A Survey on Security of Multimedia using Various Soft Computing Techniques

Arun Kumar Bisoyi  
Dept. of CSE, College of Engineering and Technology, Bhubaneswar, Odisha, India

Subhalaxmi Das  
Dept. of CSE, College of Engineering and Technology, Bhubaneswar, Odisha, India

Amitav Mahapatra  
Dept. of CSE, College of Engineering and Technology, Bhubaneswar, Odisha, India

ABSTRACT
In the recent years, with the development of network and multimedia technology, multimedia data, especially image, audio and video data are used more and more widely in human society. In this regard, to protect multimedia contents, cryptography, which appears to be an effective way for information security, has been employed in many practical applications. This Paper, make a survey on various techniques for securities of Multimedia content using soft computing approach.

Keywords

1. INTRODUCTION
Around 40% of the world population has an internet connection today [1]. The popularity of internet increases in every year; this is because of use of entertainment, politics, economics, militaries, industries, education, multimedia data, etc. [2]. To provide security to those applications is the main issue over internet. To shutout those issue many encryption techniques are proposed to give security to multimedia data. Transferring multimedia data through communication networks requires the transmission in secured manner. Multimedia security aims at protecting the media from distortion by various attackers while it is being sent over the communication network [3]. There are many schemes and ongoing techniques for encryption of multimedia data, but in this paper we propose various encryption techniques using soft computing approaches for providing security to multimedia data (Color Image, Satellite Image and Video data)[4][5][6]. Encryption method can be the immediate solution to protect information against hacker or interceptor. Such type of techniques required encryption of an image through some sort of mathematical algorithm, where only the real party that shares the image could possible decrypt to use the image. Encryption is a algorithm to convert original message, known as plaintext, into cipher text, in the encrypted form. These are widely use of images in different-different processes [10]. Therefore, the security of image data from hackers and unauthorized users is important. Image encryption plays an important role in the field of information hiding and data security. Image hiding or encrypting method and algorithm can be very from simple methods to more complicated and reliable frequency method. Mainly available encryption algorithms used for text data cannot be suitable for multimedia data so there is need to develop image encryption based algorithms.

However, multimedia content such as digital images undergo various attacks like compression, format transformation and graphical alteration by filters and others. To provide security against attacks Neural network [11] and cryptography [12] together can make a great help in field of networks security [9]. The key generated by neural network is in the form of weights and neural functions which is hard to break. In this context, content data would be used as an input data for cryptography so that data become unreadable for attackers and remain secure from them. In this paper we make a survey of on various soft computing approaches associated encryption and decryption technique with different neural network techniques.

2. LITERATURE REVIEW
Shiguo Lian et al [4] has proposed a chaotic neural network based encryption processes, which provide high security at low cost. A fast video encryption scheme is proposed in this paper, which combines encryption process with MPEG4 encoding. This scheme supports direct bit-rate control, and keeps the error-robustness, compression ration and file format unchanged. This scheme is suitable for real-time applications like video transmission, video access. In this encryption scheme, some sensitive code blocks and bit-planes are encrypted, which make the decoded images too chaotic to be understood. Thus, the encryption scheme can get high security in perception.

In MPEG4 codec, video data are consists of many Video Sessions(VSs). Again each VS has three layers: 1) Video Object (VO), Video Object Layer(VOL) and Video Object Plane(VOP). Here, VOP is encoded by the embedded block coding with optimized truncation (EBCOT) [13] and, VOL encoding and EBCOT encoding are in relation with progressive property. The relationship between them is shown in Fig. 1. Where, each VOL represent a level of resolution.

