Republic of Iraq
Ministry of Higher Education
and Scientific Research
University of Diyala
College of Science
Computer Department

Lightweight Salsa Cryptography Algorithm for Data Security and Authentication

A Thesis
Submitted to the Computer Science Department \College of Science \University of Diyala
In a Partial Fulfillment of the requirements for the Degree of Master of Science in Computer

By

Samah Jalil Saba

Supervised By

Assistant Prof. Dr. Bashar Talib Hameed

2021 A.D 1443 A.H

بسم الله الرحمن الرحيم

وَلاَ يحِيطُون بِشَيءٍ مَن عِلمِهِ إلاَ بِمَا شاء وسِعَ
كُرسِيُهُ اُلسَمَاوات وَالأرضَ وَلا يَؤُدُهُ حِفّظُهُمَا
وَهُو العَليّ العَظِيمُ

صدق الله العَليّ العظيم

سورة البقرة: من الاية ٢٥٥

Acknowledgment

Firstly, all my prayers go to (Allah), the Almighty, for the successive blessings, divine providence, and my success in this research.

I would like to express my thanks and gratitude to my supervisor Assist Prof. Dr. Bashar Talib Hamid for supervising this research and for the bounty, patience, and continued guidance throughout the work.

My thanks to all academics and administrative staff at the Department of computer science.

Last and not least, thanks a lot go to my family, my friends, and anyone who helped me in one way or another.

SAMAH

Dedication

I would like to dedicate this
Work To:
The soul of both my father and my brother,
and also dedicate to my mother, my son,
and the rest of my family

SAMAH

Linguistic Certification

This is to certify that this thesis entitled ***"Lightweight Salsa Cryptography Algorithm for Data Security and Authentication "*** was prepared by ***"Samah Jalil Saba"*** at the University of Diyala/ Department of Computer Science, is reviewed linguistically. Its language was amended to meet the style of the English language.

Signature:

Name :

Date : / / 2021

Scientific Amendment

I certify that the thesis entitled “***Lightweight Salsa Cryptography Algorithm for Data Security and Authentication***” was prepared by "***Samah Jalil Saba***" has been evaluated scientifically; therefore, it is suitable for debate by the examining committee.

Signature:

Name :

Date : / / 2021

Supervisor's Certification

I certify that this thesis entitled "***Lightweight Salsa Cryptography Algorithm for Data Security and Authentication***" was prepared by **"*Samah Jalil Saba*"** Under my supervisions at the University of Diyala, Faculty of Science Department of Computer Science, as partial fulfillment of the requirement needed to award the degree of Master of Science in Computer Science.

Signature:

Name: Assistant Prof. Dr. Bashar Talib Hamid

Date: / /2021

Approved by the University of Diyala Faculty of Science Department of Computer Science.

Signature:

Name: Assistant Prof. Dr. Bashar Talib Hamid

Date: /6 /2021

(Head of Computer Science Department)

Examination Committee Certification

We certify that we have read the thesis entitled “***Lightweight Salsa Cryptography Algorithm for Data Security and Authentication***” and an examination committee, examined the student “***Samah Jalil Saba***” in the thesis content and that in our opinion, it is adequate as fulfill the requirement for the Degree of Master in Computer Science Department, University of Diyala.

(Chairman)
Signature:
Name**: Prof. Naji Mutar Sahib**
Date: **/ / 2021**

Signature:
Name: **Assist. Prof. Dr. Ali Muhsin Muhammad** **(Member)**
Date: **/ / 2021**

Signature:
Name: **Assist. Prof. Dr. Adil Ibrahim Khalil** **(Member)**
Date: **/ / 2021**

Signature:
Name: **Asst. Prof. Dr. Bashar Talib Hamid** **(Supervisor)**
Date: **/ / 2021**

Approved by the **Dean** of College of Science, University of Diyala

(The Dean)
Signature:
Name: **Prof. Dr. Tahseen H. Mubarak**

Date: **/ / 2021**

Abstract

Confidentiality and authenticity of data are two primary and critical information security services, and they are generally interrelated and share similar objectives for the protection of privacy, and reliability of accessing information. In order to protect data against unauthorized or unintentional disclosure, cryptography is used during data exchange and when data is stored.

