Autoencoder Artificial Neural Network Public Key Cryptography in Unsecure Public channel Communication

Vikas Sagar, Krishan Kumar

ABSTRACT: As we all known that cryptography is a procedure to hide data so that it can’t be access or modified by any unauthorized entity. At the present digital world security is a main concern. To maintain this security there are many cryptographic algorithm exist. But the world technology grew each and every day so we have to find some new algorithms to maintain the security at higher level. In the proposed and implemented work used artificial neural network to increase the security during data communication in digital world. Autoencoder Neural Network is a new approach in the era of digital world so that used here in cryptographic algorithm to increase the strength of the security. There are three basic aims of cryptography availability, privacy and integrity easily achieved by this new approach. This work examine that the attacker can’t get access the data however he/she exist in the same network or not. Neural Network’s Uncertainty property can be easily achieved. This approach also examined on different data size and key size. Proposed work used the autoencoder for encryption and decryption. The final experimental result show our proposed algorithm efficient and accurate and also show how this approach perform better. Proposed and implemented algorithm can be easily used for secure data communication with more efficiently.

KEYWORDS: Cryptography, Symmetric key, Security, Autoencoder neural network.

I. INTRODUCTION

From a long time Cryptography used to maintain security in military and secret operations during information transformation. But nowadays all the transfer and communication done over the internet. 21st is a digital world and here Cryptography play a very important role for any kind of information transfer. That information may be for a particular person or for an organization. There important data or information keep safe and sound by the help of the cryptography from an unauthorized entity. In the cryptographic algorithms mathematical calculation used to secure the data. With mathematical operations secret data that’s called plain text convert in to scramble data that called cipher text. The plain text convert in to cipher text to maintain the confidentiality during the data transformation. There are two main stream to categorized cryptographic algorithm: one is symmetric key cryptography and other one is asymmetric key cryptography. In the symmetric key cryptography or private key cryptography a single key is used for both, encryption procedure and decryption procedure.

On the contrast asymmetric key cryptography used two different key, one for encryption procedure and other one for decryption procedure [1]. There are 4 primary goals of any cryptographic algorithm. These are:

1. **Confidentiality**
2. **Data integrity**
3. **Authentication**
4. **Non-repudiation**

**Confidentiality** refer to the data can’t be access by an unauthorized entity during the transformation from source to goal. Data should not be modified during the transfer by all means that called **Data integrity**. Modification include deletion, insertion and alternation. Third goal is the **Authentication** that means source should be verified by the receiver. Last one is **Non-repudiation**. **Non-repudiation** ensure that source or sink both cannot deny from their action of sending and receiving data.

Artificial Neural Network (ANN) is a mathematical or artificial model. We can say it’s a kind of simulation of biological neuron system. The simple and real example is human brain system. ANN is copy the human brain system, how brain network work and process each and every information itself. ANN is just perform the mimic behavior of that system.

Here we use the Autoencoder neural network to create a new cryptographic algorithm. Auto encoder is an unsupervised scheme of ANN. We compare this algorithm with previous exist algorithm and find out which algorithm perform better in different situation with different parameters to send secret information over an unsecure channel.

II. OUTLINE OF PUBLIC KEY CRYPTOGRAPHY

As we already know that a single key used in public key or Symmetric cryptography for encryption procedure and decryption procedure [1]. The main components of any cryptography scheme:

1. Plaintext
2. Encryption algorithm:
3. Secret Key:
4. Ciphertext:
5. Decryption Algorithm

Let’s take an example of information transformation between two entity A and B to understand these terms.
Autoencoder Artificial Neural Network Public Key Cryptography in Unsecure Public channel Communication

First of all **plaintext** is the main text or important data that entity A want to send to entity B with a security that no one can access except B. To achieve this security an **Encryption algorithm** used. In the encryption scheme different mathematical operation take place to scramble the data. To increase the security **Secret Key** is used with encryption scheme. After that encryption process the output data is called **Ciphertext** which scrambled and can’t be access by unauthorized entity. Then that cipher text send to the entity B over the public channel. B received the ciphertext and convert ciphertext in to the plaintext with the help of same secret key and **Decryption Algorithm**. Decryption algorithm is just reverse process of the encryption algorithm.

