ABSTRACT

Digital speech signal is one of the most important types of data used due to the large number of computerized applications that are needed for mankind. Some computerized applications that use digital signals need a high level of security to protect the signal, and to turn it into a vague and incomprehensible signal to any third party, and that is why we have to seek a secure method of cryptography to protect the speech signal. In this research paper we will introduce a method based wavelet packet tree decomposition and reconstruction to generate an encrypted speech, the method will reorder the original speech using a secret number of decomposition levels, the obtained wavelet decomposition sizes will be used to form a speech segments, which must be rearranged in a secret order to get the encrypted speech, the introduced method will be implemented and evaluated to prove the security issues.

1. INTRODUCTION

The digital signal [3], [4], [5] such as sound [1] and digital image [2], [6], [7] is considered one of the most important types of data currently used and one of the most widely circulated data [8], [9], [10]. The digital audio signal may be confidential, it may be of a personal nature, or it may contain important data [11], [12], [13] which requires distortion and making it incomprehensible to any third party that is not authorized to deal with this signal. Digital signals are very important because they are used in vital applications such as in fingerprint recognition [14], [15], [16] systems and facial recognition systems [15], [16] to identify specific person.

Cryptography means encryption-decryption. The process of data encryption [17], [18] is only the process of destroying the original data so that this data becomes distorted and incomprehensible to any third party who is not authorized [19], [20] and this process is usually carried out by carrying out specific processing operations on the original data and by using a secret key (as shown in figure 1) [21], [22] that is known only by the sender and its recipient. As for the decryption process, it is executed on the encrypted data using the secret key and the specific processing operations to obtain data identical to the original without losing or losing any part of the information [23], [24], [25].
Speech Signal Cryptography Using WPT

The encryption and decryption method is good if you achieve the following things [26], [27], [28], [29], [42]:

- Security and protection so that it is difficult for any third party to penetrate data.
- Achieve a high level of distortion so that the error value between the original data and the encoded data is high.
- Achieving a very high percentage of correspondence between the original data and the data that was decoded so that the error ratio between them falls to zero.
- Do not lose or lose any part of the information during encryption and decryption.
- Ease of implementation.
- The speed of the encryption and decryption process.

For data protect [30] many methods were introduced to encrypt-decrypt digital signals. Some methods were based on signal segmentation [31], [32] others were based on adding and subtracting fixed noise to the signal [33], [34]. Some proposed methods were based on matrix multiplication and XORING Using huge private key [35], [36], [37] while others were based on signal blocking, dividing the original signal into blocks then each block was encrypted alone [38];

2. SIGNAL DECOMPOSITION USING WPT

Wavelet packet tree method (WPT) [39], [40], [41] can be easily used to decomposed digital speech signal into approximation and details applying the matlab function wavedec as shown in figure 2:

![Wavelet Packet Tree Method Diagram](image-url)
Here by selecting the decomposition level we can obtain a set of approximations and details (C in the figure) with a specified length of each (L in the figure), these components can be used to divide a speech signal into segment with a defined length, these segments can be rearranged in order to generate an encrypted speech signal, figure 3 shows an example of signal decomposition using WPT:

![Image](image1.png)

**Figure 3:** Speech signal decomposition example

### 3. THE PROPOSED METHOD

The proposed method of encryption-decryption is based on WPT decomposition, and the encryption phase as shown in figure 4 can be implemented applying the following steps:

1. Get the original speech signal.
2. Reshape the signal matrix (whether it is mono or stereo) to one row matrix.
3. Select a level of decomposition to be used as private key (PK1).
4. Apply wavedec
5. Get the length of each approximation.
6. Use the length to divide the speech matrix into segments.
7. Select a rearrangement order to rearrange the segment (this is to be used as PK2).
8. Rearrange the speech matrix to get the encrypted one.
9. Reshape back the matrix to get the encrypted speech signal.

![Image](image2.png)

**Figure 4:** Encryption phase
The decryption phase as shown in figure 5 can be implemented applying the following steps:
1) Get the encrypted speech signal.
2) Reshape the signal matrix (whether it is mono or stereo) to one row matrix.
3) Retrieve the level of decomposition to be used as private key (PK1).
4) Apply wavedec
5) Get the length of each approximation.
6) Use the length to divide the speech matrix into segments.
7) Retrieve the rearrangement order to rearrange the segment (this is to be used as PK2).
8) Rearrange the speech matrix to get the decrypted one.
9) Reshape back the matrix to get the decrypted speech signal.

![Figure 5: Decryption phase](image)

4. IMPLEMENTATION AND EXPERIMENTAL RESULTS

The following speeches shown in table 1 were recorded using sampling frequency 44100 samples per second; these speeches were used in our experiments:

<table>
<thead>
<tr>
<th>Speech number</th>
<th>Recorded speech</th>
<th>Size(sample)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Al-Balqa Applied University is a Jordanian university</td>
<td>260635</td>
</tr>
<tr>
<td>2</td>
<td>Speech encryption decryption method</td>
<td>181102</td>
</tr>
<tr>
<td>3</td>
<td>Amman is the capital city of Jordan</td>
<td>183544</td>
</tr>
<tr>
<td>4</td>
<td>Ziad Abdel kareem Alqadi</td>
<td>171261</td>
</tr>
<tr>
<td>5</td>
<td>Mohammad Khrisat</td>
<td>110368</td>
</tr>
<tr>
<td>6</td>
<td>Yousif Eltous</td>
<td>135114</td>
</tr>
<tr>
<td>7</td>
<td>Jihad Nader</td>
<td>96690</td>
</tr>
<tr>
<td>8</td>
<td>Saleh Khawatreh</td>
<td>109474</td>
</tr>
<tr>
<td>9</td>
<td>Aqaba is a wonderful city located on the red sea</td>
<td>223861</td>
</tr>
<tr>
<td>10</td>
<td>May God protect Jordanians and all mankind</td>
<td>220897</td>
</tr>
</tbody>
</table>
Figure 6 shows the plot of speech 1 and the encrypted version of this speech.

