Ilorin Journal of Science
Volume 1, Number 1, 2014, pp. 50 – 60 (Printed in Nigeria)
ISSN: 2408 – 4840 © 2014 Faculty of Physical Sciences, University of Ilorin

ILORIN
JOURNAL OF SCEINCE

ILJS-14-002

Design and Implementation of a Yoruba Language Crytosystem

Abikoye, O. C., Akintola, A. G. and Ibrahim, S.

Department of Computer Science, University of Ilorin, Ilorin, Nigeria

Abstract

Security has always been a bedrock for transaction and protection of human lives and until the existence of artificial intelligence, humans have learned to trust themselves for each other safety or safety of information except for few times where security was threatened *by* breach of trust or betrayal by persons which often led to enormous crisis. Recent research works have been able to achieve some high level of success in security of information. However, communication in Yoruba language using the Yoruba alphabet and symbols is still at its infancy. In other to achieve acceptable quality of information security of Yoruba alphabet and symbols in text, there is need for a robust security system. There are multiple layers of security, requirements and features which need automated technicality for assurance of effective security. In this paper, encryption and decryption of information written in Yoruba as text is achieved using Caesar Cipher Algorithm to carry the Yoruba people along in the advancement taking place in the world of information technology. It adopted a symmetric encryption algorithm that uses one key for both encryption and decryption and used digits as its ciphertext which is a novel means of ensuring difficulty in breaking the codes as patterns with digits are not easily recognizable.

Keywords: Ciphertext, Encryption, Decryption, Cryptography, Yoruba

1. Introduction

Increased level of security gives comfort to people, especially a time like this when the Internet provides essential communication between many people and is being increasingly used as a tool for commerce, education, information and networking. Security becomes an extremely important issue to deal with. There are many aspects to security and many applications, secure commerce and payments to private communications, authentication and protecting passwords. Within the context of any application-to-application communication, there are some specific security requirements, including (Bishop, 2005):

Corresponding Author: Abikoye, O.C.
Email: abikoye.o@unilorin.edu.ng

- *Authentication:* The process of proving one's identity.
- *Privacy/confidentiality:* Ensuring that no one can read the message except the intended receiver.
- *Integrity:* Assuring the receiver that the received message has not been altered in any way from the original.
- *Non-repudiation:* A mechanism to prove that the sender really sent this message.

An essential aspect for secure communications is that of cryptography.

Cryptography is is the science of securing data. It deals with using mathematics to encrypt and decrypt data. It enables one to store sensitive information or transmit it across insecure networks so that it cannot be read by anyone except the one for which it was intended (Bradley & Irwin, 2009).

In data communication and telecommunications, cryptography is essential when communicating over any untrusted medium, which includes just all networks, particularly the Internet.

Cryptography not only protects data from theft or alteration, but can also be used for user authentication. In general, three types of cryptographic schemes are used to accomplish these goals: secret key (or symmetric) cryptography, public-key (or asymmetric) cryptography, and hash functions. In all cases, the initial unencrypted data is referred to as plaintext. It is encrypted into ciphertext, which will in turn be decrypted into usable plaintext (Bishop, 2005).

A cipher is a pair of algorithms that create the encryption and decryption. The detailed operation of a cipher is controlled by both the algorithm and in each instance by a "key". This is a secret (ideally known only to the communicants), usually a short string of characters, which is needed to decrypt the ciphertext. Cryptography can be strong or weak; its strength is measured in the time and resources it would require to recover the plaintext. The result of strong cryptography is cipher text that is very difficult to decipher without possession of the appropriate decoding tool. A "cryptosystem" is the ordered list of elements of finite possible plaintexts, finite possible ciphertext, finite possible keys, and the encryption and decryption algorithms which correspond to each key. Keys are important, as ciphers without variable keys can be trivially broken with only the knowledge of the cipher used and are therefore useless (or even counter-productive) for most purposes. Historically, ciphers were often used directly for encryption or decryption without additional procedures such as authentication or integrity checks (Calloway, 2008).

