



Naif Arab University for Security Sciences
Journal of Information Security & Cybercrimes Research

مجلة بحوث أمن المعلومات والجرائم السيبرانية
<https://journals.nau.edu.sa/index.php/JISCR>

JISCR

Integrating Light-Weight Cryptography with Diacritics Arabic Text Steganography Improved for Practical Security Applications



CrossMark

Malak G Alkhudaydi, Adnan A. Gutub*

Computer Engineering Department, College of Computer and Information Systems, UMM Al-QURA University, Makkah, Saudi Arabia.

Received 12 Oct. 2020; Accepted 25 Nov. 2020; Available Online 20 Dec. 2020

Abstract

Cryptography and steganography are combined to provide practical data security. This paper proposes integrating light-weight cryptography with improved Arabic text steganography for optimizing security applications. It uses light-weight cryptography to cope with current limited device capabilities, to provide acceptable required security. The work tests hiding encrypted secret information within Arabic stego-cover texts, using all common diacritics found naturally in the Arabic language. The study considers different challenging situations and scenarios in order to evaluate security practicality. It further carries out simulations on some short texts from the Holy Quran, taking them as standard authentic texts, that are fixed and trusted, therefore providing realistic study feedback that is worth monitoring. Our improved approach features preferred capacity and security, surpassing the best previous diacritics stego approach, showing interesting potential results for attractive enlightening exploration to come.

I. INTRODUCTION

The concept of security is very important for many real-world applications. It is currently considered that there is a critical demand for safety and privacy. Everything is becoming digital as a huge vital move for today's technology, which ever raises new vulnerabilities. Most of our personal and identity information has been digitalized, such as passport or identity information, email messages, health information, private family pictures, bank information, and credit card information. It is vital that all this digital information is not dishonestly exposed, i.e. making the need for suitable security more and more urgent.

Security is needed for storing data and for transmitting it as well. In fact, data transfer exposes information

to security breaches more than when they stand still [1]. Cryptography encryption is found to be useful in securing data transfer. Encryption means converting the data to difficult (complex) form to protect data and information from unauthorized benefit. Steganography is another method for hiding the presence of data, regarding storage and securing data-transfer [2]. Steganography is different from cryptography in that cryptography is concerned with removing the power of a hacker or a third party, while steganography is concerned with hiding the data itself from a third party. Cryptography alone mainly serves confidentiality; and steganography normally helps preserve availability [3], we are integrating them to benefit from them both, in addition to adding robustness, raising the system service and providing more security.

Keywords: Information Security, Arabic Text Steganography, Diacritics Steganography, Data encryption, Light-Weight cryptography, Cybercrimes.



Production and hosting by NAUSS



* Corresponding Author: Adnan A. Gutub

Email: aagutub@uqu.edu.sa

doi: [10.26735/FMIT1649](https://doi.org/10.26735/FMIT1649)

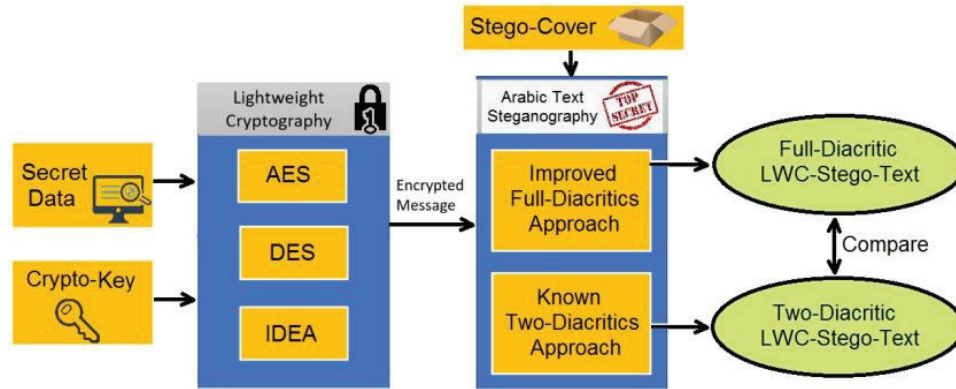


Fig. 1. Overview of the integrated system.

The work further proposed adopting light-weight cryptography to cope with current limited device capabilities, providing acceptable required practical security. Recently, many lightweight cryptography (LWC) research has presented well-known algorithms [4]. Such techniques work on many devices that have insufficient resources, like connectivity hardware and software, adequate, power consumption, memory size and execution time, i.e. targeting pervasive devices with limited resources [5]. The focus of lightweight cryptography (LWC) is to provide practical security and privacy in resource-constrained applications, embedded systems, Internet-of-Things (IoT), and cyber-physical systems, including Radio Frequency Identification (RFID) systems, wireless sensor networks, vehicle ad-hoc networks, and digital healthcare needs [6]. This work involving LWC is studied theoretically, as a separate layer, to ensure its suitable to be combined with steganography, to form effective security protection.

Steganography techniques that have been developed lately are getting more attention than ever before [7]. Steganography is found to be promising and useful in personal processing and the storage of e-data, especially when assuming private individual confidentiality, though it is not the only secrecy solution. However, it can be one of the efficient approaches recommended in today's real-life e-services, as discussed differently within Unicode text steganography [8] and multi-image secret hiding [9], as well as within educational platforms [10] and online account systems [11]. Therefore, this proposed research work can be assumed as establishing an essential research position, to work on upcoming novel schemes, being adjusted and applied as desired, by specific e-applications.

Recall as introduced, the main target of involving

steganography is to hide the existence of information from unintended users, i.e. the information will be hidden in such a way that none can reveal its presence. In this work, we will combine integrated cryptography and steganography techniques to secure sensitive information within Arabic text data. Arabic text crypto-stego schemes ensure that even if the embedded text is discovered, the content remains unknown in meaning or unbeneficial, as the hidden information has been encrypted via LWC [12]. The use of LWC is selected to ensure using the minimum amount of computation resources, helping to maintain practical limited device capabilities, fulfilling the user experience of today. The study contribution tests the integration of different common LWC schemes (AES, DES, IDEA) with interesting variation in Arabic text steganography approaches (two-Diacritic, improved full Diacritics), as shown in the overview of Fig. 1. The Cryptography input is secret data and pre-shared crypto-key; the steganography input is stego-cover and the output of cryptography. The steganography output can be noted as the integrated system results, as marked in Fig. 1. To be concise, this study proposes an Arabic text crypto-stego system where sensitive secrets are encrypted and hidden within Arabic text diacritics, which are tested providing positive results.

The flow of the paper is as follows. Section 2 will cover the related studies of LWC and linked steganography, with emphasis on Arabic text, along with some technical details. Section 3 will describe the research approach of the integrated system, explaining both hiding and retrieving security processes. Section 4 provides the implementation details and the experiment results, presenting the performance and efficiency measurements, followed by the conclusion of the work in Section 5.

II. RELATED BACKGROUND

This section reviews some technical details about Light Weight Cryptography (LWC) and highlights knowledge on Steganography. The section gives focus to Arabic Language texts, which have helpful features used in our integration system. The work performance is tested in a latter section, via security and capacity. The term capacity represents the size of the bits of data that can be outfaced by the stego-cover, as steganography carriers of standard cover media such as pictures [11], video clips [10], music, and sound [13], where our work focuses on diacritic Arabic text steganography. Regarding the security measurements, attackers can handle the detection of hidden information, using its perceptual appearance. In that case, if the gap between stego-text and the cover text is large, higher security levels can be achieved, as will be elaborated on later during the comparison and analysis section.

A. Light-weight cryptography (LWC)

The term cryptography started to be used at the start of the 19th century. It uses encryption to conceal data, and decryption to recover information, as main activities within the field of cryptograph. Encryption changes the plain or ordinary text into complex encoded data called ciphertext. Decryption is the reverse process; which is reverting the ciphertext into plain text. Cipher (or Cypher) are the algorithms used for encryption and decryption. There are two categories for Cryptography: Symmetric and Asymmetric. Regarding the symmetric type, the process of encryption and decryption is handled using one same key. That means that both sender and receiver would have pre-shared their key. For the asymmetric type of cryptography, the process uses two different keys, public and private. The public key is shared while the private key is used by the receiver only. The transferred data is encrypted using the public key and decrypted using its paired private key [1].

Lightweight cryptography (LWC) is added as an enhancement to many famous encryption methods such as AES (Advance Encryption Standard) and DES (Data Encryption Standard) cryptography [3]. Such algorithms are developed mainly for devices with insufficient computational resources, or high-power availability. LWC gathers between lightness and security. The goal of lightweight cryptography (LWC) is to provide security and privacy in resource-constrained applications, embedded systems, Internet-of-Things (IoT), and cyber-physical systems, in-

cluding Radio Frequency Identification (RFID) systems, wireless sensor networks, vehicle ad-hoc networks, and healthcare [4].

Block size is small in lightweight cryptography, usually 64 bits or 80 bits. In addition, the key is set to be small, typically less than 90 bits. Rounds are less complex, and S-boxes often contain 4 bits. Considering hashing methods in LWC, SPONGENT [4], PHOTON [5], and Lesamanta-LW [6] are popular recent examples. Differently, PRESENT [3] and CLEFIA [2] are other scenarios for block-crypto methods other than Enocoro and Trivium representing stream methods [6]. In this research, we benefited from all the previously mentioned attempts, testing theoretically, the three popular LWC algorithms, AES, DES and IDEA, as separate layers to be integrated with steganography, aiming for better performance and higher security. This use of LWC is studied hypothetically, as a separate layer, to ensure its appropriateness when merged with steganography, developing operative security protection.