The base VOL is composed of MxN sized VOPs. And each I-VOP is split into XXY sized code blocks, and each code block is encoded from the most significant bit plane to the least significant one. EBCOT codec encodes code blocks into
progressive bit-streams. The encryption scheme encrypts compressed bit-streams directly. The base VOL is composed of \(M\times N\) sized VOPs. And each I-VOP is split into \(X\times Y\) sized code blocks. The VOLS are numbered from basic layer to the highest layer, and bit-planes are numbered from the least significant one to the most significant one. Thus, the bit-planes can consist of a set in progressive order, \(P_{0,0,0}, P_{0,0,1}, \ldots, P_{0,1,0}, P_{0,1,1}, \ldots, P_{1,0,0}, P_{1,0,1}, \ldots, P_{0,0,0}, P_{0,0,1}, \ldots, P_{0,0,0}\). The video produced by encrypting the later bit-planes (for example, from \(P_{0,0,0}\) to \(P_{0,0,0}\)) are more chaotic than the one produced by encrypting the former bit-planes (for example, form \(P_{0,0,0}\) to \(P_{0,0,0}\)). Based on this set, selective encryption can be constructed by encrypting only some significant bit-planes. Thus, in order to obtain high security and high speed, they proposed the following encryption principles:

1) **File format information**
   The file format information, such as file header, packet header, and so on, should be left unencrypted in order to support such operation as image browsing or bit-rate control.

2) **VO Encryption**
   Only VOL\(_0\) is encrypted greatly, other VOLs are only encrypted by motion vector encryption. In VOL\(_0\), I-VOP is encrypted greatly, while P-VOP, BVOP or S-VOP is only encrypted by motion vector encryption. For each I-VOP, only the code blocks in the lowest sub-band are completely encrypted, and other ones are only encrypted by selective encryption that encrypts only \(q\) of the most significant bit-planes. Through various experiments, they recommend to select \(q=5\), which can get good tradeoff between security and time-efficiency.

3) **Cipher selection**
   Considering that the encoding-passes are often of variable size, block ciphers with fixed plaintext-size are not suitable here. This is because that the bit-rate control and error resilience are both based on encoding-passes that can be cut directly without affecting on the decoding process. Therefore, stream ciphers are more suitable here. However, no suitable random generators can obtain really random sequence now. The ones based on chaotic system and neural networks [24, 25] are regarded to have higher security compared with traditional pseudo-random number generators. Therefore, they propose to use the stream cipher based on chaotic neural network that is modified from the block cipher proposed in [26]. And it is often used to encrypt motion vectors signs. The encryption process is

\[
e_{j} = f(\omega_{i} p_{j} + \omega_{j-1} e_{j-1} + \theta_{j})
\]

Here, \(p_{i}\) (‘0’ or ‘1’) and \(c_{i}\) (‘0’ or ‘1’) are the i-th plaintext and cipher text respectively, \(f(x)\) is the function whose value is 1 if \(x \geq 0\) and 0 otherwise, and the parameters \(\omega_{j}\) and \(\theta_{j}\) are determined by a chaotic binary sequence \(B\) .

\[
\omega_{j} = \begin{cases} 
1 & b(2i) = 0 \\
-1 & b(2i) = 1
\end{cases}
\quad \text{and} \quad \theta_{j} = \begin{cases} 
1/2 & b(2i+1) = 0 \\
-1/2 & b(2i+1) = 1
\end{cases}
\]

Here, the chaotic binary sequence is generated based on Logistic map

\[
x(k+1) = \mu x(k)(1-x(k)).
\]

Here, they take \(\mu = 4\) and \(k \in (0,1, n-1)\). Initial state \(x(0)\) is the key, and the chaotic sequence \(x(0), x(1), \ldots, x(n)\) is produced through iterated chaotic map. If \(x(i) = 0.b_{0}(b_{1})(b_{2})(b_{3})\ldots\) then binary chaotic sequence \(b(0), b(1), b(2), \ldots \) may be constructed by extracting the first \(m\) bits of \(x(i)\). The bigger \(m\) is, the higher the key sensitivity is, and while the higher the computer resolution is required. Through various experiments, we recommend \(7 < m < 17\). The decryption process is symmetric to the encryption one, that is, the plaintext \(p_{i}\) is replaced by the cipher text \(c_{i}\) and the cipher text \(c_{i}\) by the plaintext \(p_{i}\). Thus, the cipher is a symmetric one, in which, the initial value \(x(0)\) is key that is assigned 128 bits.