The overall public key cryptography process shown in figure:

![Public Key Cryptography](image)

**III. OUTLINE OF ANN**

Neural networks were first introduced by M. C. Culloch Pitts in 1943. ANN is inspired from biological neurons. ANN basically consists of small computing units, called neurons, arranged in different layers and interconnected with each other with some nerves connections called synapses. Simple mathematical computations are performed in each neuron. ANN tries to mimic the behavior of biological neurons and their interconnections. These are widely used in different areas together with other technologies for solving artificial intelligent problems, such as functions approximation, modeling, brain mapping, air traffic control, stock exchange, control systems, financial modeling, data compression, classification tasks or data processing [3], [6].

ANN divided in to two categories

**Supervised learning:** In supervised learning, the model work over the input and produce the output that is predefined or known.

**Unsupervised learning:** In contrast of supervised learning the unsupervised learning doesn’t have the known output. So that its work only on input by itself.

**IV. METHODOLOGY OF AUTOENCODER NEURAL NETWORK PUBLIC KEY CRYPTOGRAPHY**

ANN is a modern technique which perform better day by day in all real life areas. In this paper we proposed a new approach based on the Autoencoder artificial neural network (AANN). ANN can also perform an important role in the field of cryptography nowadays. There are many approaches already exist with ANN in cryptography. ANN is work on the mimic behavior of neurons which react according the real time events so that it can be gave effective result in the cryptography area.

**V. AUTOENCODER**

An Autoencoder also called autoassociator or Diabolo network is an ANN that is fall in the unsupervised learning category. Autoencoder neural network applies backpropagation process to setting the target values to be equal to the inputs. Because of its unsupervised, unpredictability and randomness property it’s very helpful in security [3], [4].

In simple words we can say that autoencoder is used to copy the input to output. Autoencoder compress their input by the procedure that’s called **latent-space representation**, and then regenerate the output same as input. Autoencoder is a combination of two individual units. That are-

1. **Encoder:** This part of the autoencoder neural network compresses the input into a latent-space representation. This encoding function entitled by
   \[ h = f(x) \]

2. **Decoder:** This part of autoencoder reassemble data that is generate by the encoder which is in latent space representation form. This decoding function entitled by
   \[ r = g(h) \]

![Autoencoder Architecture](image)

**Figure 2: Basic Architecture of an Autoencoder [16]**

**5.1 Classification of Autoencoder**

There are different types of Autoencoders:

1. **Undercomplete Autoencoders:** An autoencoder whose code measurement is not as much as the info measurement. Adapting such an autoencoder drives it to catch the most remarkable highlights. Be that as it may, utilizing a major encoder and decoder in the absence of enough preparing information enables the system to retain the errand and discards learning helpful highlights. In the event of having straight decoder it can go about as PCA. Nonetheless, adding nonlinear enactment capacities to the system makes it a nonlinear speculation of PCA.

2. **Regularized Autoencoders:** Rather than restricting the span of autoencoder and the code measurement for include learning, we can add a misfortune capacity to avert it retaining th assignment and the preparation data.
• Meager Autoencoders: An autoencoder which has a sparsity punishment in the preparation misfortune notwithstanding the reproduction mistake. They normally being utilized for the purpose of different assignments, for example, arrangement. The misfortune is not as straightforward as different regularizes, and we will talk about it in another post later.
• Denoising Autoencoders (DAE): The contribution of a DAE is a tainted duplicate of the genuine info which should be remade. Therefore, a DAE needs to fix the debasement (commotion) and additionally recreation.
• Contractive Autoencoders (CAE): The fundamental thought behind these kind of autoencoders is to take in a portrayal of the information which is powerful to little changes of the info.

3. Variational Autoencoders: They amplify the likelihood of the preparation information as opposed to duplicating the contribution to the yield and in this manner does not require regularization to catch helpful data.

5.2 Structure of Autoencoder
An autoencoder is combination of two, symmetrical neural network that may consist four or five superficial layers. The half layers of the network used for the encoding and second set superficial layers carry out the operation of decoding [12], [14].

The layers are limited Boltzmann machines, the building squares of profound conviction systems, with a few eccentricities that we'll talk about beneath. General structure of an autoencoder
- Maps an input $x$ to an output $r$ (called reconstruction) through an internal representation code $h$
- It has a hidden layer $h$ that describes a code used to represent the input
- The network has two parts
- The encoder function $h = f(x)$
- A decoder that produces a reconstruction $r = g(h)$

Here we’ll explain the simplified structure of the autoencoder schema.