B. Arabic text Steganography

Steganography is a technique to keep data hidden in storage or during communication. Designing steganography systems can include three criteria: perceptual transparency, robustness, and hiding capacity [9]. The purpose of stego-security is to protect confidential information that is contained in sensitive texts, and ensure it is not retrieved easily [14]. Steganography can hide data within any multimedia digital format [11]. This work targets text steganography, considered as a challenging method for limited unwanted (redundant) information, i.e. in text files more than in other types of cover media [15]. The composition of text files is common and the structure of other types of network digital forms may be quite dissimilar from what the media monitors. The beauty of writing a word document in the connection, is its practical, lowest memory, utilization of sources [8]. Therefore, our choice of text steganography found various steganographic techniques geared towards specific languages, depending on the text linguistic structure [15], motivating our work towards improving new relevant adjustments.

Several studies have lately been devoting steganography to Arabic texts, i.e. to take advantage of adding, editing or changing letters or diacritics and make use of its Arabic linguistic rules [8]. Arabic words are proven to be robust and difficult to discover via steganography [16].



TABLE I
EXAMPLE OF DIFFERENT LWC WITHIN THE INTEGRATED CRYPTO-STEGO SYSTEM.

LWC	AES	IDEA	DES
Original Secret Text	Malak	Malak	Malak
En-crypted message	E£fEج ظة:5ض	Jé,qr}رذوغ\$D	TOT\$PV\$µ^D
Secret bits of en-crypted message	1110111110111011101111111011 00010111000001000001101100010 1010011101100110001111101100 01001101100110101110110001011 01100110110011011000101110100 01000001101100010101110110110 01100011001100001010100100110 11000100011000100010111000010 10100011	1110111110111011101111111000 110000100010000101110111000 10100000010011010110110101 001100011011000101001001101 100010111010001001000111001 001111011111000101000000010 001111000001000100101011000 011101010011101100110001000 01110001	111011111011101110111111 110010111000011000100000 110000101010001000100000 010001000010010011000010 101101010101000001010110 001001001100001110101111 110000101010011011000010 101101100100111101010100
Stego-text as Integrated system output	عَلَى قَدْرٍ أَهْلَ الْعَزْمِ تَأْتِي الْعَرَائِمُ وَتَعْظُمُ فِي عَيْنِ الصَّغِيرِ صَغَارُهَا وَتَأْتِي عَلَى قَدْرِ الْكِرَامِ الْمَكَارِمُ وَتُصْغَرُ فِي عَيْنِ الْعَظِيمِ الْعِظَامُ يُكَلِّفُ سَيْفَ الدَّوْلَةِ الْجَيْشَ هَمَّهُ وَقَدْ عَجَزَتْ عَنْهُ الْجَيْشُ الْخِضَارُ وَيَطْلُبُ عِنْدَ النَّاسِ مَا عِنْدَ نَفْسِهِ وَذَلِكَ مَا لَا تَدْعِيهِ الضَّرَاعِمُ يُفْدِي أُنْمَ الطَّيْرِ عُمْرًا سِلَاحَهُ تُسَوِّرُ الْفِلاَ أَحْدَانُهَا وَالْقِشَاعِمُ وَمَا ضَرَّهَا خَلْقٌ بِغَيْرِ مَخَالِبِ وَقَدْ خَلَقَتْ أَسْيَافَهُ وَالْقَوَائِمُ هَلِ الْحَدِيثُ الْحَمْرَاءُ تَعْرِفُ لَوْنَهَا وَتَعْلَمُ أَيَّ السَّافِيَيْنِ الْغَمَائِمُ سَقَّتْهَا الْغَمَامُ الْغَرَّ قَبْلَ نَزْوَلِهِ فَلَمَّا دَنَا مِنْهَا سَقَّتْهَا الْجَمَاجِمُ بَنَاهَا فَأَعْلَى وَالْقَنَا يَنْزِعُ الْقَنَا وَمَوْجُ الْمَنَابِيا حَوْلَهَا مُتَلَاظِمُ	عَلَى قَدْرٍ أَهْلَ الْعَزْمِ تَأْتِي الْعَرَائِمُ وَتَعْظُمُ فِي عَيْنِ الصَّغِيرِ صَغَارُهَا وَتَأْتِي عَلَى قَدْرِ الْكِرَامِ الْمَكَارِمُ وَتُصْغَرُ فِي عَيْنِ الْعَظِيمِ الْعِظَامُ يُكَلِّفُ سَيْفَ الدَّوْلَةِ الْجَيْشَ هَمَّهُ وَقَدْ عَجَزَتْ عَنْهُ الْجَيْشُ الْخِضَارُ وَيَطْلُبُ عِنْدَ النَّاسِ مَا عِنْدَ نَفْسِهِ وَذَلِكَ مَا لَا تَدْعِيهِ الضَّرَاعِمُ يُفْدِي أُنْمَ الطَّيْرِ عُمْرًا سِلَاحَهُ تُسَوِّرُ الْفِلاَ أَحْدَانُهَا وَالْقِشَاعِمُ وَمَا ضَرَّهَا خَلْقٌ بِغَيْرِ مَخَالِبِ وَقَدْ خَلَقَتْ أَسْيَافَهُ وَالْقَوَائِمُ هَلِ الْحَدِيثُ الْحَمْرَاءُ تَعْرِفُ لَوْنَهَا وَتَعْلَمُ أَيَّ السَّافِيَيْنِ الْغَمَائِمُ سَقَّتْهَا الْغَمَامُ الْغَرَّ قَبْلَ نَزْوَلِهِ فَلَمَّا دَنَا مِنْهَا سَقَّتْهَا الْجَمَاجِمُ بَنَاهَا فَأَعْلَى وَالْقَنَا يَنْزِعُ الْقَنَا وَمَوْجُ الْمَنَابِيا حَوْلَهَا مُتَلَاظِمُ	عَلَى قَدْرٍ أَهْلَ الْعَزْمِ تَأْتِي الْعَرَائِمُ وَتَعْظُمُ فِي عَيْنِ الصَّغِيرِ صَغَارُهَا وَتَأْتِي عَلَى قَدْرِ الْكِرَامِ الْمَكَارِمُ وَتُصْغَرُ فِي عَيْنِ الْعَظِيمِ الْعِظَامُ يُكَلِّفُ سَيْفَ الدَّوْلَةِ الْجَيْشَ هَمَّهُ وَقَدْ عَجَزَتْ عَنْهُ الْجَيْشُ الْخِضَارُ وَيَطْلُبُ عِنْدَ النَّاسِ مَا عِنْدَ نَفْسِهِ وَذَلِكَ مَا لَا تَدْعِيهِ الضَّرَاعِمُ يُفْدِي أُنْمَ الطَّيْرِ عُمْرًا سِلَاحَهُ تُسَوِّرُ الْفِلاَ أَحْدَانُهَا وَالْقِشَاعِمُ وَمَا ضَرَّهَا خَلْقٌ بِغَيْرِ مَخَالِبِ وَقَدْ خَلَقَتْ أَسْيَافَهُ وَالْقَوَائِمُ هَلِ الْحَدِيثُ الْحَمْرَاءُ تَعْرِفُ لَوْنَهَا وَتَعْلَمُ أَيَّ السَّافِيَيْنِ الْغَمَائِمُ سَقَّتْهَا الْغَمَامُ الْغَرَّ قَبْلَ نَزْوَلِهِ فَلَمَّا دَنَا مِنْهَا سَقَّتْهَا الْجَمَاجِمُ بَنَاهَا فَأَعْلَى وَالْقَنَا يَنْزِعُ الْقَنَا وَمَوْجُ الْمَنَابِيا حَوْلَهَا مُتَلَاظِمُ

- Understanding the system necessities along with its implementation details. Determining the cyp-to-key as an important, well-hidden element, so that no third party can discover it. Similarly, the secret data or messages should be hidden such that none can reveal its presence or their contents.
- In case any hidden information is detected illegally, the message should have been encrypted properly, assuming its key is hard to find out, or used for secret/messages decryption.
- No technical or useful information in recognizing

- the message is permitted to be revealed or detected; and any other ways that disclose the contents of secret messages shouldn't be revealed.
- Fulfill system requirements that may include secret/message requirements, container requirements, stego-container requirements, algorithm requirements and security requirements. In brief, to achieve secret/message transfer without any blocks or issues and most importantly with security goals set.
- Any secret/message that goes under steganogra-



phy should be resilient to the distortion caused by various transformations, such as decrease or increase, conversion to another formats, change in resolutions or addition of objects. This maintains the information robustness.

- The error correction code may be used to maintain the integrity of the embedded information. Some messages can be duplicated to increase reliability.
- It's vital to bear in mind the degree to which the message is secret (imperceptibility), along with the size of the message embedded within it (maximum size and capacity).

The integrated approach proposed in this work used both cryptography and steganography to accrue their benefits, as shown in Fig. 4. The integration is heavily experimented, confirming practical performance and utilization, measuring performance via capacity and security. The proposed study is performed by applying the three LWC schemes on the same involved secret text, as shown by the simplified example, detailed in TABLE I. Then, the LWC selection is performed, based on the resultant efficiency, followed by the conversion to binary for stego imbedding within diacritics. The hiding is performed by preserving (keeping) the diacritics whenever a one is found, removing the zero diacritics, and this is performed continuously on the bit stream until the end of the secret data binary bits. In the following subsections, we will detail the suggested algorithm in some depth, within the architecture components and testing data, in order to verify the work.