4) **Pass encryption**
   Each encoding-pass is encrypted with different key. That is, for each encoding-pass, the chaotic binary sequence is generated from different initial-condition. Thus, if one encoding-pass cannot be synchronized because of transmission errors, the other ones can still be decrypted correctly.

The proposed encryption algorithm assures that the changed bit-stream can be still be decrypted correctly, which benefits from its advantage that the encoding passes are encrypted independently. In this encryption scheme, I-VOP frames are encrypted by selective bit-plane encryption, while other frames are encrypted by motion vector encryption. For I-VOP frames, the computational complexity of encryption process is in relation with the number of bit-planes to be encrypted: \(q\). Taking various videos with different sizes for example, we test the time ratio between encryption process and the summation of encryption and encoding process. Where, the number of encrypted bit-planes is \(q=5\). As a result they got that encryption time ratio is not bigger than 10%, which shows that the encryption process is of low cost. These properties make it suitable for real-time applications with different security or bit-rate requirement.

Ismail et al.[5] has proposed a back-propagation artificial neural network on the encryption of huge-sized satellite images. In artificial neural networks (ANN), feed-forward Multilayer perceptions (MLPs) utilize the back-propagation (BP) training algorithm [14]. A back-propagation network uses the back-propagation learning algorithm to learn a key to be used for encryption and decryption [15]. The used network is of \(N\times M \times N\) neurons representing the input, hidden, and output layers, respectively. The network is trained by adjusting the weights while the bias is given a constant value between 0 and 1 after normalizing the values. The bias between the input layer and the hidden layer works as the first key (K1), while the bias between the hidden layer and the output layer represents a second key (K2). Then, the network is used to encrypt and decrypt normal satellite images. This paper introduces a new idea for using neural back-propagation (NBP) in cryptography; we can securely exchange data between two sites \(S\) and \(R\) using NBP. The site \(S\) represents the sender (encryption), and the site \(R\) represents receiver (decryption). Weight and bias of the NBP represent public and private keys, respectively. The public key is given, the network is trained after the weight initialized and the bias may be taken a constant value after normalizing the values.

The encryption key is at least one bias (bias1 or bias2 or the two biases). The keys K1 or K2 or both will be constant (not updated) during training. The number of keys is determined according to the number of hidden layers in the network configuration. The key length depends on the number of neurons in the hidden layer for bias1 and the number of neurons in the output layer or input layer, for bias2. The keys must be numeric; if the keys are characters or strings,
The encryption algorithm works as follows.

1. The input image is segmented into L windows, or sub-images. Then, each sub-image is normalized [16].

2. The N normalized sub-images are input to the network, and the output of the hidden layer is computed. The output of the hidden layer is a vector of size ML, where M is the number of hidden neurons.

3. The N inputs are 8-bit integer numbers, and the outputs of hidden layer are M real values to transmit as the cipher text. Then the outputs of the hidden layer are quantized before transmission. This operation is done by rounding after multiplying each neuron’s output by 255.

4. The encrypted image is obtained by transforming the outputs of the hidden layer from one-dimensional array into a two-dimensional matrix.

The decryption algorithm works as follows.

1. The M elements of the encrypted image are input to the output layer of the decryption unit.

2. Using the response of the output layer, which is of size NL pixels, the decrypted image is extracted by transforming the array into a two-dimensional matrix.

The experimental result shows based on different SAR images, the goodness of fit (quality of decryption) between the original images and the decrypted ones was at least 98%, as per the proposed result. By experiment they found that the network is not affected by geometrical image distortions like translation, size, and rotation.

Yanling Liu et al [7] has proposed Cohen-Grossberg neural networks for removal of noise generated during transmission of encrypted color image. The hiding and storing of hidden data are possible because of Arnold transform (AT) [17] and Cohen-Grossberg neural network respectively. AT is used to hide the secret data. Then, the secret data is designed as the stable equilibria of Cohen-Grossberg neural networks in order to be stored. In the destination, when the message with noise is input to the network, the network can recall the original stored data. At last, via AT, the secret data without noise is decrypted into the original image.