Statement of the Problem

There's a whole lot of information that we don't want other people to see or have access to, such as: Credit-card information, Private correspondence, Personal details, Sensitive company information, Bank-account information, confidential information about clients and a whole lot more. This brings about the need to know who we are sharing confidential information with guard against unauthorized access to information.

In other to take full advantage of technology, an encryption and decryption system is desired for Yoruba letters.

2. Materials and Methods

In the midst of terrorism and other forms of crimes, it is important to improve protection of information and persons against loss or damage. The problem of a secure means of information transmission over any network channel is of major concern and has received major attention from researchers over the world with each trying striving towards advancement in creating permanent and more effective means of ensuring security.

Security is the prevention and protection against assault, damage, fire, fraud, invasion of privacy, theft, unlawful entry and other such occurrences caused by deliberate actions (Cohen, 1995). In computing it's the extent to which a computer system is protected from data corruption, destruction, interception, loss or authorized access (Harris, 2010).

The value of information is determined by the characteristics it possesses. Information security has the duty of making sure that the value of information doesn't decrease due to any form of damage caused physically or electronically. In view of this objective information security is conceptualized in the following terms

- Confidentiality*:* Ensuring that no one can read the message except the intended receiver. For most type of data set (transactions, communication, entertainment) confidentiality is required.
- Integrity*:* This refers to the ability of data to be complete or undivided. Integrity could be related to as consistency in the available information or data. It assures the receiver that the received message has not been altered in any way from the original.
- Availability**:** in information security availability refers to the accessibility of information by authorized persons for any information to serve its purpose, the information must be available when needed.
- Authenticity: it is very important in information security that information available for use is real and genuine. Authentication is used to verify that a person is also who he or

she claims to be. In Authenticating a user or person some form of proof is required e.g. passwords, finger print, voice or face recognition all these ensures that the information given to who it can be trusted with. Authentication and authorization works hand in hand. Authorization is an act of given consent or approval to a user on what he or she can do with information available to them.

- Non-repudiation: in securing information, security is said to be strong when users of authentication cannot be refuted. This is a mechanism to prove that the sender really sent this message.
- Accuracy: this refers to when Information is free from mistakes or errors and it has the value that the end user expects. If information has been intentionally or unintentionally modified, it is no longer accurate.
- Utility: utility of information refers to the quality or state of having value for some purpose or end. Information has value when it can serve a particular purpose. This means that if information is available, but not in a format meaningful to the end user, hence it is not useful.
- Possession: possession of information is the quality or state of ownership or control of some other characteristics. While a breach of confidentiality always results in a breach of possession, a breach of possession does not always result in a breach of confidentiality. These major concepts of securing information are interdependent they rely on each other for ensuring strong security of the information, data and transactions. In all users should be assured that they can trust the information they use (History of Cryptography, 2013).

Forms of Security Measures

Implementing security measures can be seen as ways of enforcing security. This measures take different forms depending on the effect of the perceived risk of attackers or damages to the said information although all measures of security taken are done objectively to achieve the aim of protecting unauthorized persons from gaining access to resources and ensuring that authorized persons can access the information they need. Implementing security in this context is done through Authentication, Encryption, and Cryptography.

Authentication

This is the process of demonstrating or of proving one's identity. Authentication also known as User identification is based on three factors – the knowledge factor, the possession factor and the biometrics factors. The knowledge factor uses what is known by a person, e.g. a PIN or password. The possession factor uses what is possessed by a person, e.g. a smart card. The biometrics factor uses a biological or behavioral characteristic about a person, such as a written signature or a fingerprint.

Encryption

Encryption is the process of encoding information in such a way that only the person (or computer) with the **key** can decode it.