A. System Architecture

Consider the encryption via LWC shown in Fig. 4, the system methods used here are AES (Advance Encryption Standard), DES (Data Encryption Standard) and IDEA (International Data Encryption Algorithm). These algorithms are selected for the study, assumed to be practically efficient and suitable for use with limited resources [26]. It's noticeable that the encryption key is private and pre-shared, so the encryption is symmetric, as the key is needed for decryption when the secret data is to be retrieved once more.

The sensitive information is encrypted and sent to be hidden via steganography, following Fig. 5. The hiding process depends mainly on the full Arabic text diacritics (Fathah, Kasrah, Dammah, Fathatan, Dammatan, Kasratan, Sukun and Shaddah); where the results are in

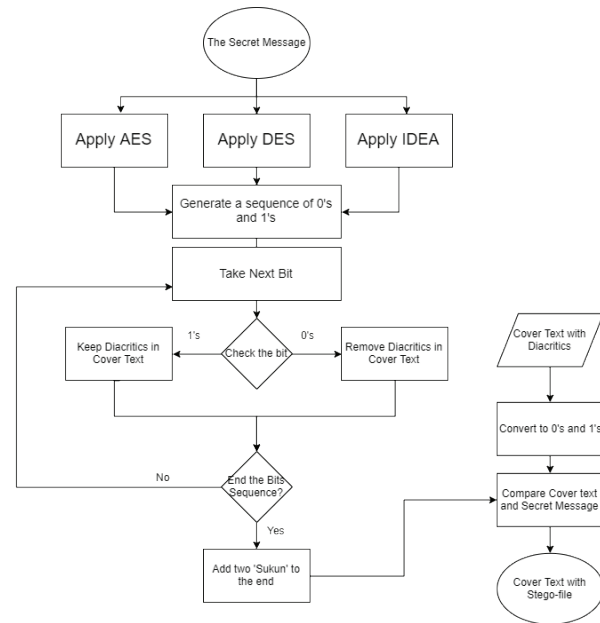


Fig. 4. Block diagram of the integrated security system.

the form of stego-text that is safe for public distribution, assuming its secret cannot be easily detected or even known to be present. This stego hiding within Arabic text is similar in principle to the efficient hiding of [27] but utilizing diacritics instead of Unicode features.

When receiving or retrieving the sensitive information, the stego-text is used to reveal the hidden text and revert to the original secret, but remain encrypted. Then, the crypto-key plays its role in decrypting it, revealing the secret once more. Recall, using the LWC purpose of simplifying the computational, power, capability and memory requirements, to be suitable for devices with limited resources [28]; as the main variance from utilizing complex encryption, such as elliptic curve cryptography, with stego combinations [29]. The use of AES, DES and IDEA are good choices as LWC encryption algorithms for this study, in addition to research pilot stages [28]; though these LWC procedures may not be the optimal solutions for heavy-duty daily applications that can be kept for future studies.

In summary, diacritics is an important feature of Arabic text and it's mainly used here for hiding the messages. The Arabic letters and diacritics are then converted to binary bits and the cypher secret text is stego-stored. The binary code will identify the cover text that will be used in the hiding process. Diacritics, in our work, are so vital in playing an important role in ensuring that the place-

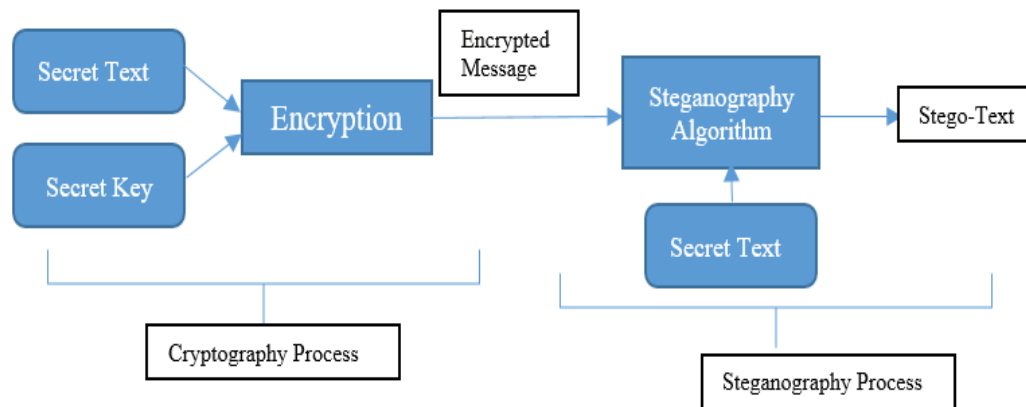


Fig. 5. Storing the sensitive secret data.

ment of the secret bits are set within the usable range of the cover text. In fact, this hiding research needs to be combined to secure access technology [30], in order to provide more effective overall privacy solutions.

B. Experimental Steps and Hiding Process

The system is integrated with its embedding process, as shown in Fig. 5. The software development starts by selecting an input file that contains the information or messages to be securely stored or sent. Then, the LWC encryption algorithm is applied to the message, i.e. using AES or DES and IDEA, to get the encrypted data. The encrypted results are compared, as shown in TABLE I, providing a possibility for preferential choice. TABLE I encrypted data from the different LWC schemes are compared by looking at their size, as well as their secrecy, via automated histogram and PSNR analysis. It is to be made clear that there is no obvious observed relationship between the different encryption cyphertext results and the original plain text. Therefore, automatic analysis is necessary, as will be elaborated on in the next section's comparison.

The encrypted data are then hidden within Arabic text diacritics via the two methods. Our example in TABLE I shows a hiding process sample, using our improved full diacritics stego scheme. The performance of every strategy is tested to sense practical security and capacity calculations, i.e. of bit hiding, as the focus of the research. The output of this process will be a stego-text that will be stored or received through data storage or transfer for reverse processing to follow.

The proposed algorithm starts by converting the cover-text, which is used for the hiding process, into binary bits. The bits are stored in an array, where every 16 bits represent Arabic letters or diacritics, which are placed together visually. 8-bits are added, representing non-Arabic characters, to complete the programming proper investigation. Recall that the secret message is encrypted by the intended LWC algorithms. The algorithms read the first bit of the secret message and then compare it with the first diacritics in the cover-text. If, for instance, the first bit to be hidden is a 'one', this first diacritic, 'Fathah' for example, will remain; otherwise the 'Fathah' will be removed. This process is repeated until all the secret bits in the secret message have been handled. When the secret message ends, two 'Sukun' are added to the cover-text, which is not found in Arabic writing, used as an indication of ending the hiding process. Algorithm 1 outlines the hiding process as a pseudo-code.

It's important to make sure that all possible diacritics are present in the cover-text, in order to hide some data in a cover text, using the proposed technique. Then, the algorithm sequentially matches every diacritic to a bit from the secret bits, as shown by the example in Fig. 7.

It is to be noted that the secret bits should be hidden in the cover text in proper order, i.e. from right-to-left, following the Arabic writing style. To run our experimentations fully, we used the example of a prepared, simplified poem as a suitable cover text, see below:

عَلَى قَدْرٍ أَهْلَ الْعَزَمِ تَأْتِي الْعَزَائِمُ وَتَأْتِي عَلَى قَدْرِ
الْكَرَامِ الْمَكَارِمِ وَتُعْظِمُ فِي عَيْنِ الصَّغِيرِ صَغَارَهَا وَتُصَغِّرُ فِي
عَيْنِ الْعَظِيمِ الْعِظَائِمَ يَكْلِفُ سَيْفُ الدَّوْلَةِ الْجَيْشَ هَمَّهُ وَقَدْ



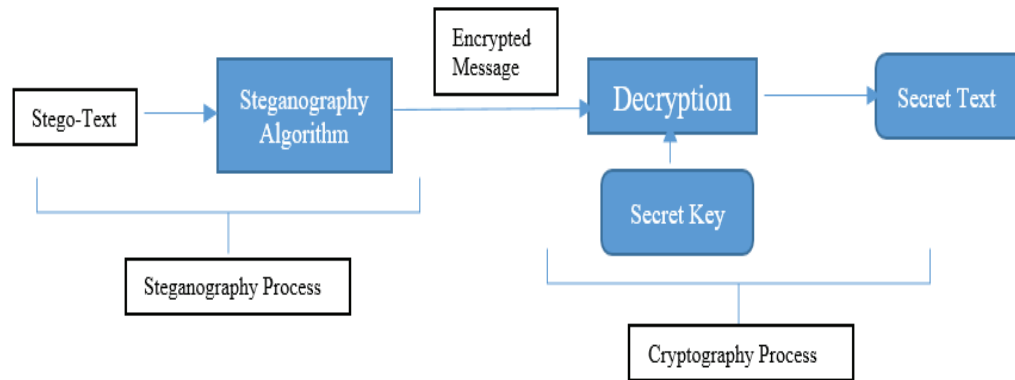


Fig. 6. Retrieving back the secret data.

