ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 5, Issue 6, December 2015

DECRYPTION SYSTEM OF THEMATIC TEXTS IN SPANISH USING FREQUENCY ANALYSIS INCLUDING UNIGRAMS, BIGRAMS AND TRIGRAMS

Bárbara Emma Sánchez Rinza, María del Rocío Guadalupe Morales Salgado*, Pablo León Morales. Benemérita Universidad Autónoma de Puebla, Universidad Popular Autónoma del Estado de Puebla*.
Facultad de Ciencias de la Computación, 14 Sur and Avenida San Claudio

Abstract*——*This paper describes a cryptosystem for the Spanish language. All languages have some words that are more common than others to make connections between phrases or sentences. In the Spanish language there are several types of words like prepositions, articles that are words of one, two or three frequently used letters. The decryption by syllabic frequencies is an algorithm based on Spanish grammar rules, which is done by a statistical study of frequencies of words of one, two and three most common letters. The method comprises: selecting a theme from a text of about 10,000 words, and calculate the frequencies of these words of one, two and three most used letters in the Spanish language, we will name these types of words unigrams, bigrams and trigrams and comparing with the ciphered text that has to have the same subject. For the process of encryption we will use the Vigenère algorithm to encode the encrypted text previously and subsequently it will be decoded using this technique [1].

I. INTRODUCTION

The word cryptography is a term that describes all techniques to encrypt messages or make them intelligible without resorting to a specific action.

Cryptography is based on arithmetic's: In the case of a text, consists in transforming the letters that make up the message into a series of numbers (in the form of bits since computers use the binary system) and then perform calculations with these numbers in order to:

- To modify them and make them incomprehensible, the result of this modification (the encrypted message) is called ciphered text as opposed to the initial message, called clear text.
- To make sure the recipient can decrypt them, the action of encoding a message to make it secret is called encryption and the inverse method, which is to recover the original message, is called decryption.

Cryptanalysis involves the reconstruction of an encrypted clear text message using mathematical methods. Therefore, all cryptosystems must be resistant to cryptanalysis methods. When a cryptanalysis method allows to decrypt an encrypted by using a cryptosystem message, we say that the encryption algorithm has been decoded.

II. REALIZATION OF THE CRYPTOSYSTEM

The algorithm used consisted, first in take a training text of more than 10 000 words of a specific topic and frequencies of unigrams, bigrams and trigrams was obtained (it should be emphasized that this text is not encrypted). Subsequently the same subject is encrypted by the Vigenère algorithm. Similarly the frequencies of the ciphered text were obtained. After having the two tables of frequency of the training text and the ciphered text, we will replace first by monosyllables, two-syllables and three-syllables. Subsequently a word processor yielding the percentage relationship of letters and words between the original text and decryption is used.

To implement the decryption of text by analyzing the frequencies of mono, bi- and tri-syllabic words, the process was to analyze the input ciphered text to determine the frequency of each of the letters, the frequency of the words of one-letter (unigram), two-letter words (bigram) and three-letter words (trigrams). From the training text. Once we have these frequencies, we know the most common elements in the ciphered text. Finally we proceed to replace these items with the most common in the Spanish language. The Java programming language was used for coding the algorithm. Diagrams used are shown.

Use Case Diagram

In Figure 1 it is shown a use case diagram in which the user interacts with the system. Below is a description of each use case is:

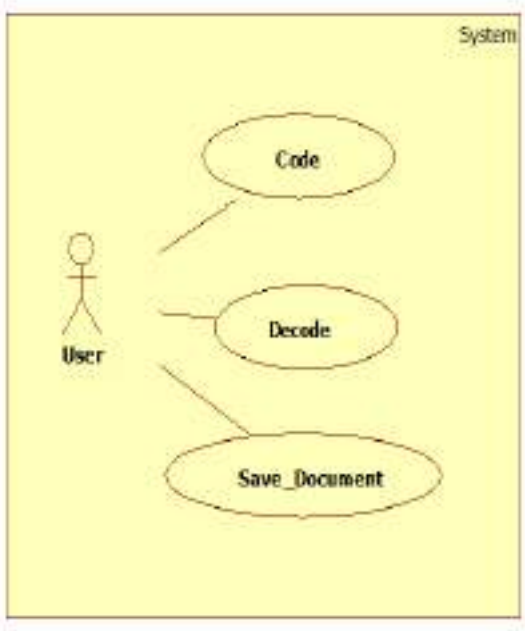

Fig 1. Use Case. Source: Own elaboration.