Different ciphering methods have been introduced as a solution to provide confidentiality and to play a greater role in information security systems. Salsa20 is one of the popular Stream cipher lightweight algorithms that provides high security and is effective in many modern applications.

A Hybrid Lightweight Salsa20 and Twofish Algorithm (HLSTA) and A Hybrid Lightweight Salsa20 and Present Algorithm (HLSPA) are proposed in this thesis to take advantage of the powerful properties of a block cipher to make Salsa20 more robust against potential attack types with same time preserving the original Salsa20 structure to increase the security of the algorithm.

In addition, this work introduced two Lightweight Authentication Algorithms by add Hashing layer to HLSTA and HLSPA. The aim of the Authentication Hybrid Lightweight Salsa20 and Twofish Algorithm (AHLSTA) and A Hybrid Lightweight Salsa20 and Present Algorithm (Authentication HLSPA) is providing more efficient and secure data authentication.

Experiments and tests for both proposed HLSTA and HLSPA algorithms were applied on two types of data (text and image) with different size, overall performance evaluation for both proposed algorithms based on results of encryption, encryption image histogram, NIST test, and average secrecy are proved both proposed algorithms are more secure than the original Salsa20 algorithm. Also, both proposed HLSPA and HLSTA have highest cipher time more than the original Slasa20 for both text and image .This mean good results for both proposed HLSPA and HLSTA, because it’s increasing the complexity and make it more resistant to attackers.

The proposed HLSPA has better performance on most error sensitivity measures compared to the proposed HLSTA for cipher image, the best average value of the Mean Square Error (MSE) was **11834.21**, the Peak Signal to Noise Ratio (PSNR) was **7.782304**, the Mean Absolute Error (MAE) was **106.2743**, and finally, the Mean Signal to Noise Ratio (MSNR) was **1.433523** with HLSPA**,** the Shannon Entropy **(**SE) was **5.38973,** and the Correlation Coefficient (CC) was **0.008124** with HLSTA**.**

Also The proposed HLSPA has better performance on most error sensitivity measures compared to the proposed HLSTA for cipher text, the best average values of the Mean Square Error (MSE) was **2073.623**, MAE was **39.831**,SE was **2.750865** with HLSPA, CC was **0.013015**, and finally best average value of the CS = **0.943269** with HLSPA.

Both proposed HLSTA and HLSPA have better performance on average secrecy measure compared to the original Salsa20, the best average values of the average secrecy was **0.522612** with HLSTA and **0.517971** with HLSPA.

Experiments and tests for both proposed AHLSTA and AHLSPA algorithms were applied on biometric fingerprint image. The performance evaluation of the AHLSTA and AHLSPA compared with original SHA256 based on results of error sensitivity measurements proved both proposed authentication algorithms have ability to generation a unique and random SHA256.

List of Contents

Cont.

Cont.

List of Figures

Cont.

List of Tables

Cont.

Abbreviations

AES	Advanced Encryption Standard
AHLSPA	Authentication Hybrid Lightweight Salsa20 and present Algorithm
AHLSTA	Authentication Hybrid Lightweight Salsa20 and Twofish Algorithm
AS	Average Secrecy
BER	Bit Error Rate
CC	Correlation Coefficient
CIA	Confidentiality- Integrity – Availability
CS	Cosine Similarity
DES	Data Encryption Standard
ECC	Elliptic Curve Cryptography
ECDH	Elliptic Curve Diffie–Hellman
ECRYPT	European Network of Excellence for Cryptology
ESTREAM	ECRYPT STREAM
GBPA	Generador de Bits Pseudo Aleatorios
GF	Galois Field
GOST	Government Standard
HDM	Hamming Distance Measure
HLSPA	Hybrid Lightweight Salsa20 and Present Algorithm
HLSTA	Hybrid Lightweight Salsa20 and Twofish Algorithm
LTE	Long term Evolution
LTE-A	LTE-Advanced
Lu	Lu chaotic system
LWC	Light Weight Cryptography
MAE	Mean Absolute Error
MARS	Mutually Agreed Resignation Scheme
MDS	Maximum Distance Separation
MSE	Mean Square Error
MSNR	Mean Signal -to-Noise Ratio
NIST	National Institute of Standards and Technology

