Scientific Research Publishing

Journal of Information Security, 2022, 13, 312-322
https://www.scirp.org/journal/jis
ISSN Online: 2153-1242
ISSN Print: 2153-1234

Algorithms and Tools for Securing and Protecting Academic Data in the Democratic Republic of the Congo

Gulain Mugaruka Buduge[1,2], Jérémie Ndikumagenge[2], Justin Buhendwa Nyenyezi[3]

[1]Department of Management Computing for Companies, Section of Commercial and Computing Sciences, Teachers' Training Collage of Bukavu, Bukavu, Democratic Republic of the Congo
[2]Department of ICT, Faculty of Engineering Sciences, University of Burundi, Bujumbura, Burundi
[3]Department of Maths-Physics, Section of Exact Sciences, Teachers' Training College of Bukavu, Bukavu, Democratic Republic of the Congo
Email: mugarukabuduge@gmail.com, jeremie.ndikumagenge@ub.edu.bi, justinnyenyezi@gmail.com

How to cite this paper: Mugaruka Buduge, G., Ndikumagenge, J. and Buhendwa Nyenyezi, J. (2022) Algorithms and Tools for Securing and Protecting Academic Data in the Democratic Republic of the Congo. *Journal of Information Security*, **13**, 312-322.
https://doi.org/10.4236/jis.2022.134017

Received: June 1, 2022
Accepted: September 27, 2022
Published: September 30, 2022

Abstract

This paper deals with the implementation of algorithms and tools for the security of academic data protection in the Democratic Republic of the Congo. It consists principally in implementing two algorithms and two distinct tools to secure data and in this particular case, academic data of higher and university education in the Democratic Republic of the Congo. The design of algorithms meets the approach that any researcher in data encryption must use during the development of a computer system. Briefly, these algorithms are steps to follow to encrypt information in any programming language. These algorithms are based on symmetric and asymmetric encryption, the first one uses Christopher Hill's algorithm, which uses texts in the form of matrices before they are encrypted and RSA as one of the asymmetric algorithms, it uses the prime numbers that we have encoded on more than 512 bits. As for tools, we have developed them in php which is only a programming language taken as an example because it is impossible to use all of them. The tools implemented are based on the algorithms of Caesar, Christopher Hill, and RSA showing how the encryption operations are carried out thanks to graphical interfaces. They are only tools for pedagogical reasons to help students and other researchers learn how to use developed algorithms. We have not developed them for pleasure but rather to be used in any information system, which would prevent and limit unauthorized access to computer systems. They will not be used only for the management of academic fees but for any other information system, which explains and shows the complexity of the tools developed. We have not been able to solve the problems of versions for

the developed prototype, because if there is a new version later some functions may be obsolete, which would constitute the limitation of these tools. This work targets primarily the Ministry of Higher Education and Universities, which will make these results its own and implement them in order to solve the problem of intrusions, and unauthorized access to developers and researchers who will use tools already made instead of thinking about their development. We are trying to demonstrate the steps and the methodology that allowed us to reach our results, in the following lines.

Keywords

Computer, Security, Crypto Systems and Analysis, Algorithms, Encryption/Decryption

1. Introduction

Information technology currently occupies a prime position in almost all sectors of human activity. Business intelligence is a key factor in the economic stability of modern companies, especially multinationals. Information systems of companies are faced with big, namely numerous and voluminous data that must often undergo complex processing on a daily basis [1]. On the one hand, accumulated over time, the data proves to be an inexhaustible source of knowledge for decision-makers and managers of these institutions and therefore constitutes reliable decision-making support, routine activities of the company, and a strategic plan on the other hand.

Automated information processing offers to man tools, and advanced artificial means with a range of exceptional possibilities of operation which improve the speed of execution, the power of treatment, the precision of calculations, etc. They allow, among other things, to perform tasks whose realization was, only a few years ago, extremely complex, difficult, long, tedious, and sometimes costly if not unfeasible [2].

However, with the expansion of computer networks and the growing and increasing volume and number of data exchanged on a daily basis, many threats to data have emerged and have varied with variable characteristics. As an example, we will mention accidental threats that occur without premeditation, passive and active intentional threats that take place under voluntary intrusion, passive threats that are harmless in nature, and threats that disrupt the functioning of the environment system [3].

Note that active threats belong mainly to four categories including disruption, interception, modification, and fabrication [3].