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 5, Issue 1, July 2015

Security System for Sending Information Containing Hidden Voice Data by Steganography (SIOVE) Using Matlab

Bárbara Emma Sánchez Rinza, María del Rocío Guadalupe Morales Salgado*, Cristian Omar Cortez Olguín
Faculty of Computer Science, Benemérita Universidad Autónoma de Puebla, Universidad Popular Autónoma del Estado de Puebla Puebla, Puebla

Abstract*—with the modern use of technology and communication, data and information travel through many channels. During transmission they may be vulnerable to passive or active attacks, in which messages are used by criminal groups to undermine the integrity of individuals and institutions. In this study, a solution is proposed that uses a system constructed in MATLAB that is able to hide voice signals in an image. That is, the data sent is an image but it carries a protected voice message within it. This solution seeks to ensure the integrity of the data. To contextualize this work, the concepts of steganography and voice signals are defined. The implementation of the SIOVE system and its application are also presented

I. INTRODUCTION

Computer security consists of ensuring that an organization's computer resources are being utilized in the manner determined by the organization, and that data access and modification can only be carried out by authorized people and within the limits of their authorization.[1,2]
The main objectives of computer security are [3]:

- Detect potential problems and security threats, minimizing and managing risks.
- Ensure proper use of resources and application of systems.
- Limit losses and implement system recovery in the event of a security incident.
- Comply with the legal framework and with overall organizational requirements

A. Steganography

The word 'steganography' comes from the Greek words steganos (hidden) and graphos (writing), and can be defined as a technique for hiding information in a covert channel in order to prevent a hidden message from being detected. When a message is transmitted, someone spying on the communication transmission may not be able to decipher it, but they will at least know that an encrypted message has been sent, when it was sent, and how much encrypted data was exchanged. When this knowledge constitutes a threat to the organization, we can employ steganography [4]. As explained above, steganography consists of hiding a secret message inside another message that is not secret. The very existence of the secret message is concealed. If the steganographic message is also encrypted, we can keep its content secret even if its existence is detected upon transmission. It is easier to conceal a message in information that can be expressed with a variable amount of data, such as a photograph, audio or video. An example of steganography is hiding a message inside another message consisting of a hidden "signature" in a set of data. When an unauthorized copy is made of the data, the source can be detected by comparing characteristics such as size and existence of associated codes using the LSB (least significant bit) technique. No "watermark" algorithm has yet been found that is resistant to every possible data manipulation, such as the introduction of noise, changing image resolution, overwriting the signature, or other techniques. The elements or actors in steganography are [4, 5, 6]:

- Container: The object that is used to carry the hidden message.
- Stego-object: The container plus the concealed message.
- Adversary: Any of the entities from which the concealed information is being hidden.
- Steganalysis: The science of detecting (passive attacks) and/or rendering harmless (active attacks) information hidden in some sort of container, and of finding useful information within the container (existence and size).

In general terms, steganography is divided into two types [7, 8,9]:

- Linguistic steganography.
- Technical steganography.

Linguistic steganography uses a written text as the carrier, while technical steganography uses any other type of carrier, which may be audio, images, video, or other data. This paper deals with technical steganography [10].

B. Voice Signals

The voice is a signal that transmits conscious, intelligent information produced by humans in such a manner that listeners can obtain information directly without the need for any further source of information such as images or text.

For voice signals, the most efficient encoders use a speech production model consisting first of an excitation that models air flowing from the lungs and vibration of the vocal cords and secondly a filter that represents the oral and nasal cavity. Both voice and audio signals use a model of an auditory perception system that indicates which components of the signal do not have to be kept because they are not heard, as they are masked in both time and frequency by neighboring higher-energy components of the signal [12]. A spectral representation serves as a simple way to represent relevant

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 5, Issue 6, December 2015

DECRYPTION SYSTEM OF THEMATIC TEXTS IN SPANISH USING FREQUENCY ANALYSIS INCLUDING UNIGRAMS, BIGRAMS AND TRIGRAMS

Bárbara Emma Sánchez Rinza, María del Rocío Guadalupe Morales Salgado*, Pablo León Morales. Benemérita Universidad Autónoma de Puebla, Universidad Popular Autónoma del Estado de Puebla*.
Facultad de Ciencias de la Computación, 14 Sur and Avenida San Claudio