Cont.

PHT	Pseudo Hadamard Transform
P-Layer	Permutation layer
PSNR	Peak Signal to Noise Ratio
RC6	Rivest Cipher
RFID	Radio Frequency Identification
RS	Reed Solomon code
S-boxes	Substitution boxes
SE	Shannon Entropy Analysis
SHA256	Secure Hash Algorithms 256
SHA-3	Secure Hash Algorithm 3
SM	Smart Meter

List of Algorithms

Chapter One

Introduction

1.1 Overview

The security of the computer is the protection of the system and the data from (modification, unauthorized access, and deletion) attacks. Computer security should provide confidentiality, integrity, and availability. Data protection refers to the methodologies that are planned and enforced to protect sensitive information or data from disturbance, modification, degradation, disclosure, misuse, usage, and unauthorized access. Information security is a collection of practices intended to keep data safe from unauthorized access or alterations [1].

Cryptography is used for the information safety of digital reality today. It means hidden secrets are concerned with encryption. Cryptography is the science of hiding information. More broadly,, it is about designing and analyzing protocols that obstruct the enemy [2]. Modern cryptography concerns with confidentiality (information cannot be understood), integrity (the information cannot be altered), non-repudiation (sender cannot deny transmission of the information), and at a later stage authentication (sender and receiver can confirm each cryptography is used in many applications) [3].

The Cryptography system can be divided into two parts: first, symmetric-key cryptography and the second is public-key cryptography. The algorithms used for symmetric key cryptography are named symmetric key algorithms. Two kinds of symmetrical algorithms are available: stream cipher and block cipher. Stream ciphers encrypt data one by one but block ciphers by dividing information into blocks [4].

Algorithms in cryptography take an important role in providing data security against malignant attacks. The competence of the cryptographic algorithm does not only rely on its time taken for encryption and decryption, and it also accounts for the number of stages utilized to obtain the cipher-text from an original text [5].

Most of the recent encryption algorithms encrypt and decrypt text data that has a small size that does not meet the real requirements of image data. Therefore, it is paradoxical to use the traditional method of text encryption that has large computational requirements [6]. Thus, recent studies and researches aim to minimize these requirements to secure digital images in multimedia distribution. Many studies have proposed different image encryption methods to overcome the encryption problems of digital images [7].

The stream cipher was implemented to obtain the high security of different types of data. The salsa20, as a stream cipher, has an efficient structure that might present the best encryption method of digital images because it requires reasonable hardware according to its simple structure, it means Lightweight hardware encryption addresses increased performance criteria such as power consumption and the area size of the devices. The researchers have applied a series of tests to enhance and justify the efficiency of the salsa20 cipher in visual encryption applications [8]. On the other side, many types of research have paid attention to the weak points of the salsa20 cipher by proposing attack algorithms [9].

In this thesis, proposed new two versions of the salsa20 lightweight algorithm based on hybridized the mechanisms of the original salsa20 stream cipher and Twofish block cipher to a proposed first version called Hybrid Lightweight Salsa20 and Twofish Algorithm which refer a shortly by HLSTA. The proposed second version called Hybrid Lightweight

Salsa20 and Present block cipher Algorithm which refer a shortly by HLSPA. The proposed of HLSTA and HLSPA take advantage of the properties block cipher for both algorithms Twofish and present and preserving the original Salsa20 structure to increase the security of the algorithm while at the same time maintaining the speed of the algorithm within the acceptable range to achieve standards for light encryption algorithms, which are high security and can be used in devices with specific sources.