As far as the integration of new technologies into their systems is concerned, companies believe that computerized management facilitates data processing. However, some intruders take advantage of this to disrupt the functioning of the information system and thus the entire company. Every company is a potential

target for hackers [4]. These threats or attacks make information technology systems vulnerable and lead some companies to incur considerable expenditure, often unbudgeted, sometimes causing them to go bankrupt. To remedy this, companies in collaboration with research institutions are combining efforts to put in place tools, techniques, technologies and means in order to secure and protect data and information saved or exchanged through various communication networks [5]. It is within this framework that the purpose of this article is to provide part of the solutions by designing and implementing encryption and decryption methods and tools likely to strengthen the security and protection of company data and information by means of almost unbreakable cryptographic algorithms.

2. Materials and Methods

2.1. Materials

Currently, any organization, company, or institution swears by the use of computer technology to ensure the backup and exchange of data. Academic data of the students of the DRC are for the most part stored on physical supports. As the process of digitization of the latter is at an advanced stage, we have an obligation to anticipate potential problems related to the digitization of sensitive data that would be subject to attacks and multifaceted threats and therefore vulnerable in the future. The topic of this article is therefore the security of academic data of all involved students throughout the DRC. This paper is based on the design and implementation of encryption algorithms for computer data stored or exchanged across various transmission networks using the RSA and Christopher Hill encryption algorithms [6]. Since the latter uses a certain number of theoretical concepts in their cryptosystems, a brief overview of this would provide an entry point into the tools and methods used.

Given the problems of managing the circulation of information on the network, we, teachers and computer security enthusiasts, found it urgent to look for a security mechanism for the data transiting on the web, especially in the digital platform of higher and university education in the Democratic Republic of the Congo.

To achieve our objectives, we resorted to symmetric encryption using the C. Hill algorithm and asymmetric encryption using the RSA algorithm.

2.2. Methods

Several methods are used to encode and decode information. These methods are developing in parallel with the constant evolution of computer software and tools for intercepting information for or from a third person. Among the commonly used methods, we will mention the **symmetric encryption** which is known as the secret key algorithm. It is a unique decoding method that must be provided to the recipient before the message can be decoded [7]. On the other hand, **asymmetric encryption** uses two different keys, public and private, that

are mathematically related and therefore interdependent to some extent. Specifically, the keys consist only of large numbers that have been paired together but not identical. In view of the above, we will implement two sets of cryptographic algorithms, one based on the Hill Symmetric cryptosystem and the other based on the RSA cryptosystem.

3. Results

3.1. Experimental Results

The first crypto system developed is made with the Hill algorithm which uses an n*n matrix as key and an n*m matrix as encrypted text, where the m is variable depending on the encrypted text. The second one was done with the RSA algorithm using prime numbers encoded on 512 bits. Since these will be used to secure academic data, we will nuance these two approaches by using the listings sowed on the encryption interfaces.

1) Example: Here is the text to be encrypted with C. Hill's algorithm in Figure 1.

L'informatique prend aujourd'hui.

The numerical result in matrix form is:

The randomly selected key matrix is as follows in Figure 2.

The result of the encryption in an encrypted matrix in Figure 3.

The result of the original text encryption transformed into a matrix is presented below in Figure 4.

This way of presenting numerical result shows only that the matrices were used [8]. We have just finished the encryption with the algorithm of C. Hill, then decryption will follow. In every decryption, one uses encrypted result. For this case, we will use the encrypted matrix then search for the inverse of the key matrix which will be multplied to obtain a clear matrix (Transformed into a source text). With these two matrices (the first obtained after encryption and the second from the key matrix), we can decrypt the message without any problems.

76	39	105	110	102	111	114	109	97	116	105
113	117	101	32	112	114	101	110	100	32	97
117	106	111	117	114	100	39	104	117	105	46

Figure 1. The text matrix to be encrypted.

5	7	6
2	3	4
2	5	3

Figure 2. The encryption key matrix.

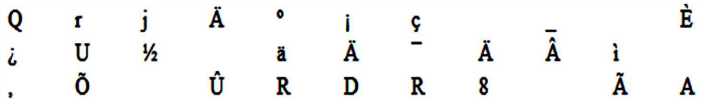
81	114	106	196	186	161	231	147	95	154	200
191	85	189	162	228	196	175	196	194	236	173
44	213	24	219	82	68	82	56	21	195	65

Figure 3. The numerically converted result of the numerical matrix.

Figure 4. The result converted into text of the numerical matrix.

a) First Matrix

This figure is the same as Figure 4.



This one in a numerical matrix:

This figure is also the same as Figure 3.