Cryptography

Cryptography is the study of methods of sending messages in disguised form so that only the intended recipients can remove the disguise and read the message. The message may be converted using a 'code' (in which case each character or group of characters is substituted by an alternative one) or 'cipher' (in which case the message as a whole is converted, rather than individual characters. (Kessler, 1998)

Cryptography can be classified into three (Symmetric, Asymmetric and Hash function):

Symmetric Method:

This is also known as ‘secret key’ encryption. Symmetric encryption transforms plaintext into ciphertext using a secret key and an encryption algorithm. Using the same key and a decryption algorithm, the plaintext is recovered from the ciphertext. The key needs to be passed on to the recipient. This method uses a single key for both encryption and decryption. The security of the symmetric encryption method is completely dependent on how well users protect the key. If a key is compromised, then all messages encrypted with that key can be decrypted and read by an intruder. Also the fact that there is one key (or in some cases the two keys are directly related to each other), the key can be broken.

This is complicated further by how symmetric keys are actually shared and updated when necessary, because both users use the same key to encrypt and decrypt messages, symmetric cryptosystems can provide confidentiality, but they cannot provide authentication or non-repudiation. There is no way to prove who actually sent a message if two people are using the exact same key.

Algorithm that makes use of symmetric key cryptography includes Data Encryption Standard (DES), Triple DES (3DES), Blowfish, IDEA, RC4, RC5, and RC6

Asymmetric Method

The Asymmetric method is also known as a public-key method. The key holder has two keys--a private key (which only they know) and a public key (which is uploaded to a key server or given to people they want to correspond with). In this method of encryption each entity has different keys, or asymmetric keys. The two different asymmetric keys are mathematically related. If a message is encrypted by one key, the other key is required to decrypt the message.

When a person wants to encrypt a message or file, they use the Public key of the recipient to encrypt it. This ensures that only the recipient (or anyone with the private key) can read the message or file. If the sender wants to guarantee that they were the sender, they will use their private key to sign the message (and the recipient will use their public key to verify that it was sent by them).

It has better scalability and key distribution than symmetric systems, although it's slower than the symmetric system. (Pesante, 2008) Algorithms that uses asymmetric key includes RSA, Elliptic Curve Cryptosystem (ECC), Diffie-Hellman, El Gamal, Digital Signature Standard (DSS)

Hash functions

Hash functions, also called message digests and one-way encryption, and are algorithms that, in some sense, use no key. Instead, a fixed-length hash value is computed based upon the plaintext that makes it impossible for either the contents or length of the plaintext to be recovered.

Cipher

It is an algorithm for performing encryption and decryption. Ciphers make textual communication a mystery to anyone who might unduly intercept it. Hence, a cipher is a method used to encode characters to hide their values (Dulany, 2009).

Types of Cipher

There are two basic types of encryption ciphers namely: substitution and transposition (permutation).

Substitution Cipher

The substitution cipher replaces bits, characters, or blocks of characters with different bits, characters, or blocks. A substitution cipher uses a key to know how the substitution should be carried out. In the Caesar Cipher, each letter is replaced with the letter three places beyond it in the alphabet. This is referred to as a shift alphabet.

Transposition Cipher

The transposition cipher does not replace the original text with different text, but moves the original text around. It rearranges the bits, characters, or blocks of characters to hide the original meaning.

In a transposition cipher, permutation is used, meaning that letters are scrambled. The key determines the positions that the characters are moved to and rearranged as. This is a simplistic example of a transposition cipher and only shows one way of performing transposition. When introduced with complex mathematical functions, transpositions can become quite sophisticated and difficult to break (Raphael & Sundaram, 2012).

Methodology

The system is a Yoruba language cryptosystem, first we establish the letters in Yoruba alphabet which are a total of 25 (7 vowels & 18 consonants). These vowels carry diacritics to indicate the tone of the language; an acute accent (`) for the high tone commonly called “Do”, a grave accent (′) for the low tone which is also known as “Mi” and a macron accent(¯) for the middle tone which is often left blank and is called “Re”. These letters 39 in all for the plain text (i.e. text to be encoded) and the digits 1-39 is the ciphertext.