Abstract*——*This paper describes a cryptosystem for the Spanish language. All languages have some words that are more common than others to make connections between phrases or sentences. In the Spanish language there are several types of words like prepositions, articles that are words of one, two or three frequently used letters. The decryption by syllabic frequencies is an algorithm based on Spanish grammar rules, which is done by a statistical study of frequencies of words of one, two and three most common letters. The method comprises: selecting a theme from a text of about 10,000 words, and calculate the frequencies of these words of one, two and three most used letters in the Spanish language, we will name these types of words unigrams, bigrams and trigrams and comparing with the ciphered text that has to have the same subject. For the process of encryption we will use the Vigenère algorithm to encode the encrypted text previously and subsequently it will be decoded using this technique [1].

I. INTRODUCTION

The word cryptography is a term that describes all techniques to encrypt messages or make them intelligible without resorting to a specific action.

Cryptography is based on arithmetic's: In the case of a text, consists in transforming the letters that make up the message into a series of numbers (in the form of bits since computers use the binary system) and then perform calculations with these numbers in order to:

- To modify them and make them incomprehensible, the result of this modification (the encrypted message) is called ciphered text as opposed to the initial message, called clear text.
- To make sure the recipient can decrypt them, the action of encoding a message to make it secret is called encryption and the inverse method, which is to recover the original message, is called decryption.

Cryptanalysis involves the reconstruction of an encrypted clear text message using mathematical methods. Therefore, all cryptosystems must be resistant to cryptanalysis methods. When a cryptanalysis method allows to decrypt an encrypted by using a cryptosystem message, we say that the encryption algorithm has been decoded.

II. REALIZATION OF THE CRYPTOSYSTEM

The algorithm used consisted, first in take a training text of more than 10 000 words of a specific topic and frequencies of unigrams, bigrams and trigrams was obtained (it should be emphasized that this text is not encrypted). Subsequently the same subject is encrypted by the Vigenère algorithm. Similarly the frequencies of the ciphered text were obtained. After having the two tables of frequency of the training text and the ciphered text, we will replace first by monosyllables, two-syllables and three-syllables. Subsequently a word processor yielding the percentage relationship of letters and words between the original text and decryption is used.

To implement the decryption of text by analyzing the frequencies of mono, bi- and tri-syllabic words, the process was to analyze the input ciphered text to determine the frequency of each of the letters, the frequency of the words of one-letter (unigram), two-letter words (bigram) and three-letter words (trigrams). From the training text. Once we have these frequencies, we know the most common elements in the ciphered text. Finally we proceed to replace these items with the most common in the Spanish language. The Java programming language was used for coding the algorithm. Diagrams used are shown.

Use Case Diagram

In Figure 1 it is shown a use case diagram in which the user interacts with the system. Below is a description of each use case is:

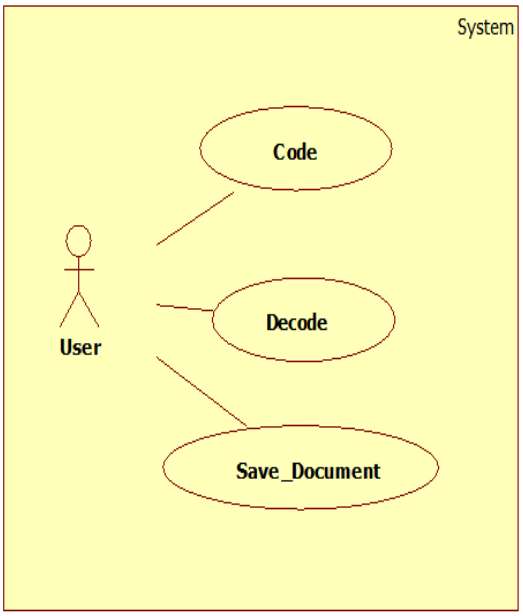

Fig 1. Use Case. Source: Own elaboration.

A Graph Theoretical Approach to Structured Programming based on the Cyclomatic Number

Gonzalo URCID-SERRANO
Optics Department, INAOE
Tonantzintla, Puebla, PUE 72000, Mexico

Rocío MORALES-SALGADO*
Information Technologies Department, UPAEP
Santiago, Puebla, PUE 72160, Mexico

ABSTRACT

In this paper we propose a novel approach for structured programming based on a graph theoretical notion known as the cyclomatic number of an undirected connected graph. Since the cyclomatic number is readily associated with a graphical representation of a program module from its control flow diagram, it provides an intuitive point of view that may complement and enhance common existing techniques currently used in language programming activities that are based mainly on the structured model. Our focus in the present research is three fold, first we introduce the cyclomatic number as a mathematical useful notion for program module description. Secondly, algebraic decomposition of the cyclomatic number as a numerical function is used for program module classification. Finally, illustrative examples are provided to demonstrate the scope and appeal of our proposed graph theoretical framework.