81	114	106	196	186	161	231	147	95	154	200
191	85	189	162	228	196	175	196	194	236	173
44	213	24	219	82	68	82	56	21	195	65

b) Second Matrix in Figure 5

At this level, the decryption consists of multiplying the second matrix by the first (key matrix inverse to encrypted matrix) in order to come back to the initial matrix that will be the clear text. The result is the following in Figure 6.

2) Example of the RSA algorithm

The RSA algorithm is an asymmetric cryptography algorithm; this means that it uses a *public* key and a *private* key (*i.e.* two different, mathematically linked keys). As their names suggest, a public key is shared publicly, while a private key is secret and must not be shared with anyone.

Principle of this algorithm

1) Key generation

The RSA works from two prime numbers, which we will call p and q. These two numbers must be very large, because they are the keystone of our encryption. Today, we use keys from 128 to more than 1024 bits, which represents decimal numbers from 38 to more than 308 digits! Once these two numbers are determined, let's multiply them. We note n the product $n = p \times q$, and $(p-1) \times (q-1)$.

Let us now look for a number e (less than $(p-1) \times (q-1)$), which must necessarily be prime with $(p-1) \times (q-1)$.

Let us then compute the inverse of e modulo $(p-1) \times (q-1)$), which we will note d.

$$d \equiv e^{-1} \bmod\big((p-1)(q-1)\big)$$

The couple (e, n) is the public key, and (d, n) is the private key.

2) Encryption and decryption

To encrypt a number, you just have to put it to the power e. The remainder modulo n represents the number once encrypted.

91	135	150
30	45	136
60	91	15

Figure 5. The decryption key matrix.

76	39	105	110	102	111	114	109	97	116	105
113	117	101	32	112	114	101	110	100	32	97
117	106	111	117	114	100	39	104	117	105	46

Figure 6. Initial result deciphered in matrix.

Encrypted message = Be encrypted message ^{e} mod n

To decrypt, we use the same operation, but to the power of d:

Be encrypted message = Encrypted message ^{d} mod n

Once e, d and n are computed, we can destroy p, q and $((p-1)(q-1))$, which are not necessary to encrypt and decrypt. Even worse, the private key d can be computed very quickly from p and q, so these numbers should not be kept.

Note: In general, the private key is then encrypted using symmetric encryption. This makes it possible to keep it in a safe way, because the key used by the symmetrical encryption does not have to be transmitted, and thus does not risk to be intercepted. At our level, we have applied this algorithm with other more complex aspects, as demonstrated in the following lines. We now focus on the data stored in the database management system. The RSA encryption algorithm is used to secure the identifiers [9]. First, the generation of random numbers encoded on 512 bits gives the result shown in the listing below. Since the RSA algorithm uses prime numbers, we have developed a tool to generate them randomly as much as we can. Its interface is in **Figure 7**.

The tool that we have created allows to generate randomly coded prime numbers on 512 bits, it is the first RSA tool that we implemented. It generates the numbers p, q, and e. The generated numbers are:

Numbers of the capture:

$p =$ 17142378547352120628412840096806956351432295584058942228482302549347254182341063401160449796329162443485535846702677433451768521881334928097448674126813997

$q =$ 33590833169965604257193504685961145912941241088036813986015524833327643894608621111405826404490315971695314820693605913783617344190008542990496397710666557

$e =$ 96046669129965744089072755795837504024332587048398215654876389341155525633219275116556597202867811523476157394537921676219119844670842244023877763801759877

Suppose we want to encrypt 65 which is the ASCII value of A. To begin with we need to calculate the value of n which is the product of p and q. Then calculate $(p-1)(q-1)$.

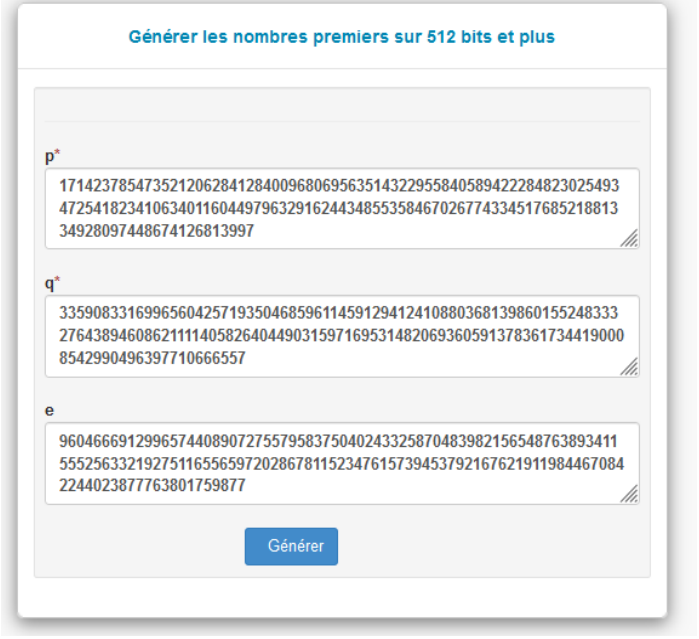


Figure 7. Tool designed for data encryption under RSA.