Algorithm 1: Hiding secret message bits into Arabic cover-text

Inputs: cover-text, input file

Outputs: stego-text

1. Conduct the secret message encryption by different LWC algorithms.
2. Choose the preferred cyphertext from LWC outputs
2. Convert the encrypted message (it is a sequence of binary bits 0's and 1's).
3. Insert the cover-text with Diacritics, and converting it to Binary Bits. n.
4. storing it in an array
5. Check value of bits of secret cyphertext message to be hidden.
 - a. if it is '1' the 'Diacritic' is kept.
 - b. if it is '0' remove 'Diacritic'.
6. Repeat step 5 until the secret message ends.
7. As secret message ended, involve two 'Sukun' diacritics in cover-text indicating ending embedding.
8. Show the text before/after the stego process, i.e. before/after encryption.

عَجَزَتْ عَنْهُ الْجِيُوشُ الْخَضَارِمُ وَيَطْلُبُ عِنْدَ النَّاسِ مَا عِنْدَ
نَفْسِهِ وَذَلِكَ مَا لَا تَدْعِيهِ الضَّرَاغِمُ يُفِدِي أُنْمُ الطَّيْرِ عَمْرًا
سَلَاحَهُ نَسُورُ الْفَلَاحِ أَحْدَانُهَا وَالْقَشَاعِمُ وَمَا ضَرَّهَا خَلْقٌ بَغِيرَ
مَخَالِبٍ وَقَدْ خَلَقَتْ أَسْيَافُهُ وَالقَوَائِمُ هَلِ الْحَدِيثُ الْحَمْرَاءُ
تَعْرِفُ لُونَهَا وَتَعْلَمُ أَيَّ السَّافِيَيْنِ الْفَعَائِمُ سَقَّتْهَا الْفَمَامُ الْفَرُ
قَبْلَ نَزُولِهِ فَلَمَّا دَنَا مِنْهَا سَقَّتْهَا الْجَمَاجِمُ بِنَاهَا فَأَعْلَى وَالقَنَا
يَقْرَعُ القَنَا وَمَوْجُ المَنَايَا حَوْلَهَا مِتْلَاطِمُ

The integration process sequentially matches every diacritic to a bit from the secret bits. Then, we apply the hiding process, removing diacritics related to zero bits, whilst keeping the ones. The algorithm encrypted the secret by the three LWC algorithms represented in the text, as well as binary bits, as detailed in TABLE I. The simple

Cover text: عَلَيَّ قَدْرُ أَهْلِ الْعَزْمِ تَأْتِي الْعَرَائِمُ
Secret bit: 1010011010101010110001111
Stego-text: عَلَيَّ قَدْرُ أَهْلِ الْعَزْمِ تَأْتِي الْعَرَائِمُ

Fig. 7. Example of the proposed hiding stego technique.

secret data used for this clarification has the normal name "Malak".

C. Extracting Process

The extracting process is the reverse of the hiding bit procedure, as shown in Fig. 6, observed by the receiver as retrieving the secret message. It involves the stego-text, which is an output of the hiding process, starting as an input secret message for decryption, using the pre-shared crypto-key. Our proposed software platform is programmed to show the verification data, i.e. by clicking "show data" link, the program will retrieve the secret and sensitive text data. It shows the binary bits of the encrypted hidden text within the diacritics. Then, the software converts the binary bits to its original text. Note that the results from the platform show the encrypted text that needs to be decoded via the intended LWC algorithm, i.e. requiring the secret crypto-key as an input to the reverse cryptography process. The system decrypts the cipher text, regenerating the secret sensitive data, as example shown in Fig. 8.

In general, to simplify extracting the data from the cover text, using the proposed technique, we spot the unavailable diacritics as hiding bits of zeros, while others are simply ones. The retrieving Algorithm 2 shows the extraction process as a pseudo-code.



Stego-text: عَنَ قَدْرَ أَهْلِ الْعَزْمِ تَأْتِي الْعَرَائِمُ
 Cover text: عَنَ قَدْرَ أَهْلِ الْعَزْمِ تَأْتِي الْعَرَائِمُ
 Secret bit: 111100011010101010101100101

Fig. 8. Example of the extracting process.

Algorithm 2: Extract secret message, the original text

Inputs: stego-text

Outputs: secret message

1. Check the existence of the two ‘Sukun’ in the stego-file to mark the ending of hiding secrecy.
 2. Start with the first word in the stego-file which has the binary bits of the encrypted hidden text within the diacritics.
 3. Check the letters of the stego-text
 - a. If it has Diacritic, a value one is recorded.
 - b. If the letter is with no Diacritic, a value zero is recorded.
 4. Repeat step 3 until two ‘Sukun’ is sensed.
 5. Collect the binary recorded bits as cyphertext data.
 6. Decrypt cyphertext using pre-shared key via the known algorithm (AES or DES or IDEA)
 7. Convert binary to text.
 8. Extract the original message.
-

IV. COMPARISONS AND ANALYSIS

The integrated security system for hiding sensitive text data, on personal computers, is implemented on a visual studio-programming platform. The algorithm is programmed by C# language, using UTF-8 Encoding, due to its convenience fully supporting Arabic text. The data used in the experiment are samples from the holy Quran, for process testing analysis. Note that the number of letters in the chosen Quran Surahs are different, leading to different capacity values. Therefore, the data used from the Quran provide the concept linking to the work in [23]. However, the thorough progress of the integrated system comparisons, utilized samples from the fixed size poem. The poem selected was written by AlMotnabi, a famous Arabic Poet. In both cases, the Diacritics involvement showed variation in results, that helped our experiment demonstrate the effects of the hiding process, i.e. testing utilizing Diacritics for secrecy. In other words, the implementation targeted testing the proposed methodology for hiding secret information bits within all diacritics, after conducting the text encryption by different LWC algorithms, as shown via the platform in Fig. 9. Note that the LWC keys are pre-shared, well-thought-out and known by the participants before the process. In other words,

this sharing of LWC keys is part of the common key management challenge considered outside of the scope of our work at this stage.

As previously described, the Arabic language uses Diacritics in the holy Quran, as well as religious and historical scripts. As mentioned, we use study [23] as a comparison, which uses Two Diacritics schemes (‘Fathah’ and ‘Kasrah’). Our experimentation selects the secret bits to be hidden and then stores them in the cover text, using all Diacritic approaches. TABLE II shows the bits used in encryption, as well as their values in 0’s and 1’s. TABLE III demonstrates the difference in using the two approaches, applying historical Arabic poetry, which is full of diacritics, with examples using 9-bits, 20-bits, 30-bits, 50-bits, 60-bits, 100-bits, 120-bits, 240-bits and 480-bits hidden. For 480 secret bits, the secret message can’t be hidden, as the cover text has 296 bits, as it’s not big enough to conceal the secret message within its content, during the hiding process.

A. Capacity Comparison

Initially, the capacity we needed was determined by different sensitive data sizes or secret messages. Then, the number of diacritics available for each cover text was specified, which are the bounds of the Holy Quran (Al-Kawthar, Al-Sharh, AL-Shams and Al-Fatihih). For example, Surat Al-Kawthar contains 54, which is the sum of the valid Fathah and Kasrah that can be hidden or concealed. The capacity is then calculated by the usage of the diacritics that will be removed from the cover text. For example, if 9 bits are hidden and contain five zeros, the five zeros are the number of diacritics that will be removed from the Cover text. So $54 - 5 = 49$ remaining diacritics.

In this experiment, four different Surahs of the Holy Qur’an have been extracted and tested to hide and conceal sensitive data, of different sizes. In this capacity comparison, the larger amount or volume of sensitive data that needs to be hidden, the more diacritics need to be used. In large amounts of data, more than two diacritics need to be used, as in study [23]. Fig. 10 and Fig. 11 show hiding 9 bits (100010110) in the cover text, which is Arabic poetry, using 2 diacritics and all respective diacritic approaches.

Observing the results from running the system on the Holy Quran Surahs as cover texts, we note that Surat Al-Fatihah has the largest number of diacritics. It has the lowest count of Fathah and Kasrah diacritics. Surat