TABLE 1: YORUBA LETTERS AND THE EQUIVALENT DIGITS

A	B	D	E	Ẹ	F	G	Gb	H	I	J	K	L	M	N	O	Ọ	P	R	S	Ṣ	T	U	W	Y
a	b	d	E	ẹ	F	g	gb	h	i	J	K	l	m	n	O	ọ	p	r	s	ṣ	t	u	W	y
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

TABLE 2: YORUBA VOWEL (CARRYING DAICRITICS) AND THEIR EQUIVALENT DIGITS.

Á	À	É	È	Ẹ́	Ẹ̀	Í	Ì	Ó	Ò	Ọ́	Ọ̀	Ú	Ù
Á	À	É	è	ẹ́	ẹ̀	Í	ì	ó	Ò	ọ́	ọ̀	Ú	Ù

26	27	28	29	30	31	32	33	34	35	36	37	38	39

The implementation is done using Java programming tools.

Strength

- The simplicity of the above algorithm makes it a first choice; it employs a simple non technical method of shifting plain text to its corresponding cipher text. Basically there are 2 methods for implementing a shift cipher: either count through the alphabet letter after letter, or write down the plaintext letter with a new ciphertext lettet for each shift amount.
- This system is also very flexible and allows the user decide on the number of shifts for substitution.
- System Model

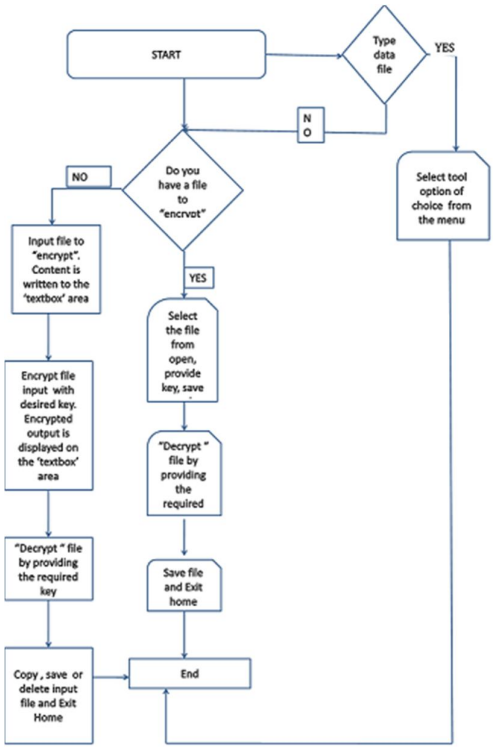

FIGURE 1: MODEL SHOWING THE FLOW OF THE SYSTEM

Caesar Algorithm

Step1: Initialize all variables

Step2: For I ranging from 1 to duration

Step3: Set chr to the ith letter in the duration

Step4: Find the key of chr in the Yoruba_text

Step5: Merge the key'th letter in the cipher_letters

Step6: Return ciphertext

System sample code

This is the step by step instruction for implementing the program the proposed system. It shows a well commented algorithm used for encrypting and decrypting.

Encryption code

```
For i → RTE0 To charsInFile – 1
  Letter → RichTextBox2.Text.Substring(i, 1)
    StreamToWrite.Write (Asc(letter)Xor code)
     StreamToWrite.Write(" ")
     Next
```

Decryption code

```
For i →0 To numbers.Length – 1
    Number-CShort(numbers(i))
      Ch-Chr(Number Xor code)
    Decrypt-Decrypt & ch
    Next
    RichTextBox2.Text →Decrypt/Display the text in the textbox area/
```

3. Results and Discussion

Input File

The figure below shows a sample plain text written in Yoruba.