$n = 575826777920502404431124967283521690715726393197842077648265$
$666166568712315613822711737906855092641318870121538155559039$
$147344443125260510490572366339777441557608867798006325258029$
$450412761026445488325714606523492010645310759155796269464670$
$141765871839950343680658581653689838997451205055168843586693$
627398329

$(p-1)(q-1)$
$= 575826777920502404431124967283521690715726393197842077648265 66$
$616656871231561382271173790685509264131887012153815555903914 73$
$444431252605104905723663397773908243971504802814396516846676 44$
$658762071951653618850308994183262635861078846584952103865565 05$
$236153516282999118537034260361158513371169775564162178991777 6$

Having already n and e we can then generate the public key which is the couple formed by n and e, and from the public key we can encrypt the message.

Encrypted message = be encrypted message exponent e modulo n

Be encrypted message = 65

$e = 960466691299657440890727557958375040243325870483982156548763$
$893411555256332192751165565972028678115234761573945379216762$
$191198446708422440238777638017 5987$

$n = 575826777920502404431124967283521690715726393197842077648265$
$666166568712315613822711737906855092641318870121538155559039$
$147344443125260510490572366339777441557608867798006325258029$
$450412761026445488325714606523492010645310759155796269464670$
$141765871839950343680658581653689838997451205055168843586693$
627398329

Encrypted message
= 455512971847695085469730793079761372161640341956123148378709134836135570460781843615961144067452392205355394992027740884780923165365746220165759474995244726525240964927536246190868532357355268008138783755817748406007667398273004660367727944974495691250223467510127723412883131509987185727074934506336489690221250223467510127723412883131509987185727074934506336489690221

With the previous example, we find that if Mr. Bob sends message A to Alice with RSA encryption using the above parameters, Alice will receive 455512971847695085469730793079761372161640341956123148378709134836135570460781843615961144067452392205355394992027740884780923165365746220165759474995244726525240964927536246190868532357355268008138783755817748406007667398273004660367727944974495691250223467510127723412883131509987185727074934506336489690221 instead of A, as for her to decrypt the message if she has the private key. In the following part we will show how to decrypt the message.

Before decrypting the message in RSA, we need to calculate the value of d, which is the inverse of e modulo $(p-1)(q-1)$.

d = 770310650716094684905263403650319109705367415620859422729266185743414560946120644604519772368027323956964506786031689161149204317703603543474263464168753022251193366280111700542310118172050956687595721698308163374076552396504072499466513651806670296581175855788692893718291049182384305707470088978135 83621805

With the value of d, we can generate the private key which is the couple n and d.

So to decrypt the message, we use the following formula:

Having the necessary information n, d we can find the clear message which is the starting message.

After making all the calculations, the clear message is: 65. In sum, we have just presented two tools (cryptographic systems), the first is based on symmetric encryption while the second is based on asymmetric encryption. The first one implements the Hill algorithm which is based on matrices, which allowed us to rely on the transformation of the entered text into a matrix and the second one implements the encryption which is based on prime numbers, we used prime numbers encoded on 512 bits. In computer security, the more we think about increasing the bits of the prime numbers to be encrypted, the more the security of the information increases.

3.2. Algorithms

We managed to deploy two crypto systems that use symmetric and asymmetric encryption respectively. The first one used Hill's algorithm while the second implemented RSA's. Let us now note the particularities and nuances constraints of the two implemented algorithms.

For the algorithm designed based on Hill's crypto system:

1) The text to be encrypted of n characters;

2) The invertible square matrix;

3) The matrix determinant must be > 0;

4) The number of rows in the encrypted matrix depends on the columns of the key, while the number of columns is the resultant of the length of the text to be encrypted and the number of columns in the key if the former is a multiple of the latter;

5) The encrypted text converted into an ASII code value is transformed into a matrix according to the key matrix;

6) The product of the key matrix and the text to be encrypted.

