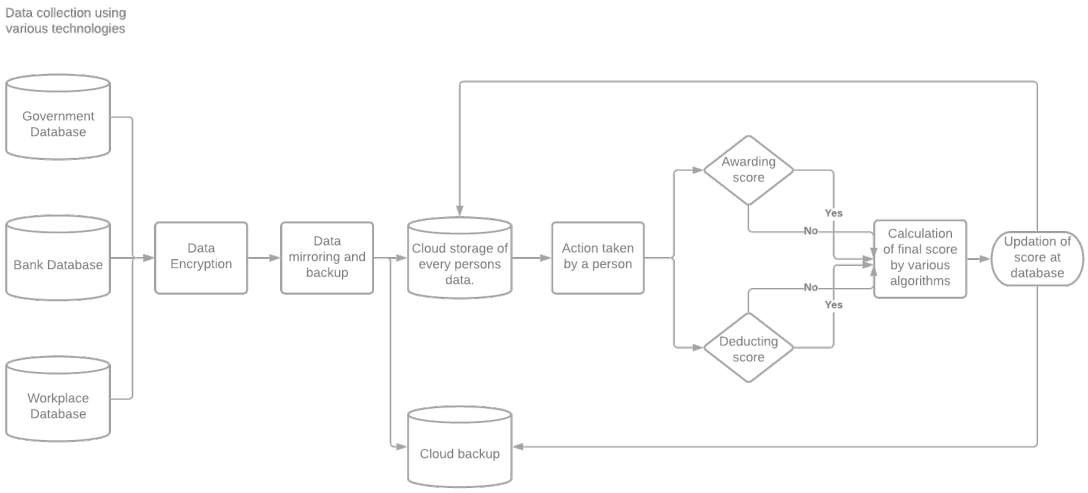

Figure 2: Implementation of Aadhaar based social credit system for individuals

- **For companies**

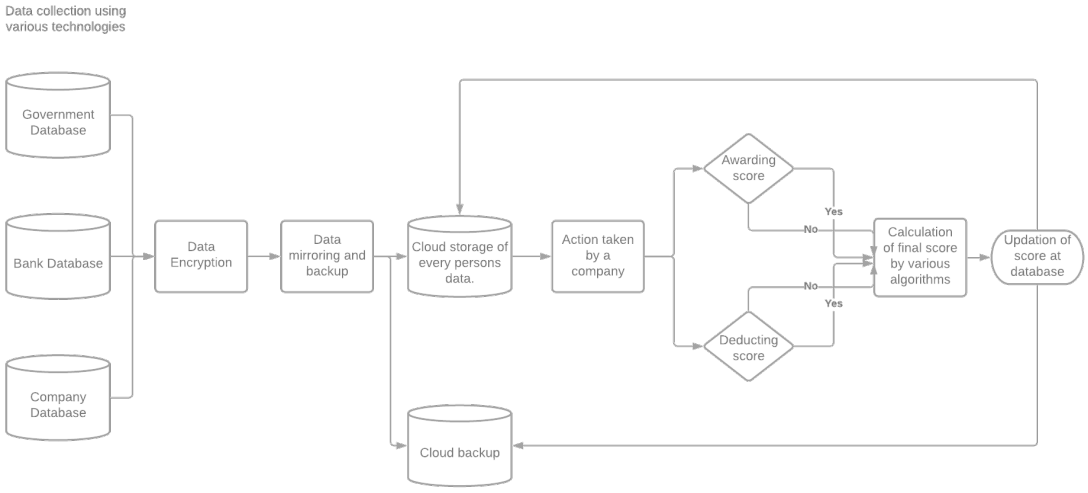

Figure 3: Implementation of Aadhaar based social credit system for companies

8.12 Issues with the social credit systems for India

The foremost problem in implementation of such a social credit system in India would not be the requirement of technological resources, but the incompatibility of such a system with the democratic set up of India and the fundamental rights of its citizens. We need to consider the implications this system will have on the citizens, not just the merits or demerits of the system in isolation. One could say that Article 21 of the Indian Constitution has been one of the most dynamic provisions. There are numerous rights that have been read into this provision such as right to livelihood, right to privacy, right to education etc. (One may find more on this here). This article would be the greatest impediment to the implementation of such a system.