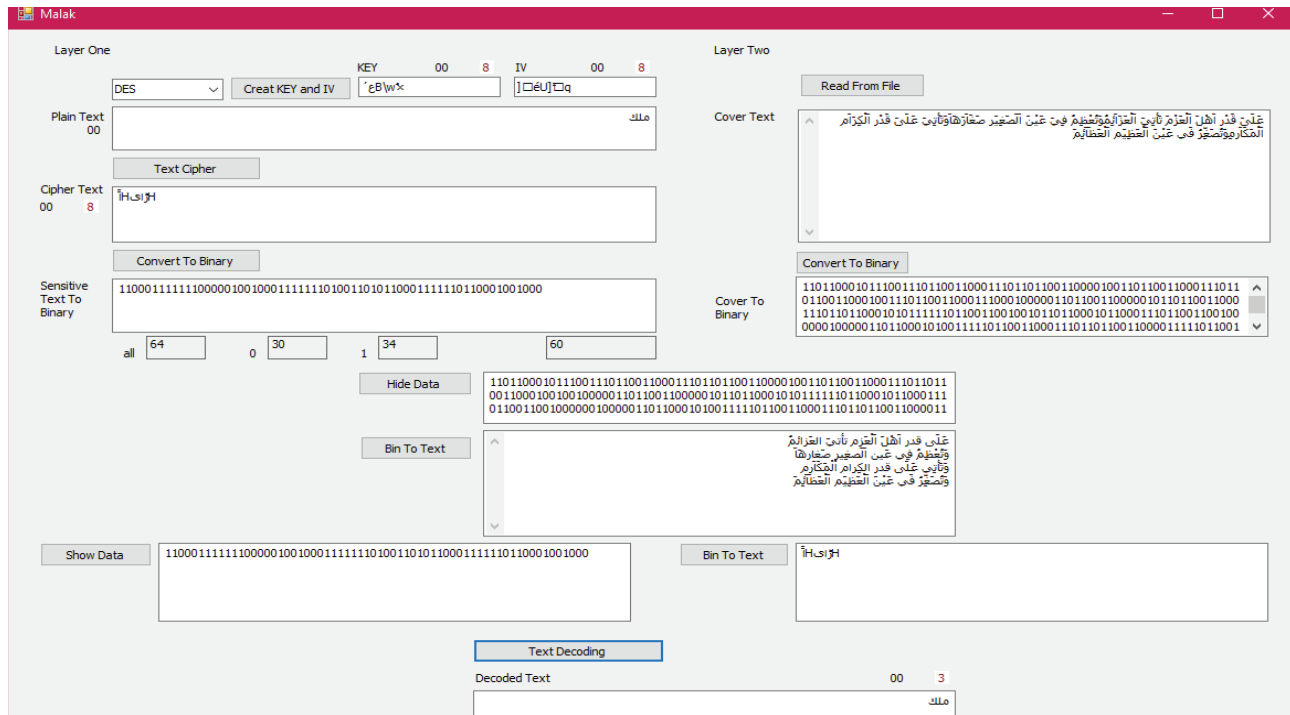


Fig. 9. The integrated system interface.

على قدر أهل العزم تأتي العزائم وتأتي على قدر الكرام المكارم وتعلم في عين الصغير صغارها وتنعمر في عين
العظيم العظيم يكافئ سيف الذؤلة الخبيث همه وقد عجزت عنه الجيوش الخصارم ويطلب عند الناس ما جند نفسه وذلك ما
لا تدعيه الصر أعم يفدي أتم الطير عمرأ سلاحه نسور الألا أحداتها والفساحم وما صرّها خلق بغير مخالب وقد خلقت
أستيفه والفوائم هل الحدت الحمراء تعرف لونها وتعلم أي السافين الغمام سفتها الغمام الغر قبل نرؤيه فلما دنا منها سفتها
الجماجم بناها فأعلى والفنا يقرع الفنا وموج المنابيا حولها متلاطم

Fig. 10. Hiding 9 bits secret (100010110) using two diacritics scheme.

Al-Fatihah represents the highest capacity, compared with the rest of the four Surahs. TABLE IV shows the number of diacritics, as well as their percentages, tested via Holy Quran Suras as cover text. TABLE V summarizes the comparison between all diacritics and the two diacritics approach tested on the Holy Quran Suras.

Noticing the results above, we conclude that concealment is highest within completed diacritics, in terms of capacity. TABLE VI shows the capacity results for the four chosen Surahs, in cases of both two diacritics, and all diacritics.

Fig. 12 shows the percentage comparison between the two approaches, which proves that it's better in our case to use all diacritics in hiding secret messages.

B. Security Comparison

In this section, we study security. We show that the

proposed integration system provides acceptable hiding results, without affecting security. It's important to make sure that automatically predicting hidden information should be complex, using the security testing methods, such as histogram investigations and peak signal-to-noise ratio (PSNR) analysis. The histogram tests are used to study the security of the three different LWC from our integration prospective. The PSNR is adopted for the security study of the two steganography procedures, i.e. our improved all diacritics stego approach, as well as the two diacritics approach, justifying the promising acceptability of our integrated Arabic text security proposal, in an interesting manner.

1) Histogram Investigations:

Histograms are useful statistical data representations of pixels to be used for comparisons. They require the



على قدر أهل العزم تأتي العزائم وتأتي على قدر الكرام المكارم وتُعظم في عين الصغیر صغارها وتُصغر في عين العظیم العظیم يكلف سيف الدولة الجیش همه وقد عجزت عنه الجنوش الخصارم ويطلب عند الناس ما عند نفسه وذلك ما لا تدأخيه الصراخم يفتدي أتم الطير حمرأ سلاحه نسور الفلا أحداثها والفناجم وما صرّها خلق بغير محالب وقد خلقت أسيافه في القوائم هل الحدث الحمرأء تعرف لونها وتعلم أي السافينين الغمام سفتها الغمام الغر قبل نزلها فلما دنا منها سفتها الجمأجم بناها فاعلى والفنا يفرغ الفنا وموؤج المنأيا حولها متلاطم

Fig. 11. Hiding 9 bits secret (100010110) using proposed all diacritics scheme.

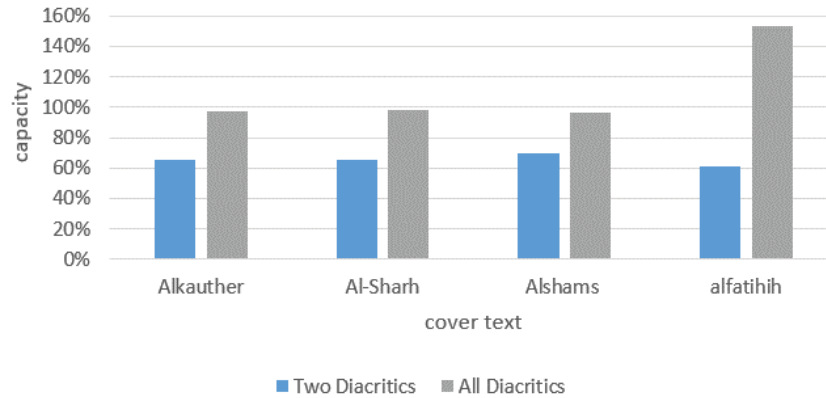


Fig. 12. Capacity comparison average graph of hiding within Holy Quran Surahs..

TABLE II
THE BITS USED IN THE EXPERIMENT AND THEIR VALUES.

Number of bits	Value
9 Bits	100010110
20 Bits	10110010101010010010
30 Bits	1011001011110010101010 01010010
50 Bits	100100111001011100110100101010 01010010010101001101
60 Bits	101100101111001010101001010010 010110010111100101010100101001
100 Bits	010110010111100101010100101001101100101111001010101001010010
120 Bits	1011001011110010101010010100100 1011001011110010101010010100110 1100101111001010101001010010010 110010111100101010100101001
240 Bits	1011001011110010101010010100100 1011001011110010101010010100111 0110010111100101010100101001001 0110010111100101010100101001011 0010111100101010 100101001001011 00101111001010101001010011011001 01111001010101001010010010110010111100101010100101001
480 Bits	11001011110010101 1100101010100101001110 10110010111100101010010100100101100101 10010100100101100101 1100101010 1011001011 0100101001001011001011110010101010010100 1001010010010110010111100101010100101001 0111100101010 111001010101001010011011001 1110010101010010100111011001011110010101 0100101100101 1010101001010 10110010111100 10010100100101100101 10110010111100101010 101010010100 1011110010 101100 010010100100 10101001010011011001011110010101001010010010110010111100101010100101001 1110010



TABLE III
DIACRITICS STEGANOGRAPHY COMPARISON.