For the algorithm based on the RSA cryptosystem:

1) The recovery of the ASCII value of each character in the text to be encrypted;

2) Triggering the concatenation procedure of the found ASCII values;

3) The created function returns the concatenated ASCII codes;

4) The transformation of the returned value into Big Integer;

5) The use of the exponent **e modulo n**;

6) The encrypted message is saved in the database.

The major and main part of our work consisted of the implementation of encryption and decryption algorithms which consist in securing all the information and/or data exchanged on the network including the academic data of the students.

In addition to these algorithms, we have implemented tools to encrypt information using symmetric and asymmetric algorithms. Initially, these tools are used to demonstrate what we have achieved, and then we will use some of these features in all the computer systems we implement. This will ensure the security of information that will be stored in the database management systems, while controlling unauthorized access and availability of data on our servers.

4. Conclusions

Because computer systems play a central role in companies, their security is vital. They, therefore, constitute a strategic issue. Thanks to the development of networks, computer systems are increasingly open to the outside world. Computer security, which is an aspect that promotes the survival and viability of any company, is therefore essential. New technologies are everywhere and the Internet is accessible almost without limit, however, we are witnessing a proliferation of viruses, malware, and various fraud attempts on the network. It is therefore very important to control access in order to protect the company's resources. In case of failure, the consequences can be both financial and legal. Data security has become a complex task in recent times, as the more important information a company has, the more likely it becomes to be attacked by hackers. Ensuring data security does not only mean protecting data from unauthorized access or

making it confidential, but also ensuring that the generic principles of information security (authentication, confidentiality, integrity, non-repudiation, availability) are respected. The first two aspects require sufficient time to implement appropriate algorithms to address them. Unfortunately, many companies do not take into account the importance of the security system and its impact on the company, whereas the security policy and architecture are in line with technological evolutions, preserve the company's information assets and, moreover, promote its development. This is why most companies call in experts after having already lost their data.

The implementation of a computer system precedes the notion of information security because we cannot secure a system that we do not have, which is the major problem of the Ministry of Higher and University Education of the DRC. Our first project consists of setting up the Ministry's computer system and the second in setting up security mechanisms that consist in securing the sector's essential information against the main online threats: unauthorized access, data hacking, intrusion, and virus. The implementation of this mechanism addresses the concerns of many users. The vulnerability of data on a national scale has been of great concern to us and has led us to reflect on the security mechanisms that guarantee the implementation of appropriate security and protection tools. Encryption and decryption algorithms designed, developed, and implemented are proving to be an important solution to the aforementioned threats, including attacks and intrusions. In general, we have managed to develop a security mechanism that allows developers to protect their information in real-time on the network and especially that stored in a database management system. When developing an information system, developers who want to secure their information must use these two tools.

These tools and techniques for securing academic data will be used in the computer system of the Ministry of Higher and University Education of the Democratic Republic of the Congo. We believe that these mechanisms will contribute to data security in the cloud in general and in the higher education space in the Democratic Republic of the Congo in particular.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Carlier, A. (2013) Business Intelligence and Management. Afnor Editions, La Plaine Saint-Denis, 80-200.

[2] Ndjate, L. (2014) Mise en place d'un crypto systeme pour la sécurité des donnée et la détection d'intrusion dans un supermarché. https://www.memoireonline.com/01/16/9388/m_Mise-en-place-dun-crypto-systeme-pour-la-securite-des-donnee-et-la-detection-dintrusion-da0.html

[3] Fung, K.T. (2005) Network Security Technologies. Auerbach Publications, Boca

Raton.

[4] Alaba, F.A., Othman, M., Hashem, I.A.T. and Alotaibi, F. (2017) Internet of Things Security: A Survey. *Journal of Network and Computer Applications*, **88**, 10-28. https://www.sciencedirect.com/science/article/abs/pii/S1084804517301455 https://doi.org/10.1016/j.jnca.2017.04.002

[5] Eastaway, R. and Wyndham, J. (2001) Pourquoi les bus arrivent-il toujours par trois? *Flammarion*, **3**, 95-107. https://doi.org/10.1007/978-88-470-1122-9_14

[6] Hill, L.S. (1929) Cryptography in an Algebraic Alphabet. *The American Mathematical Monthly*, **36**, 306-312. https://doi.org/10.1080/00029890.1929.11986963

[7] Internet Society (2021) Comprendre les bases du chiffrement. https://isoc.org.gn/event/comprendre-les-bases-du-chiffrement/

[8] Edward, L.R. (2000) Cryptological Mathematics. The Mathematical Association of America, Washigton, 124-140.

[9] Douglas, S. (2001) Cryptographie, Théorie et pratique. Vuibert, Paris, 12-16.