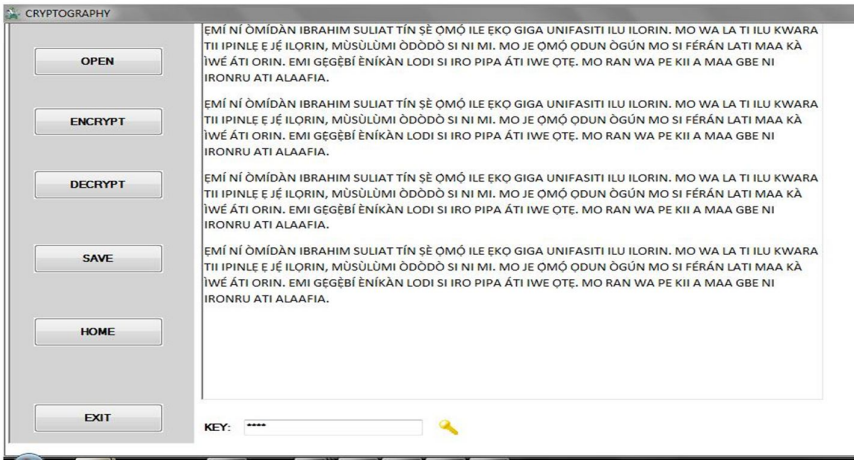

FIGURE 2: SAMPLE INPUT TEXT

Output File

The figure below shows the resulting cipher output when the plain text has been encrypted with a key.

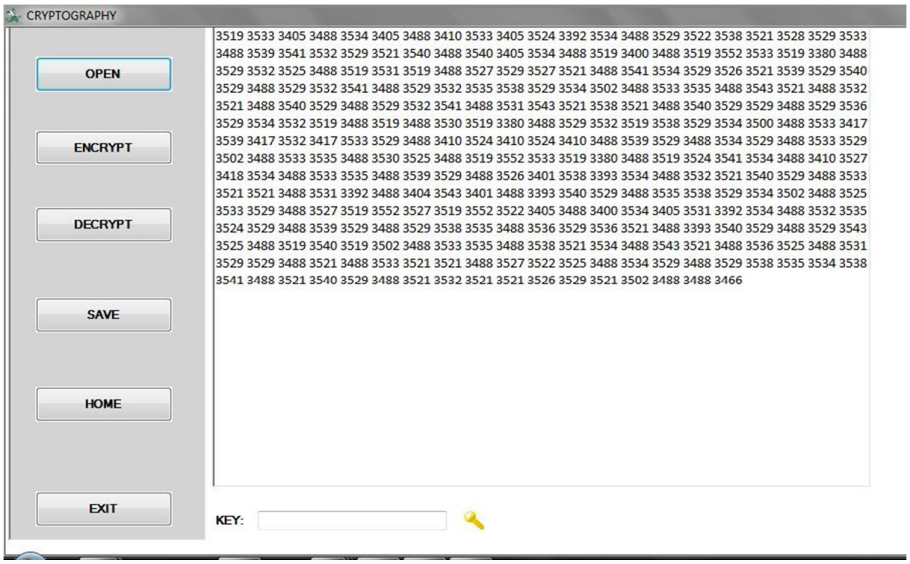

FIGURE 3: ENCRYPTED TEXT

Decryption

The figure below shows the decrypted file which is the same as the original plain text encrypted.