![Figure 3: Schematic structure of an autoencoder with 3 fully-connected layers](Image)

Encoder $f$ and decoder $g$

$f : X \rightarrow h$
$g : h \rightarrow X$

$\arg \min ||X -(f \circ g) X||^2$

One hidden layer
• Non-linear encoder
• Takes input $x \in \mathbb{R}^d$
• Maps into output $h \in \mathbb{R}^e$

$h = \sigma_1(Wx + b)$
$x' = \sigma_2(W'h + b')$

Where $\sigma$ is an element-wise activation function such as sigmoid. Trained to minimize reconstruction error (such as sum of squared errors)

$L(x, x') = ||x - x'||^2 = ||x - \sigma_2(W'(\sigma_1(Wx + b) + b'))||^2$

Provides a compressed representation of the input $x$

VI. AUTOENCODER ARTIFICIAL NEURAL NETWORK PUBLIC KEY CRYPTOGRAPHY (AANNPKC)

Our algorithm based on the Autoencoder neural network. As we already know that autoencoder artificial neural network is an unsupervised learning scheme which is worked like backpropagation algorithm. The simplest example for an autoencoder is feed-forward neural network or no-recurrent neural network. Autoencoder based on encoder and decoder two symmetric individual neural network parts. Here encoder and decoder both used the same key. Now in term of cryptography here we suppose that two entity Alice and Bob want to communicate confidentially. To attain this objective Alice used autoencoder artificial neural network. Alice generate a training set for this. This training set based on the binary values of all alphabets, numbers and special symbols. Padding and key values also include for this training procedure. Now this training set securely pass to the Bob. Now Alice and Bob both use the same training set to train their autoencoder neural networks. Alice use binary values as feature vector to train her autoencoder neural network and generate the ciphertext. On the other hand Bob use the ciphertext as feature vector and recreate the original binary value.
Here we describe the whole process of the algorithm with the example how it’s work with real data. The main aim of the proposed algorithm if to maintain four pillars of cryptography (Confidentiality, Data integrity, Authentication, Non-repudiation) and that is finally tasted at the end of this paper and compare this algorithm with other pre-existing algorithms.

**Encryption:**
1. Very first step we input data from the sender that’s called Plain text or original text.
2. In the second step the original text turn in to ASCII form.
3. Here we turn the ASCII in to Binary format.
4. Now we use padding and secret key on this Binary data to make data satisfy for (Encoder) autoencoder neural network layer.
5. Apply the autoencoder and found the cipher text which send to the receiver as it is.

**Decryption:**
1. Cipher text receives at the receiver side and send directly to the (Decoder) autoencoder neural network layer with the secret key.
2. Then the output data convert to the Binary value and send to next step.
3. Here Binary value turn in to ASCII format and finally ASCII convert to corresponding to plain text.

**Encryption process with example on sender side:**

**STEP 1:** Firstly take the simple text massage.

Hi Bob! It is a very important text

**STEP 2:** Now in second step convert this initial value in to ASCII format


**STEP 3:** Convert the ASCII format in to binary format.

00110111 00100000 00110001 00110000 00100000 00110011 00110010 00100000 00110001 00110000 00100000 00110011 00110010 00100000 00110001 00110001 00110100 00110001 00110011 00110010 00100000 00110001 00110001 00110101 00100000 00110011 00110010 00100000 00110011 00110011 00110001 00100000 00111001 00111000 00100000 00110001 00110110 00100000 00110001 00110000

**STEP 4:** Now we use padding and secret key on this Binary data to make data satisfy for (Encoder) autoencoder neural network layer.

00110111 00110110 00100000 00111001 00110001 00110000 00110101 00100000 00110001 00110001 00110110 00110110 00100000 00110001 00110001 00110000 00100000 00110011 00110010 00100000 00110001 00110001 00110001 00110110 00110110 00100000 00110001 00110001 00110001 00110110 00110110 00100000 00110001 00110001 00110001 00110111 00110010 00100000 00110001 00110000

**STEP 5:** Apply the autoencoder and found the cipher text which send to the receiver as it is.

00100000 00110001 00110001 00110101 00110000 00110111 00100000 00110001 00110001 00110001 00110110 00110110 00100000 00110001 00110001 00110001 00110111 00110010 00100000 00110001 00110000

**Published By:**
Blue Eyes Intelligence Engineering & Sciences Publication

Retrieval Number: K1456098119/20190/BEIESP
DOI: 10.35940/ijitee.K1456.098119
Hi Bob! It is a very important text

This is the whole procedure how this algorithm work. Now let’s suppose that someone want to spy Alice and Bob’s conversation. There are two methods for that. First one spy entity create the same to sane two autoencoder with same training set that is almost impossible. Second is somehow spy entity try to crack the ciphertext with guessing or brute force attack that also take a huge amount of time. So that this algorithm is very much secure.