Secret Bits Hidden	Two Diacritics previous scheme	Proposed All Diacritics approach
9 Bits	عَلَى قَدْرِ أَهْلِ الْعَزْمِ تَأْتِي الْعَزَائِمُ وَتَأْتِي عَلَى قَدْرِ الْكِرَامِ الْمَكَارِمُ وَتُعْطِمُ فِي عَيْنِ الصَّغِيرِ	عَلَى قَدْرِ أَهْلِ الْعَزْمِ تَأْتِي الْعَزَائِمُ وَتَأْتِي عَلَى قَدْرِ الْكِرَامِ الْمَكَارِمُ وَتُعْطِمُ فِي عَيْنِ الصَّغِيرِ
20 Bits	عَلَى قَدْرِ أَهْلِ الْعَزْمِ تَأْتِي الْعَزَائِمُ وَتَأْتِي عَلَى قَدْرِ الْكِرَامِ الْمَكَارِمُ وَتُعْطِمُ فِي عَيْنِ الصَّغِيرِ	عَلَى قَدْرِ أَهْلِ الْعَزْمِ تَأْتِي الْعَزَائِمُ وَتَأْتِي عَلَى قَدْرِ الْكِرَامِ الْمَكَارِمُ وَتُعْطِمُ فِي عَيْنِ الصَّغِيرِ
30 Bits	عَلَى قَدْرِ أَهْلِ الْعَزْمِ تَأْتِي الْعَزَائِمُ وَتَأْتِي عَلَى قَدْرِ الْكِرَامِ الْمَكَارِمُ وَتُعْطِمُ فِي عَيْنِ الصَّغِيرِ	عَلَى قَدْرِ أَهْلِ الْعَزْمِ تَأْتِي الْعَزَائِمُ وَتَأْتِي عَلَى قَدْرِ الْكِرَامِ الْمَكَارِمُ وَتُعْطِمُ فِي عَيْنِ الصَّغِيرِ
50 Bits	عَلَى قَدْرِ أَهْلِ الْعَزْمِ تَأْتِي الْعَزَائِمُ وَتَأْتِي عَلَى قَدْرِ الْكِرَامِ الْمَكَارِمُ وَتُعْطِمُ فِي عَيْنِ الصَّغِيرِ	عَلَى قَدْرِ أَهْلِ الْعَزْمِ تَأْتِي الْعَزَائِمُ وَتَأْتِي عَلَى قَدْرِ الْكِرَامِ الْمَكَارِمُ وَتُعْطِمُ فِي عَيْنِ الصَّغِيرِ
60 Bits	عَلَى قَدْرِ أَهْلِ الْعَزْمِ تَأْتِي الْعَزَائِمُ وَتَأْتِي عَلَى قَدْرِ الْكِرَامِ الْمَكَارِمُ وَتُعْطِمُ فِي عَيْنِ الصَّغِيرِ صَغَارَهَا وَتُصَغِّرُ فِي عَيْنِ	عَلَى قَدْرِ أَهْلِ الْعَزْمِ تَأْتِي الْعَزَائِمُ وَتَأْتِي عَلَى قَدْرِ الْكِرَامِ الْمَكَارِمُ وَتُعْطِمُ فِي عَيْنِ الصَّغِيرِ صَغَارَهَا وَتُصَغِّرُ فِي عَيْنِ
100 Bits	عَلَى قَدْرِ أَهْلِ الْعَزْمِ تَأْتِي الْعَزَائِمُ وَتَأْتِي عَلَى قَدْرِ الْكِرَامِ الْمَكَارِمُ وَتُعْطِمُ فِي عَيْنِ الصَّغِيرِ صَغَارَهَا وَتُصَغِّرُ فِي	عَلَى قَدْرِ أَهْلِ الْعَزْمِ تَأْتِي الْعَزَائِمُ وَتَأْتِي عَلَى قَدْرِ الْكِرَامِ الْمَكَارِمُ وَتُعْطِمُ فِي عَيْنِ الصَّغِيرِ
120 Bits	عَلَى قَدْرِ أَهْلِ الْعَزْمِ تَأْتِي الْعَزَائِمُ وَتَأْتِي عَلَى قَدْرِ الْكِرَامِ الْمَكَارِمُ وَتُعْطِمُ فِي عَيْنِ الصَّغِيرِ صَغَارَهَا وَتُصَغِّرُ فِي عَيْنِ الْعَظِيمِ الْعَظَائِمُ يُكَلِّفُ سَيْفَ الدَّوْلَةِ الْجَيْشَ هَمَّهُ وَقَدْ عَجَزَتْ عَنْهُ الْجَيْشُ الْخَضَارِمُ وَيَطْلُبُ عِنْدَ النَّاسِ مَا عِنْدَ نَفْسِهِ وَذَلِكَ مَا	عَلَى قَدْرِ أَهْلِ الْعَزْمِ تَأْتِي الْعَزَائِمُ وَتَأْتِي عَلَى قَدْرِ الْكِرَامِ الْمَكَارِمُ وَتُعْطِمُ فِي عَيْنِ الصَّغِيرِ صَغَارَهَا وَتُصَغِّرُ فِي عَيْنِ الْعَظِيمِ الْعَظَائِمُ يُكَلِّفُ سَيْفَ الدَّوْلَةِ الْجَيْشَ هَمَّهُ وَقَدْ عَجَزَتْ
240 Bits	عَلَى قَدْرِ أَهْلِ الْعَزْمِ تَأْتِي الْعَزَائِمُ وَتَأْتِي عَلَى قَدْرِ الْكِرَامِ الْمَكَارِمُ وَتُعْطِمُ فِي عَيْنِ الصَّغِيرِ صَغَارَهَا وَتُصَغِّرُ فِي عَيْنِ الْعَظِيمِ الْعَظَائِمُ يُكَلِّفُ سَيْفَ الدَّوْلَةِ الْجَيْشَ هَمَّهُ وَقَدْ عَجَزَتْ عَنْهُ الْجَيْشُ الْخَضَارِمُ وَيَطْلُبُ عِنْدَ النَّاسِ مَا عِنْدَ نَفْسِهِ وَذَلِكَ مَا لَا تَدْعِيهِ الضَّرَاعِمُ يَفْدِي أَمَّ الطَّيْرِ عَمْرًا سَلَاحَهُ نَسُورَ الْفَلَائِحِ وَأَخْدَانُهَا وَالْقَشَاعِمُ وَمَا ضَرَّهَا خَلْقٌ بِغَيْرِ مَخَالِبٍ وَقَدْ خَلَقَتْ أَسْيَافُهُ وَالْقَوَائِمُ هَلِ الْحَدِيثُ الْحَمْرَاءُ تُعْرَفُ لَوْنُهَا وَتُعَلَّمُ أَيُّ السَّافِيَيْنِ الْغَمَامُ الْغُرُقَيْلُ	عَلَى قَدْرِ أَهْلِ الْعَزْمِ تَأْتِي الْعَزَائِمُ وَتَأْتِي عَلَى قَدْرِ الْكِرَامِ الْمَكَارِمُ وَتُعْطِمُ فِي عَيْنِ الصَّغِيرِ صَغَارَهَا وَتُصَغِّرُ فِي عَيْنِ الْعَظِيمِ الْعَظَائِمُ يُكَلِّفُ سَيْفَ الدَّوْلَةِ الْجَيْشَ هَمَّهُ وَقَدْ عَجَزَتْ عَنْهُ الْجَيْشُ الْخَضَارِمُ وَيَطْلُبُ عِنْدَ النَّاسِ مَا عِنْدَ نَفْسِهِ وَذَلِكَ مَا لَا تَدْعِيهِ الضَّرَاعِمُ يَفْدِي أَمَّ الطَّيْرِ عَمْرًا سَلَاحَهُ نَسُورَ الْفَلَائِحِ وَالْقَشَاعِمُ وَمَا ضَرَّهَا خَلْقٌ بِغَيْرِ

TABLE IV
COUNTING PERCENTAGES OF THE AUTHENTIC HOLY QURAN SURAHS AS COVER TEXT.

	Al-Kawthar		Al-Sharh		Al-Shams		Al-Fatihah	
	Count	Percentage	Count	Percentage	Count	Percentage	Count	Percentage
Fathah	42	(42/82)*100=51%	52	(52/94)*100=55%	125	(125/201)*100=62%	1256	(1256/2929)*100=42%
Kasrah	12	(12/82)*100=14%	10	(10/94)*100=10%	18	(18/201)*100=8%	576	(576/2929)*100=19%
Dammah	4	(4/82)*100=4%	4	(4/94)*100=4%	16	(16/201)*100=7%	360	(360/2929)*100=12%
Sukun	16	(16/82)*100=19%	22	(22/94)*100=23%	27	(27/201)*100=13%	447	(447/2929)*100=15%
Tanwin Fathah	0	0	2	(2/94)*100=2%	0	---	23	(23/2929)*100=0.7%
Tanwin Kasrah	0	0	0	0	1	(1/201)*100=0.4%	42	(42/2929)*100=1.4%
Tanwin Dammah	0	0	0	0	0	---	35	(35/2929)*100=1.1%
Shaddah	8	(8/82)*100=9%	4	(4/94)*100=4%	14	(14/201)*100=6%	190	(190/2929)*100=6.4%



TABLE V
CAPACITY COMPARISON PERCENTAGES VIA HOLY QURAN SURAHS AS COVER TEXT.

	Al-Kawthar	Al-Sharh	Al-Shams	Al-Fatihih
Two Diacritics 'Fathah and Kasrah'	65%	67%	70%	61%
All Diacritics	97%	98%	96%	153%

TABLE VI
CAPACITY COMPARISON OF HIDING DIFFERENT SECRETS WITHIN HOLY QURAN SURAHS.

Secret Bits Hidden/ Number of Diacritics	Capacity = (number of zeros *16) /8	Al-Kawthar		Al-Sharh		Al-Shams		Al-Fatihih	
		Two	Full	Two	Full	Two	Full	Two	Full
		54	82	62	94	143	201	1832	2929
9 Bits	$(5*16)/8 = 10$ byte	49	77	57	89	138	196	1872	2924
20 Bits	$(11*16)/8 = 22$ byte	43	71	51	83	132	190	1821	2918
30 Bits	$(15*16)/8 = 30$ byte	39	67	47	79	128	186	1817	2914
50 Bits	$(26*16)/8 = 52$ byte	28	56	36	68	117	175	1806	2903
60 Bits	$(30*16)/8 = 60$ byte	24	52	32	64	113	171	1802	2899
120 Bits	$(60*16)/8 = 120$ byte	-	22	2	34	83	141	1802	2869
240 Bits	$(120*16)/8 = 240$ byte	-	-	-	-	23	81	1712	2809
480 Bits	$(240*16)/8 = 480$ byte	-	-	-	-	-	-	1592	2689



data to be formed as images to enable comparison, providing a security difference estimation relationship. The histogram analysis covered the tests for all scenarios of multi-bits letters encryption, through only displaying the even numbers (as evidence) in TABLE VII. The original cover text testing, represented as an image for this histogram study depiction, is shown in the first column in TABLE VII. Note that we run the histogram investigation for the secret process using MATLAB 2015 on the Windows 10 operating system of a personal computer, with 2.6 GHz core i7 CPU and 8G DDR3 ram. Illustrative samples of the MATLAB histogram results of many different tests are exemplified, as shown in TABLE VII, to help understand the LWC selection within this work. It is important to mention that this work is tested on the available PC platform, intending to provide a full comparative analysis between three known LWC structures. The work is performed for research investigation and academic elaboration, helping in providing some understanding of the concepts, via using a PC platform. However, it cannot fully represent a real-life implementation solution, as it runs on limited capability devices, as needed for LWC equivalency requirements, which is outside of the research scope and available for future investigation.