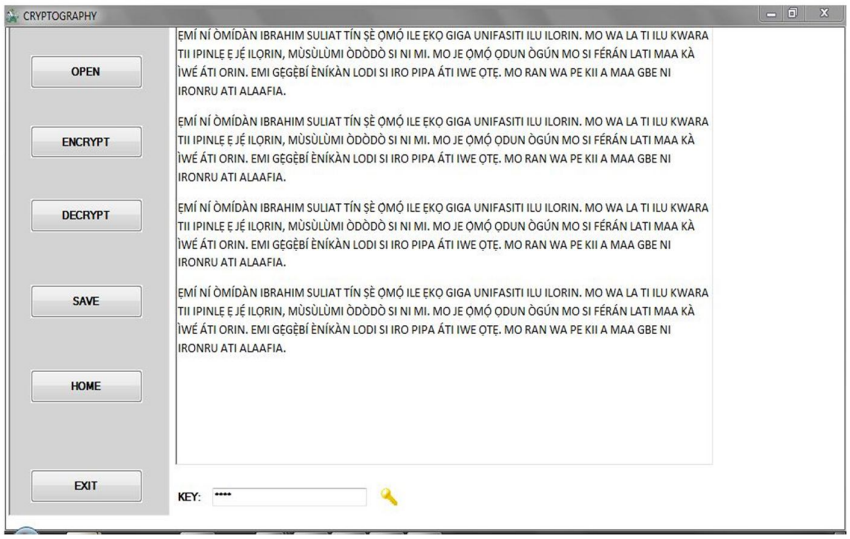

FIGURE 4: DECRYPTED FILE

Encryption techniques have been proved to give substantial security to any system on which it is implemented. The need for securing information can't be over emphasized as events and happenings that turn out due of lack of doing so could be damaging and difficult to control.

Yoruba language is spoken by over forty (40) million people in Nigeria, Togo, Benin and in communities in other parts of Africa, Europe and the America (Wikipedia, 2014). It doubles up as a means of passing information across to people whether over land or across a network and has also reached its peak of acceptance globally such that many systems are creating options for the language either by translation or voice authentication, this goes on to prove that creating a Yoruba language cryptosystem is an advancement on linguistics and data security.

Having implemented and tested this cryptosystem, it can be concluded that cryptography has an advantage over other information security techniques. The system maintains the integrity of the information as nothing new is added to the message asides transcribing as a form of security.

Acknowledgement

The authors acknowledge the comments of the reviewers for their suggestions and comments in improving the quality of the manuscripts. Also, we are grateful to the Faculty of Physical Sciences, University of Ilorin, Ilorin, Nigeria for sponsoring of this Maiden Edition.

References

Bishop, M. (2005): *Computer security: Art and Science*, 4th edn., Boston, MA. Addison Wesley. 16-21.

Bradley, C., & Irwin, B. (2009). Literature Survey: An investigation into the field of cryptography and cryptographic protocols. Retrieved February 26, 2013 from http://www.cs.ru.ac.za/research.g06c5476/Honours/LiteratureReviewCowie.pdf.

Calloway, D. (2008). Cryptography and Network Security: Literature Review of Cryptography & Network Security.Retrievedhttp://www.academia.edu/DanCalloway.

Cohen, F. (1995): A Short History of Cryptography Retrieved January 22, 2013 from http://www2.itu.edu.tr/~orssi/dersler/crypography/chap2-1.pdf.

Dulaney E. (2009): *CompTIA Security+ Study Guide*, 4th Edn., Wiley Publishing Inc., Indianapolis, Indiana. 487-489.

Harris, S. (2010). *Cryptography all-in-one exam guide* (6th ed.). New York, NY: McGraw-Hill Retrieved from http://www.ccure.org/documents/crytography/cisspallinone.pdf

History of Cryptography (2013). Retrieved February 27, 2013 from http://www.logicalsecurity.com/resources/whitepaper/cryptography.

Kessler, G.C. (1998). Handbook of applied cryptography; An overview of cryptography Retrieved March 13, 2013 from http://www.garykessler.net/librarycrypto.html.

Pesante, L. (2008): Information security basics. Retrieved February 27, 2013 from http://www.us-cert.gov/reading_room/infosecuritybasics.pdf.

Raphael, J., & Sundaram, V. (2012). Cryptography and Steganography; International Journal Computer Technology, **2**(3), 626-630.

Wikipedia (2014). Yoruba people. Retrieved March 5, 2014 from http://www.wikipedia.org/wiki/yoruba_people.