1.2 Related Works

Many researchers and mathematicians have suggested many works about the Salsa20 algorithm. The following are some studies and discussions that can so far associate works to the suggested work in this thesis:

- **M. S. Mahdi & N. F. Hassan (2018)** [10]: Proposed super Salsa keystream generator with a robust structure, using an array of (4,b) size instead of an array of (4,4), b was salsa volume (8, 12, 20). The diffusion of creating the mainstream increased because the impact of changing length, operating on the one element of the array (4, b) in each iteration, was presented with a balance between the complexity and speed as well. This has resulted in growing complexity and hesitation of linear and differential attacks. It's needed for 2^512 Probable keys to breaking super salsa. The super salsa Keystream's randomness has succeeded in overcoming the benchmark tests of five (frequency, serial, poker, runs, and autocorrelation).
- **L. Evangelina et al. in (2018)** [11]**:** Presented the stream cipher Generador de Bits Pseudo Aleatorios (GBPA) designed for the Internet of Things (IoT). The algorithm is based on the Salsa20

cipher. GBPA uses the core function of Salsa20 with less input parameters, less storage, and less processor time. The GBPA cipher is small in performance and is suitable for IoT devices. GBPA has led to lower use of program and data memory in comparison to computational requirements of GBPA with a lightweight cipher. Three statistical test suites were used to assess the randomness of the performance of the cipher: EACirc, DIEHARD, and NIST statistical test suites with excellent results. The results of both algorithms were not significantly distinguished and were also used as references for Salsa20.

- **H. K. Hoomod & A. M. Hussein (2019)** [12]: Presented a hybrid encryption approach based on the Twofish block cipher and the Salsa stream cipher to give proper security, by combining processes that operate on the salsa20 algorithm with techniques in the S-boxs of the Twofish algorithm, and replacing the key-schedule in S-boxes with the equation in the lu system, which will make breaking the code produced by this method more difficult. The findings suggest that the algorithm modified has less time before it is modified than the original algorithm. As a consequence, the modified algorithm can be used for high-speed encoding and decryption in cascading networks and others.
- **Z.M.J.Kubba & H.K. Hoomod (2019)** [13]: Proposed a hybrid algorithm based on two cryptography algorithms Present and Salsa20. In addition, a chaotic system 2D logistic map was used to produce pseudo-random keys which make the proposed chiper algorithm more complex. This algorithm was meant to provide a hybrid algorithm by enhancing the original present algorithm's complexity and performance. With fast running time, the proposed

algorithm worked efficiently. The time for the suggested algorithm was 8.45 milliseconds, while 10.13 milliseconds were taken the same-size of data using present algorithm, and the analyzed result of the generated sequence keys passed the randomness of the NIST suite.

- **S. M. Salim Reza et al. (2019)** [14]**:** Suggested used a simple lightweight stream cipher algorithm is salsa20 to be used in Smart Meter (SM) for securing the power grid. Along with Salsa20, also proposed Elliptic Curve Cryptography (ECC) based authentication before exchanging any data. This work addressed the problem that the SM is a resource-constrained electronic device that requires a lightweight security mechanism for securing the network but most of the proposed protocols or methods can not satisfy the requirements needed to be lightweight as they consume huge processing power and takes huge processing time to complete its operation. Therefore, will have suggested to used salsa20 stream cipher lightweight algorithm and ECC-based authentication. Numerically analyzed the performance in the case of energy utilization and processing time consumes very little energy and takes very less processing time which makes it suitable to be used in SM.