The experiment includes using different letters, starting from 2 letters up to 16 in the encryption phase, using the three LWC algorithms. TABLE VII represents the samples of an even number of letters encryption histogram display, showing the effect of using the same inputs on the three encryption algorithms (AES, DES and IDEA), i.e. to compare between them in the case of using all different letters. The analysis uses histograms of the stego-cover as being the most accurate, when it's the closest to the histogram of the original cover-text, representing its security acceptability.

Considering TABLE VII, some remarks can be made. For example, it is important to note that, interestingly, in the case of using AES with an 8 letter encryption, the histogram gives an extreme value, which is far from the original histogram of the cover text, indicating its low security when considering this. However, when observing all of the histograms tests together, for all trials of different letters used, it's noted that IDEA gives the best histogram graph, compared with the original one, followed by DES, giving reasonable results, i.e. in between AES and IDEA. In fact, although DES gives some variations in most letters used, this work considers it to provide a more acceptable security overview in its results than

AES, which is a remarkable conclusion.

2) PSNR Comparison:

Peak signal-to-noise ratio (PSNR) can represent the relationship between the maximum powers or values of any signal with the power of any noise corrupting that signal. In our integration system scope, PSNR represents the ratio between the original text and the text received, once the encryption and hiding process have taken place. It is used to give an indication of the applicability of our method, in ensuring the hiding process of secrets are embedded ambiguously. As this PSNR value is unusual, it suggests that the secret is detectable, and the system privacy is sabotaged.

Like histograms, PSNR is usually applied to images, and it doesn't depend on image intensity scaling. To calculate the PSNR in the opportunity of Arabic Text security, two images are compared in every trial. The first image contains the cover text as in Fig. 13, while the second image contains the stego image, i.e. the image of the cover text with the secret message encrypted and hidden within it. The PSNR is remarked in decibels, showing the difference between the two images. A higher PSNR value represents a higher quality of image achieved. If the PSNR image quality is higher, data is well-hidden and cannot easily be detected, meaning higher security.

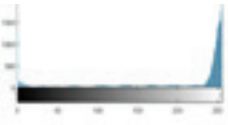
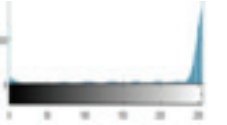

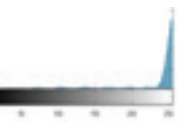
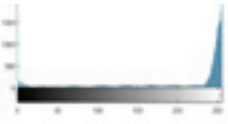



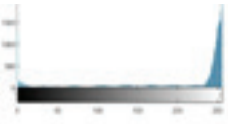

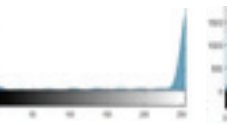
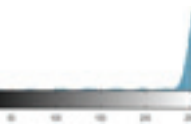
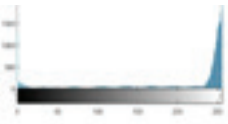







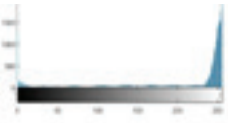


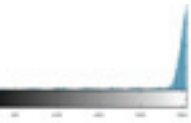
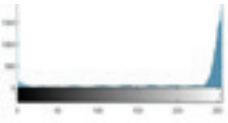



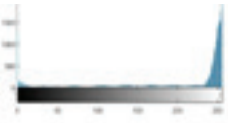



Fig. 14 briefly provides a comparison between using the stego techniques. It shows the PSNR results of using the two diacritics approach and using the all diacritics approach for hiding secret messages. The study shows that using our improved 'all diacritics' scheme provides more average PSNR than the previous two-diacritics procedure. This indicates that our work promises to retain better image quality, more like the originals, indicating higher security. This similarity between the stego-image of our system and the original image is a remarkable declaration which indicates the effectiveness of our work.

V. CONCLUSION

This paper suggests an integrated crypto-stego Arabic text security system. It integrated lightweight cryptography (LWC) within diacritics steganography to fully secure sensitive data. The work needed the cryptography procedure to protect privacy, and the steganography process to completely hide the presence of information, with full dependence on the user and their preferred available data. The selection of LWC is needed to cope with to-



TABLE VII
 SECURITY HISTOGRAMS LWC COMPARISON TESTING DIFFERENT AES, DES AND IDEA SCENARIOS.

LWC	Original	AES	DES	IDEA
2 letters encrypted				
4 letters encrypted				
6 letters encrypted				
8 letters encrypted				
10 letters encrypted				
12 letters encrypted				
14 letters encrypted				
16 letters encrypted				



عَلَى قَدْرِ أَهْلِ الْعَزْمِ تَأْتِي الْعَزَائِمُ وَتَأْتِي عَلَى قَدْرِ الْكِرَامِ الْمَكَارِمُ وَتُعْطِمُ فِي عَيْنِ الصَّغِيرِ صَغَارَهَا وَتُصَغِّرُ فِي عَيْنِ
 الْعَظِيمِ الْعَظَائِمَ يُكَافُ سَيْفُ الدَّوْلَةِ الْجَيْشَ هَمَهُ وَقَدْ عَجَزَتْ عَنْهُ الْجَبُوشُ الْخِصَارُ وَيَطْلُبُ عِنْدَ النَّاسِ مَا عِنْدَ نَفْسِهِ وَذَلِكَ مَا
 لَا تَدَّعِيهِ الصَّرَاحُ عَمَّ يَفْدِي أَمَّ الطَّيْرِ عُمْرًا سِلَاحَهُ نَسُورُ الْفَلَاحِ إِخْدَانُهَا وَالْفَسَاحِمُ وَمَا صَرَّهَا خَلْقٌ بِغَيْرِ مَخَالِبٍ وَقَدْ خَلَقَتْ
 أَسْبَابَهُ وَالْفَوَائِدُ هَلْ الْخَبْرُ تَعْرِفُ لَوْ نَهَا وَتَعْلَمُ أَيَّ السَّافِينِ الْعَمَامِ سَفَّتْهَا الْعَمَامُ الْغُرُ فَبَلِّ نَزُولِهِ فَلَمَّا دَنَا مِنْهَا سَفَّتْهَا
 الْجَمَامُ بَنَاهَا فَأَعْلَى وَالْفَنَاءُ يَفْرَحُ الْفَنَاءُ وَمَوْجُ الْمَنَابِ حَوْلَهَا مُتَلَابِطٌ

Fig. 13. Reference image of the testing original cover text.

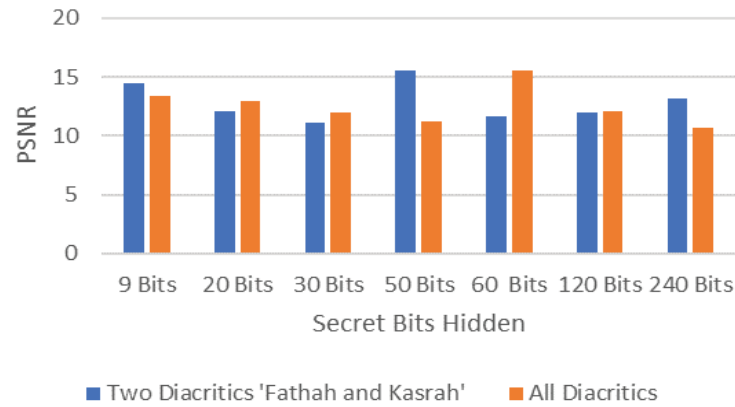


Fig. 14. Security PSNR comparison.

day's urgency for practical encryption computations, desired within mobile devices with limited resources. The proposed steganography approach is new. Its idea came from the evolution applied to the steganography method, of utilizing the double ('Kasrah' and 'Fathah') diacritics for embedding data. This Arabic steganography novelty is considered, via hiding sensitive encrypted data within all eight Arabic diacritics, i.e. to insert data within a full-scale diacritic method, integrating it with LWC.

The work experimented with integration strategies, testing three practical LWC schemes, i.e. AES, DES and IDEA, embedded via two Arabic text steganography diacritic procedures. The work tested performance by considering the capacity of hiding an amount of data within a given Arabic text, as well as investigating security via the similarities between histogram and PSNR analysis. The new full diacritics method was validated and compared with the two diacritics method, and tested on the text of a select poem, as well as some Holy Quran Surahs, taking them to be authentic fixed diacritic texts. The results showed a noticeable increase in capacity, without degrading security. The capacity results of our proposed full diacritic method showed over 90% general improvements, and a further 95% improvement, when validated using the holy Quran Surahs. From a security perspective, this proposed integration method is found to be better than others, based on PSNR analysis, preferring IDEA among the other two LWC encryption hiding processes. This

work has opened up research on real-world applications, benefitting from its simplicity in dealing with texts, i.e. for security purposes, in storing and transferring Arabic messages. The results are positive, opening the door for future research into the considerations of integrating other dissimilar Arabic steganography techniques, as well as LWC for hiding security information, in addition to using different Arabic, Urdu, or Farsi language contexts as test materials. This information hiding research is showing attractive results, which can be applied to current real-life security applications. It can be further useful to other Arabic related languages such as Urdu and Farsi, paving the way for promising text security research studies to come.

ACKNOWLEDGMENT

Thanks to Umm Al-Qura University for supporting this research via its Master program in computer sciences and Engineering.