Keywords: Modeling, Graphs, Algorithms, Cyclomatic number, Structured Programming Techniques, Programming Languages, Computer Science Education.

1. INTRODUCTION

Within the vast spectrum of theoretical and practical approaches available for supporting computer programming languages, the structured model still remains a milestone that is a valuable part for the foundation and discussion of new programming paradigms such as the *functional* [7, 13] or *object oriented* [3, 17] models. A key issue in the art of working with and applying the structured model at both the professional and educational levels, relies in how program "structure" can be visualized and measured. In order to achieve simultaneous visualization and measurement of a program module structure in a simple and direct way, we propose to use the cyclomatic number as the basis of an alternative framework for structured programming. In a more restricted sense, our approach can be considered an application of discrete mathematics to structured programming by means of a few key ideas borrowed from graph theory. It is important to remark that the cyclomatic number has been successfully applied in the field of software engineering, where it is known under the name *cyclomatic complexity* [8, 9, 10, 12]; however, to avoid controversial issues belonging to the realm of software metrics we do not use the adjective "complexity" in the present discussion [5].

The work presented here is organized as follows: Section 2, gives background material of structured programming, graphs, and function factorization in support of Section 3, where the cyclomatic number of a graph is defined, function factorization is applied to decompose it when viewed as a numerical mapping, and the final subsection establishes a representative set of prototype program modules based on this decomposition. Section 4, applies our specific proposal of using the cyclomatic number as an alternative presentation for the structured model in language programming activities; simple educational examples are provided to illustrate our approach. Finally, Section 5 states the conclusions as well as future work related to the research presented here.

2. CONCEPTUAL BACKGROUND

The material described in this section provides the necessary ideas used to develop, discuss, and apply an alternative framework for structured programming based on the concept of cyclomatic number, a graph theoretical notion treated in more detail in the next section. The presentation is divided in three subsections, namely, structured programming [12, 17], directed and undirected graphs [6, 14], and function factorization or decomposition [1].

Structured Programming

There are plenty of different programming styles from which a well trained computer scientist or engineer may select or combine them to solve real world computational problems. From a practical and educational point of view, the structured programming style, will remain a fundamental and solid background for professional programmers as well as undergraduate students that can help them to provide automated solutions to

ACERCAMIENTO A LOS USUARIOS DE SECOND LIFE DE HABLA HISPANA A TRAVÉS DE UN ESTUDIO PSICOGRÁFICO DE TIPO VALS

Lorena Mariano Gutiérrez, Universidad Popular Autónoma del Estado de Puebla
Héctor Hugo Pérez Villarreal, Universidad Popular Autónoma del Estado de Puebla
Judith Cavazos Arroyo, Universidad Popular Autónoma del Estado de Puebla
María del Rocío Guadalupe Morales Salgado, Universidad Popular Autónoma del Estado de Puebla

RESUMEN

En la actualidad es importante conocer el perfil de los consumidores de las plataformas que se están utilizando. Second Life es una plataforma virtual que ofrece un mercado global y emergente que ha sido aprovechado por pocas empresas por falta de información que hay acerca de los usuarios activos. Por ello, esta investigación buscó analizar a los usuarios de esta plataforma, mediante una investigación exploratoria de tipo VALS. Los resultados mostraron que los principales usuarios hispanos son de perfiles innovadores y experimentadores. Motivados por la autoexpresión y la búsqueda de nuevos canales

PALABRAS CLAVE: Segmentación, Second Life, VALS

APPROACH TO SECOND LIFE USERS OF HISPANIC SPEECH THROUGH A VALS-TYPE PSYCHOGRAPHIC STUDY

ABSTRACT

At present it is important to know the profile of the consumers of the platforms that are being used. Second Life is a virtual platform that offers a global and emerging market that has been exploited by few companies due to lack of information about active users. Therefore, this research sought to analyze the users of this platform, through an exploratory research of type VALS. The results showed that the main hispanic users are of innovative and experimenters profiles. Motivated by self-expression and the search for new channels.