We selected a level of decomposition equal 6 (PK 1 = 5), then we applied decomposition for each of the used speech signal, table 2 shows the obtained segment length for each speech signal:

<table>
<thead>
<tr>
<th>Speech number</th>
<th>Segments length</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>8145</td>
</tr>
<tr>
<td>2</td>
<td>5660</td>
</tr>
<tr>
<td>3</td>
<td>5736</td>
</tr>
<tr>
<td>4</td>
<td>5352</td>
</tr>
<tr>
<td>5</td>
<td>3449</td>
</tr>
<tr>
<td>6</td>
<td>4223</td>
</tr>
<tr>
<td>7</td>
<td>3022</td>
</tr>
<tr>
<td>8</td>
<td>3422</td>
</tr>
<tr>
<td>9</td>
<td>6996</td>
</tr>
<tr>
<td>10</td>
<td>6904</td>
</tr>
</tbody>
</table>

The segments were rearranged using the rearrangement order (PK 2) shown in table 3:

<table>
<thead>
<tr>
<th>Segment order</th>
<th>New order (PK2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
</tr>
</tbody>
</table>

Using the new order the speeches were encrypted decrypted, table 4 shows the obtained results:
The error was calculated using formula 1:

\[
err = \sum_{i=1}^{n} \sqrt{(x_{1i} - x_{2i})^2}
\]  

Where \(x_1\) is the original signal; And \(x_2\) is the encrypted/decrypted signal

<table>
<thead>
<tr>
<th>Speech number</th>
<th>Error between original and encrypted speeches</th>
<th>Error between original and decrypted speeches</th>
<th>Encryption time(seconds)</th>
<th>Decryption time (Seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>68.1420</td>
<td>0</td>
<td>0.148000</td>
<td>0.148000</td>
</tr>
<tr>
<td>2</td>
<td>38.0654</td>
<td>0</td>
<td>0.131000</td>
<td>0.131000</td>
</tr>
<tr>
<td>3</td>
<td>41.7242</td>
<td>0</td>
<td>0.146000</td>
<td>0.146000</td>
</tr>
<tr>
<td>4</td>
<td>43.6924</td>
<td>0</td>
<td>0.129000</td>
<td>0.129000</td>
</tr>
<tr>
<td>5</td>
<td>29.9639</td>
<td>0</td>
<td>0.121000</td>
<td>0.121000</td>
</tr>
<tr>
<td>6</td>
<td>30.1948</td>
<td>0</td>
<td>0.128000</td>
<td>0.128000</td>
</tr>
<tr>
<td>7</td>
<td>34.0084</td>
<td>0</td>
<td>0.119000</td>
<td>0.119000</td>
</tr>
<tr>
<td>8</td>
<td>38.2777</td>
<td>0</td>
<td>0.124000</td>
<td>0.124000</td>
</tr>
<tr>
<td>9</td>
<td>41.0847</td>
<td>0</td>
<td>0.141000</td>
<td>0.141000</td>
</tr>
<tr>
<td>10</td>
<td>43.8856</td>
<td>0</td>
<td>0.155000</td>
<td>0.155000</td>
</tr>
</tbody>
</table>

From the obtained experimental results we can raise the following facts:

- It is easy to use WPT decomposition for signal encryption-decryption.
- It is easy to change the number of levels and the number of segments.
- The encryption-decryption method is secure by using 2 private keys.
- It is easy to change PK2.
- The proposed method provides a good quality of encryption by providing a high error between the original and the encrypted signal.
- The proposed method is efficient by providing significantly small time for encryption-decryption.

5. CONCLUSION

A method based on WPT decomposition for speech signal encryption-decryption was proposed, implemented and tested. The obtained experimental results showed that this method is flexible in selecting the decomposition level, and in selecting the new segments order. It was shown that this method has a high quality and it was efficient in performing encryption-decryption without losing any piece of information.

SOURCES OF FUNDING

None.

CONFLICT OF INTEREST

None.

ACKNOWLEDGMENT

None.
REFERENCES


[10] Dr Rushdi S Abu Zneit, Dr Ziad AlQadi, Dr Mohammad Abu Zalata, A Methodology to Create a Fingerprint for RGB Color Image, IJCSMC, vol. 6, issue 1, pp. 205-212, 2017.


Speech Signal Cryptography Using WPT


[32] Dr. Saleh A. Khawatreh Dr. Majed Omar Dwairi, Prof. Ziad Alqadi, Dr. Mohammad S. Khrisat, Dr. Amjad Hindi, Digital color image encryption-decryption using segmentation and reordering, International Journal of Latest Research in Engineering and Technology (IJLRET), vol. 6, issue 5, pp. 6-12, 2020.


[40] Prof. Ziad Alqadi, Dr. Amjad Hindi, Dr. Majed Omar Dwairi, Dr. Mohammad S. Khrisat, Features Analysis of RGB Color Image based on Wavelet Packet Information, IJCSMC, vol. 9, issue 3, pp. 149 – 156, 2020.