REFERENCES

- [1] V. Goyal, O. Pandey, A. Sahai and B. Waters, "Attribute-based encryption for fine-grained access control of encrypted data," *Proc. 13th ACM Conf. Comput. Commun. Secur.*, Alexandria, VA, USA, Oct. 2006, pp. 89-98, doi: 10.1145/1180405.1180418.
- [2] T. Shirai, K. Shibutani, T. Akishita, S. Moriai and T. Iwata, "The 128-Bit Blockcipher CLEFIA," *Proc. 14th Int. Workshop Fast*



- Softw. Encryption FSE 2007*, A. Biryukov, Ed. Luxembourg, 2007, pp. 181-195.
- [3] A. Bogdanov et al, "PRESENT: An Ultra-Lightweight Block Cipher," *Proc. 9th Int. Conf. Cryptogr. Hardw. Embed. Syst. CHES 2007*, P. Paillier and I. Verbauwhede, Eds. Vienna, Austria, Sept. 2007, pp 450-466.
- [4] A. Bogdanov, M. Knezevic, G. Leander, D. Toz, K. Varici and I. Verbauwhede, "SPONGENT: The Design Space of Lightweight Cryptographic Hashing," in *IEEE Trans. Comput.*, vol. 62, no. 10, pp. 2041-2053, Oct. 2013, doi: 10.1109/TC.2012.196.
- [5] J. Guo, T. Peyrin and A. Poschmann, "The PHOTON Family of Lightweight Hash Functions," *Proc. 31st Annu. Cryptol. Conf. CRYPTO 2011*, P. Rogaway, Ed. Santa Barbara, CA, USA, Aug. 2001, pp. 222-239.
- [6] S. Hirose, K. Ideguchi, H. Kuwakado, T. Owada, B. Preneel and H. Yoshida, "A Lightweight 256-Bit Hash Function for Hardware and Low-End Devices: Lesamnta-LW," *Proc. 13th Int. Conf. Inf. Secur. Cryptol.*, K.-H. Rhee and D. Nyang, Eds. Seoul, Korea, Dec. 2010, pp. 151-168.
- [7] H. K. Tayyeh, M. S. Mahdi and A. S. Al-Jumaili, "Noval steganography scheme using Arabic text features in Holy Quran," in *Int. J. Electr. Comput. Eng (IJECE)*, vol. 9, no. 3, pp. 1910-1918, June 2019, doi: 10.11591/ijece.v9i3.pp1910-1918.
- [8] N. Alanazi, E. Khan and A. Gutub, "Functionality-Improved Arabic Text Steganography Based on Unicode Features," in *Arab J. Sci. Eng.*, 2020, doi: 10.1007/s13369-020-04917-5.
- [9] A. Gutub and F. Al-Shaarani, "Efficient Implementation of Multi-image Secret Hiding Based on LSB and DWT Steganography Comparisons," in *Arab J. Sci. Eng.*, vol. 45, pp. 2631-2644, Feb. 2020, doi: 10.1007/s13369-020-04413-w.
- [10] S. M. Almutairi, A. A. Gutub and N. A. Al-Juaid, "Motivating teachers to use information technology in educational process within Saudi Arabia," in *Int. J. Technol. Enhanc. Learn.*, vol. 12, no. 2, pp. 200-217, Feb. 2020, doi: 10.1504/IJTEL.2020.106286.
- [11] S. M. Almutairi, A. Gutub and M. Al-Ghamd, "Image Steganography To Facilitate Online Students Account System," in *Rev. Bus. Technol. Res.*, vol. 16, no. 2, pp. 43-49.
- [12] N. Alassaf and A. Gutub, "Simulating Light-Weight-Cryptography Implementation for IoT Healthcare Data Security Applications," in *Int. J. E-Health Med. Commun. (IJEHMC)*, vol. 10, no. 4, pp. 1-15, doi: 10.4018/IJEHMC.2019100101
- [13] N. Al-Juaid and A. Gutub, "Combining RSA and audio steganography on personal computers for enhancing security," in *SN Appl. Sci.*, vol. 1, no. 8, July 2019, Art. no. 830.
- [14] M. H. Shirali-Shahreza and M. Shirali-Shahreza, "A New Approach to Persian/Arabic Text Steganography," *5th IEEE/ACIS Int. Conf. Comput. Inf. Sci. 1st IEEE/ACIS Int. Workshop Compon.-Based Softw. Eng., Softw. Archit. Reuse (ICIS-COM SAR'06)*, Honolulu, HI, 2006, pp. 310-315, doi: 10.1109/ICIS-COM-SAR.2006.10.
- [15] W. Bender, D. Gruhl, N. Morimoto and A. Lu, "Techniques for data hiding," in *IBM Syst. J.*, vol. 35, no. 3.4, pp. 313-336, 1996, doi: 10.1147/sj.353.0313.
- [16] A. M. Al-Nofaie and A. A. Gutub, "Utilizing pseudo-spaces to improve Arabic text steganography for multimedia data communications," in *Multimed. Tools Appl.*, vol. 79, no. 1-2, pp. 19-67, doi: 10.1007/s11042-019-08025-x.
- [17] A. Gutub and K. Alaseri, "Hiding Shares of Counting-Based Secret Sharing via Arabic Text Steganography for Personal Usage," in *Arabian J. Sci. Eng.*, vol. 45, no. 4, pp. 2433-2458, July 2019, doi: 10.1007/s13369-019-04010-6.
- [18] A. Gutub and M. Fattahi, "A Novel Arabic Text Steganography Method Using Letter Points and Extensions," in *Int. J. Comput. Electr. Autom. Contr. Inf. Eng.*, vol. 1, no. 3, Mar. 2007, pp. 28-31.
- [19] S. M. Al-Nofaie, A. Gutub and M. Al Ghamdi, "Enhancing Arabic Text Steganography for Personal Usage Utilizing Pseudo-Spaces," in *J. King Saud Univ. - Comput. Inf. Sci.*, June 29, 2019, doi: 10.1016/j.jksuci.2019.06.010.
- [20] A. A. Gutub and K. A. Alaseri, "Refining Arabic Text Stego-Techniques for Shares Memorization of Counting-Based Secret Sharing," in *J. King Saud Univ. - Comput. Inf. Sci.*, June 28, 2019, doi: 10.1016/j.jksuci.2019.06.014.
- [21] S. M. Al-Nofaie, M. M. Fattani and A. A. Gutub, "Merging Two Steganography Techniques Adjusted to Improve Arabic Text Data Security," in *J. Comput. Sci. Comput. Math.*, vol. 6, no. 3, Sept. 2016, doi: 10.20967/jcscm.2016.03.004.
- [22] F. Al-Haidari, A. Gutub, K. Al-Kahsah and J. Hamodi, "Improving security and capacity for Arabic text steganography using 'Kashida' extensions," *2009 IEEE/ACS Int. Conf. Comput. Syst. Appl.*, Rabat, 2009, pp. 396-399, doi: 10.1109/AICC-SA.2009.5069355.
- [23] E. M. Ahmadoh and A. A. Gutub, "Utilization of Two Diacritics for Arabic Text Steganography to Enhance Performance," in *Lect. Notes Inf. Theory*, vol. 3, no. 1, pp. 42-47, June 2015, doi: 10.18178/lnit.3.1.42-47.
- [24] A. Odeh, A. Alzubi, Q. B. Hani and K. Elleithy, "Steganography by multipoint Arabic letters," *2012 IEEE Long Island Syst. Appl. Technol. Conf. (LISAT)*, Farmingdale, NY, 2012, pp. 1-7, doi: 10.1109/LISAT.2012.6223209.
- [25] A. A. Gutub and A. A. Al-Nazer, "High Capacity Steganography Tool for Arabic Text using 'Kashida'," in *The ISC Int'l J. Inf. Secur. (ISecure)*, vol. 2, no. 2, pp. 107-118, July 2010.
- [26] N. A. Al-Otaibi and A. A. Gutub, "2-layer security system for hiding sensitive text data on personal computers," in *Lect. Notes Inf. Theory*, vol. 2, no. 2, pp. 151-157, June 2014, doi: 10.12720/lnit.2.2.151-157.
- [27] N. Alanazi, E. Khan and A. Gutub, "Efficient security and capac-



- ity techniques for Arabic text steganography via engaging Unicode standard encoding,” in *Multimed. Tools Appl.* (2020), doi: 10.1007/s11042-020-09667-y.
- [28] N. Alassaf, A. Gutub, S. A. Parah and M. Al Ghamdi, “Enhancing Speed of SIMON: A Light-Weight-Cryptographic Algorithm for IoT Applications,” in *Multimed. Tools Appl.*, vol. 78, no. 23, Dec. 2019, doi: 10.1007/s11042-018-6801-z.
- [29] E. S. Hureib and A. A. Gutub, “Enhancing Medical Data Security via Combining Elliptic Curve Cryptography and Image Steganography,” in *Int. J. Comput. Sci. Netw. Secur. (IJCSNS)*, vol. 20, no. 8, pp. 1-8, Aug. 2020, doi: 10.22937/IJCSNS.2020.20.08.1.
- [30] N. Kheshaifaty and A. Gutub, “Preventing Multiple Accessing Attacks via Efficient Integration of Captcha Crypto Hash Functions,” in *Int. J. Comput. Sci. Netw. Secur. (IJCSNS)*, vol. 20, no. 9, pp. 16-28, Sept. 2020, doi: 10.22937/IJCSNS.2020.20.09.3.