KEY WORDS: Segmentation, Second Life, VALS

JEL: M30, M31, M39

INTRODUCCIÓN

En las últimas décadas se desarrollaron algunas plataformas apoyadas en riqueza de comunicación, colaboración virtual y creación de contenido 3-D con alto potencial para socialización, entretenimiento, educación e interacción de las marcas con sus consumidores (Zhu, Wang y Jia, 2007). A partir de ello, emergieron con mayor fuerza, la educación virtual y el marketing interactivo digital, aprovechándose nuevos canales que atrajeron un gran número de usuarios por su valor agregado (Barnes y Mattsson, 2011). Pese a su rápido crecimiento, varias de estas plataformas sociales necesitan renovar y adecuar sus modelos de negocio, a fin de continuar explotando al máximo su potencial (Castelló, 2010) y posibilitar un mayor alcance a consumidores de múltiples contextos, entre ellos los hispanoparlantes, interesados en lo que estas plataformas pueden ofrecer. El concepto de social media es el top de la agenda de muchos ejecutivos de empresas hoy en dia. Las decisiones del mercado son consultadas para identificar las formas y las

Control difuso para un convertidor CD-CD bidireccional de medio puente

Fuzzy control for a half bridge bidirectional dc-dc converter

Controle difuso para um conversor CD-CD bidirectional de media ponte

Para citar este artículo / To reference this article / Para citar este artigo: Moreno Aguilar, L. M., Peralta Sánchez, É. y Morales Salgado, M. (2016). Control difuso para un convertidor CD-CD bidireccional de medio puente. *Ingenio Magno*, 7(1), 116-132.

Luz María Moreno-Aguilar
Universidad Popular Autónoma del Estado de Puebla.
Doctorado en Ingeniería de Software
luzmaria.moreno@upaep.edu.mx

Édgar Peralta-Sánchez
Universidad Popular Autónoma del Estado de Puebla.
Laboratorio de Conversión de Energía
edgar.peralta@upaep.mx

María del Rocío Morales-Salgado
Universidad Popular Autónoma del Estado de Puebla.
Posgrados en Tecnologías de Información
e Ingeniería de Software
mariadelrocio.morales@upaep.mx

Fecha de recepción 4 de mayo de 2016
fecha de aprobación 2 de junio de 2016

Resumen

En este artículo se presenta el diseño, la simulación y la implementación de un control aplicado a un convertidor de corriente directa a corriente directa (CD-CD), bidireccional, de medio puente. La estrategia de control está basada en lógica difusa. El control modifica el valor del ciclo de trabajo del interruptor controlado mediante modulación por ancho de pulso, para asegurar un valor específico en el voltaje de salida del convertidor. El control fue validado experimentalmente en un convertidor de potencia operando tanto para funcionamiento en modo reductor como para el funcionamiento en modo elevador. La simulación se realizó en Matlab/Simulink. Un banco de pruebas fue construido con base en un convertidor CD-CD de 2 kW y el algoritmo de control fue implementado por medio de Labview y CompactRIO.

Palabras clave: algoritmo de control inteligente, control difuso, convertidor CD-CD.

Abstract

This article presents the design, simulation, and implementation of a control applied to a half bridge bidirectional direct current to direct current converter. The control strategy is based on fuzzy logic. The control modifies the values of the interrupter's work cycle controlled by pulse-width modulation in order to assure a specific value of voltage leaving the converter. The control was experimentally validated in a power converter operating both for function in step-down mode and for function in step-up mode. The simulation was carried out in Matlab/Simulink. A test bench was built based on a 2 kW DC-DC converter and the control algorithm was implemented with Labview and compactRIO.

Keywords: Intelligent control algorithm, fuzzy control, DC-DC converter.

Resumo

Neste artigo se apresenta o projeto, simulação e implantação de um controle aplicado a um conversor de corrente contínua-corrente contínua, bidirecional, de media ponte. A estrategia de controle esta baseada na lógica difusa. O controle modifica o valor do ciclo de trabalho do interruptor comandado mediante modulação por largura de pulso para garantir um valor específico de saída de tensão do conversor. O controle foi validado experimentalmente em um conversor de potência operando como agente abaixador e elevador. A simulação realizou-se em Matlab/Simulink. Um banco de ensaio foi construído baseado em um conversor CD-CD de 2kW e o algoritmo de controle foi implantado mediante Labview e compactRIO.

Palavras Chave: algoritmo de controle inteligente, controle difuso, conversor cd-cd.