- **P. K. Pand & S. Chattopadhyay (2019)** [15]: Proposed an improved authentication and security scheme for Long term evolution (LTE) and LTE-Advanced networks (LTE/LTE-A) support highly developed authentication and encryption mechanisms based on a combination of the Elliptic Curve Cryptography(ECC), Elliptic Curve Diffie–Hellman (ECDH), and Salsa20 algorithm to enhance the end-to-end security of 4G environment. In terms of many security characteristics and performance criteria, the performance of the proposed system has been contrasted with LTE-A and current systems. The comparison revealed that the scheme introduced outperforms LTE-A, and other known systems. In comparison with the current safety systems, the proposed system is reliable, stable, and powerful which can deliver reduce computing costs.

- **H. Najem et al. (2020)** [1]: Proposed a GOST block cipher and salsa stream cipher-based hybrid encrypting approach to provide adaptive security, the power consumption is lower, and the encryption speed in the system is more rapid. This proposal was designed to be used in high security required low-cost devices as it is resistant to most of the cryptanalytic attacks common to block ciphers and stream ciphers. The drawback of the GOST algorithm is simple key schedule, so that in certain circumstances be the weak point of the method of cryptanalysis. However, the proposed solution resolves this by forwarding GOST keys to the Salsa stream to ensure correct combination and stronger protection. It requires 2^256 possible keys to crack keys that should not be used as a brute force attack because of their awkward technique in this case. Five standard experiments have also exceeded the randomness of a proposed approach successfully.
- **H. Najm et al. (2021)** [16]**:** presented modification of the Secure Hash Algorithm 3 (SHA-3) with another high-speed algorithm (Salsa20), in the sensor data validation process which creates a high-speed and secure algorithm. The developed logistic method would also produce the Secure Hash Algorithm 3 (SHA-3) algorithm's initial values unknown and not identifiable by the intruder. Correspondingly, the proposed method successfully surpassed

randomness in fifteen statistical tests for National Institute of Standards and Technology (NIST).

1.3 Problem Statement

The rapid development of technology that is used in different applications and thus leads to an increase in information and data such as images and texts. To maintain the security and reliability of the data at a high level of security and at the same time to preserve the sources of devices that use these applications, lightweight encryption is the ideal solution. The salsa20 stream cipher and its simplified models are among the fastest today.

Salsa20/7 is broken by differential attacks however; Salsa20/12 is not as secure as it was previously. Therefore, most researchers and studies attend to increase confusion and diffusion of salsa20 using a different technique such as chaotic maps, but this thesis will be used a block cipher mechanism to make the original Salsa20 more secure.

1.4 Aims of the Thesis

The main goals of the proposed lightweight salsa20 cryptography for data security and authentication are illustrated as follows shown below:

1. Increase diffusion of original salsa20 stream cipher using Twofish block cipher and lightweight present cipher.
2. Create a unique and secure hash of sensitive data based on the proposed lightweight authentication algorithm.
3. Increase security level of original salsa20 lightweight algorithm with an acceptance range of execution time, which is considered one of the most important standards for lightweight encryption algorithms.

1.5 Contribution

The main contribution of this thesis is to increase the security of the original Salsa20 stream cipher lightweight algorithm using the operation of Twofish and Present block cipher algorithms. So in this work, it has been suggested which are Hybrid Lightweight Salsa20 and Twofish Algorithm (HLSTA) and Hybrid Lightweight Salsa20 and Present Algorithm (HLSPA).

In addition, the proposed authentication algorithm based on adding a hash layer for both the proposed algorithm to produced Authentication Hybrid Lightweight Salsa20 and Twofish Algorithm (AHLSTA) and Authentication Hybrid Lightweight Salsa20 and present Algorithm (AHLSPA).

1.6 Thesis Outlines

The rest of the thesis chapters are clarified as follow:

Chapter two: The theoretical background of the work

Chapter Three: The components and architecture of the propose system (HLSTA), (HLSPA), (AHLSTA), and (AHLSPA),

Chapter Four: The results and evaluation of the experimental

Chapter Five: The conclusions and future work