Arnold Transform

Generally a cryptosystem includes plaintext, cipher text, encryption algorithm, decryption algorithm and key [17]. The basic mathematical relationship of cryptosystem can be shown as

\[ Y = E_K(X) \]
\[ X = D_K(Y) \]

where X is Plaintext, Y is Cipher text, EK presents Encryption algorithm, D K presents Decryption algorithm and K is Key.

AT, proposed by Arnold in 1968, is widely used in image encryption. The transform can scramble the pixel matrix of original digital image, which achieves image hiding. The two-dimensional AT of (x, y) is showed as [18].

The research result showed that AT can be iterated and has the periodicity denoted as T, which demonstrates the original image will be recovered after T iterations. Furthermore, its good periodicity ensures that the decryption is implemented accurately and easily. That implies the iterative number t and period T of AT are keys of recovering image accurately.

Cohen-Grossberg Neural Network

The Cohen-Grossberg neural network [19] is described as :-

\[ x_i = a_i(x_i) + \sum_{j=1}^{n} t_{ij} g_j(x_j + J_i), \quad i = 1, 2, ..., n. \]

where \( x_i \in \mathbb{R} \) is the state variable of the ith neuron, \( a_i(\cdot) \) is the gain function, \( b_i(\cdot) \) is the behaved function, \( g_i(\cdot) \) is the activation function, \( t_{ij} \) is areal constant, and \( t_i \) represents the connection weight between neurons i and j with \( t_{ij} \neq t_{ji} \). In this paper, they adopt a class of reduced Cohen-Grossberg neural networks proposed by Zheng et al. in [23]. This proposed method is mainly suitable for RGB image with size \( n \times n \) pixels. This method illustrated in Fig.2-Fig.3 has three phases, encryption phase, noise removal phase and decryption phase. In noise removal phase, if the image size is small, an equilibrium is designed in the network in Fig.4. Furthermore, if the image size is big, disassemble the given image into small blocks and design corresponding numbers equilibria. In the whole process, we use Matlab image processing toolbox to read and show images.
RGB color image is firstly interpreted as data matrices of class unit 8 and then store the original data matrix, denoted as $M_o$. $M_o$ can be interpreted as three matrices of class unit 8, denoted as $R$, $G$ and $B$, respectively. Next step, AT is executed for $t_1$. (1st < T) iterative numbers for the matrix $R$ to obtain a new matrix denoted as $R_1$. Likewise, $G_1$ and $B_1$ are obtained by performing AT for $t_2$ and $t_3$, iterative numbers for the matrix $G$ and $B$, respectively. In this step, three new image data matrices $R_1$, $G_1$ and $B_1$ are stored. In addition, the iterative number $t_1$, $t_2$, $t_3$ and the period T of AT are the encryption keys. Finally, the encrypted RGB image and encryption keys can be transmitted by channels.

During the transmission, there are some disturbances caused by the existence of noise. In order to execute effectively decryption phase, Cohen-Grossberg neural networks are added to remove the noise [23][6].

Mean Absolute Error (MAE) function is adopted in the proposed paper which is the difference between the original and the decrypted image to check the decrypted image with original one.

$$MAE = \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} |x_{ij} - y_{ij}|$$

Experimental results show that the proposed method achieves effective resistance against transmission noise.

Kensuke Naoe et al [8] has proposed a new key generation model for image hashing using neural network, which does not embed any data into the content but is able to extract meaningful data from target image. This model trains artificial neural network to assign predefined code and uses this trained artificial neural network weight and the coordinates of the selected feature sub blocks of target image as keys to extract the predefined code. The proposed method has ability to output same or similar image hash value under moderate modification but the image hash value will be destroyed under great modifications.

3. PROPOSED METHOD

They explained the procedures for generation of extraction keys and generation of image hash bit patterns from target image.

Key generation process consists of following procedures:

1) Feature extraction by frequency transformation of target image

2) Selection of the feature regions according to the feature extraction attributes and saved as feature extraction key

3) Prepare image hash patterns to be related to the target image.