The social credit system proposes to bar people from travelling in trains, renting quality apartments in good neighborhoods, holding certain jobs etc. These restrictions seem impractical for a democracy like India, where people enjoy these as a matter of rights. Further, barring children from getting admission in any educational institution due to their parents having poor social rating may be in direct contravention of Article 21-A, the right to education.

Moreover, since the grading does not take into account the distinction between a civil and criminal wrong, a person's poor rating might very well be a result of a mere delay in the payment of bills or violation of the traffic rule. The system is draconian as it imperils the future prospects of an individual without making informed assessment of the activities. The Indian judiciary has held that the reformation of offenders is an integral part of the justice system in India. The idea is that an individual must be given an opportunity to reform himself to blend back with society. For a system which follows rigorous policy towards past offenders with respect to housing and other public amenities, thereby restricting their chances of merging back with mainstream society, India might not seem a very welcome place.

Another hurdle that shall come into the way of reproducing such a system into the Indian set-up would be the rich jurisprudence on an individual's right to privacy. The system involves creation of a nationwide database with comprehensive records of citizens. This entails collection, storage and categorization of data regarding each individual of the nation. The aforementioned poses the question of authorities indulging in Cyber-profiling. Profiling in any form by any institution, be it private organizations creating a consumer database or by governmental authorities, has been held as an infringement of the right to privacy in the K.S. Puttaswamy v Union of India judgement. The right to privacy has led to major restrictions on the legislature with respect to passing future legislations.

9 References

Books

- William Stallings, Cryptography and Network Security, 7th edition, Pearson Education
- Understanding Cryptography by Christof Paar

URLs

- https://searchsecurity.techtarget.com/definition/RSA
- https://www.geeksforgeeks.org/rsa-algorithm-cryptography/
- https://www.thesslstore.com/blog/what-is-homomorphic-encryption/
- https://www.venafi.com/blog/homomorphic-encryption-what-it-and-how-it-used
- https://blog.openmined.org/build-an-homomorphic-encryption-scheme-from-scratch-with-python/
- https://www.investopedia.com/terms/c/creditrating.asp
- https://www.kaspersky.com/blog/social-scoring-systems/
- https://en.wikipedia.org/wiki/Social-Credit/-System
- https://publicphilosophy.ucsc.edu/cases/chinas-social-credit-system/
- https://www.maplecroft.com/insights/analysis/chinas-social-credit-system-raises-risks-for-corporates/
- https://www.eurekastreet.com.au/article/living-in-australia-s-social-credit-dystopia
- https://www.sundayguardianlive.com/news/chinas-social-credit-program-creeps-canada
- https://www.fastcompany.com/90394048/uh-oh-silicon-valley-is-building-a-chinese-style-social-credit-system
- https://www.scmp.com/economy/china-economy/article/3096090/what-chinas-social-credit-system-and-why-it-controversial
- https://research.aimultiple.com/privacy-enhancing-technologies/
- https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/what-is-personal-data/what-is-personal-data/
- https://uidai.gov.in/my-aadhaar/avail-aadhaar-services.html
- https://www.thehindu.com/news/national/aadhaar-verdict-live-updates-supreme-court-to-decide-the-validity-of-aadhaar-scheme/article25044764.ece
- https://searchdatamanagement.techtarget.com/definition/big-data
- https://www.livemint.com/opinion/columns/opinion-why-india-needs-to-be-wary-of-china-style-social-credit-ratings-1550423726392.html

Research papers

- Chinese Views of Big Data Analytics by Derek Grossman, Christian Curriden, Logan Ma, Lindsey Polley, J.D. Williams, Cortez A. Cooper III
- https://www.oecd.org/finance/Chinas-Belt-and-Road-Initiative-in-the-global-trade-investment-and-finance-landscape.pdf
- Multiple Social Credit Systems in China By Chuncheng Liu University of California, San Diego
- Independent Study Report - Cryptography By M. Wasim Munir