4) Generation of extraction key to output image hash patterns by back propagation learning of neural network using feature extraction key.

5) Save the converged neural network classifier as image hash extraction key.

Extraction process consists of following steps:

1) Obtain feature extraction key and image hash extraction key

2) Feature extraction of target image using the feature extraction key

3) Construct a neural network model using image hash extraction key

4) Observe the output patterns from neural network using both keys

They determine the integrity, similarity or copyright by the observation of generated output patterns of neural networks of threshold value set to 0.5.

The proposed method does not limit to DCT as frequency transformation method only, one can use DFT and DWT otherwise. Also, machine learning method can be replaced with others such as Bayesian network and fuzzy.

![Image Hash Extraction and Comparison](image)

The experimental result showed that proposed method is robust to high pass filtering and image format transformation. Also, because proposed method does not alter the target content, it is applicable to steganography. Meanwhile, because the proposed method relies on the position of the feature subblocks, it is weak to geometric attacks like shrinking or rotation of the image.

Shyam Nandan Kumar et al [9] has proposed an encryption technique by using artificial neural network to protect data against unauthorized access. Based on three properties of Neural Network such as One-way Property, Learning Ability and Random Outcomes [20-22], which are suitable for Multimedia Content Protection. This paper proposes that how a complex media file (Video) encrypted and decrypted using neural network, which is difficult to break.

3.1 Designing of Neural Network

In this paper Video file are partitioned into blocks and each authentication code corresponds to a block and a neuron. For each simple neuron, a code bit s(0 or 1) and a feed forward of input training pattern, uses back propagation algorithm and set Learning rate and Momentum to 0.01 and 0.9 respectively.

Encryption process is from input layer to hidden layer and encrypted data are obtained from the output of hidden layer. Decryption process is from hidden layer to output layer and decrypted data are obtained from the output layer.
Enciphering Algorithm:
1. \( v_{ij} = p; w_{ij} = q; (p<100, q<100, i>0, j>0) \) // Initialize weights with random small value.
2. for(\( i=0; \ i<n; \ i++ \))
3. Set activation of input unit \( x_i \)
4. for(\( j=1; \ j<n; \ j++ \))
5. \( Z_j* = v_0j + \) // Encryption at hidden layer
6. \( Z_j = ( -1 ) \);
7. \( Z_j* = v_0j + \) // Encryption at hidden layer
8. }

Decryption Algorithm:
1. for(\( k=1; \ k<n; \ k++ \))
2. {
3. \( Y_k* = (w0k + \) ; // Decryption at output layer
4. \( Y_k = Y_k* \);
5. }}

4. SUMMARY

<table>
<thead>
<tr>
<th>Name of the author and paper</th>
<th>Main objective of the proposed technique</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shiguo Lian, Jinsheng Sun, Zhongxin Li, and Zhiquan Wang</td>
<td>Fast video encryption scheme is proposed, which combines encryption process with MPEG4 encoding,</td>
<td>1. This encryption scheme keeps compression ratio and file format unchanged, 2. supports direct bit-rate control, and keeps the error-robustness unchanged and Applicable to real-time applications.</td>
</tr>
<tr>
<td>Ismail I.A, Galal-Edeen, Khattab S and El Bahtity</td>
<td>To investigate the applicability of a back-propagation artificial neural network on the encryption of huge-sized satellite images</td>
<td>1) Not affected by the geometrical distortions of the input image, such as translation, scale, and rotation 2) Encryption of large sized</td>
</tr>
</tbody>
</table>

5. CONCLUSION
In this paper, we have represented some soft computing techniques which are highly required to provide more security to multimedia application. We have analyzed five different neural network techniques. Also we have elaborated the challenges, issues in those techniques.

6. ACKNOWLEDGMENTS
The authors are thankful to the department of Computer Science and Engineering, College of Engineering and Technology, Bhubaneswar.

7. REFERENCES


[16] Science, Engineering and Technology 17 2006, (pp60-64)