VII. Assessment Methodology And Analysis Results

The assessment methodology is a process to enable one better to comprehend the means expected to complete a quality evaluation. Here we simply used different algorithms codes which were generated manually some of them also downloaded from internet and some source code get from some papers. We try to maintain each and every algorithm as per rules.


Hardware Speciation: Here we use Dell Inspiron 15R 5000 series Intel Core i3 (2.40 GHz) fifth generation processor with 6GB ram with 1TB-HDD. All the algorithms are tested on Intel Core i5 (2.40 GHz).
Autoencoder Artificial Neural Network Public Key Cryptography in Unsecure Public channel Communication

Try to implement these encryption algorithm as fair as possible with all aspects. The performance of different algorithms may be change from previous data just because of hardware and software.

Figure 3: First view of the implemented software for all cryptographic algorithms.

Table 1: Cryptographic algorithms comparison with their basic properties

<table>
<thead>
<tr>
<th>Algorithm Name</th>
<th>Structure of Algorithm</th>
<th>Plain Text Length</th>
<th>Key Size (Bits)</th>
<th>S boxes</th>
<th>No. of Rounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>DES</td>
<td>Festial structure</td>
<td>64 bits</td>
<td>56</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>3DES</td>
<td>Festial structure</td>
<td>64 bits</td>
<td>56, 112, 168</td>
<td>8</td>
<td>48</td>
</tr>
<tr>
<td>Blowfish</td>
<td>Festial structure</td>
<td>64 bits</td>
<td>32-448</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>IDEA</td>
<td>Substitution-Permutation Structure (Lai-Massey scheme)</td>
<td>64 bits</td>
<td>128</td>
<td>N/A</td>
<td>8.5</td>
</tr>
<tr>
<td>AES (Rijndael)</td>
<td>Substitution-Permutation Structure</td>
<td>128 Bits</td>
<td>128, 192, 256</td>
<td>1</td>
<td>10, 12, 14</td>
</tr>
<tr>
<td>RC6</td>
<td>Festial structure</td>
<td>128 Bits</td>
<td>128, 192, 256</td>
<td>N/A</td>
<td>20</td>
</tr>
<tr>
<td>AANNPKC</td>
<td>Neural Network</td>
<td>128 Bits</td>
<td>N/A</td>
<td>N/A</td>
<td>Depends on Training Module or User</td>
</tr>
</tbody>
</table>

Here we present the average time of different cryptographic algorithms on different data size (500, 800, 1000 bytes). We take the average of 100 running executions on each data size to make the comparisons significance and perfect result.
Table 1: Average encryption and decryption time of different cryptographic algorithms on different data size.

<table>
<thead>
<tr>
<th>S. No</th>
<th>Algorithm Name</th>
<th>Key Size (bit)</th>
<th>File Size (bytes)</th>
<th>Encryption Time (Nano second) Average (100 Times)</th>
<th>Decryption Time (Nano second) Average (100 Times)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DES</td>
<td>56</td>
<td>500</td>
<td>343</td>
<td>344</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>800</td>
<td>357</td>
<td>360</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1000</td>
<td>375</td>
<td>379</td>
</tr>
<tr>
<td>2</td>
<td>3DES</td>
<td>168</td>
<td>500</td>
<td>339</td>
<td>322</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>800</td>
<td>348</td>
<td>358</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1000</td>
<td>360</td>
<td>379</td>
</tr>
<tr>
<td>3</td>
<td>Blow-fish</td>
<td>256</td>
<td>500</td>
<td>333</td>
<td>311</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>800</td>
<td>340</td>
<td>315</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1000</td>
<td>345</td>
<td>317</td>
</tr>
<tr>
<td>4</td>
<td>IDEA</td>
<td>128</td>
<td>500</td>
<td>320</td>
<td>313</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>800</td>
<td>329</td>
<td>320</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1000</td>
<td>335</td>
<td>340</td>
</tr>
<tr>
<td>5</td>
<td>AES (Rijndael)</td>
<td>256</td>
<td>500</td>
<td>340</td>
<td>369</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>800</td>
<td>360</td>
<td>370</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1000</td>
<td>370</td>
<td>391</td>
</tr>
<tr>
<td>6</td>
<td>RC6</td>
<td>256</td>
<td>500</td>
<td>334</td>
<td>340</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>800</td>
<td>341</td>
<td>350</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1000</td>
<td>342</td>
<td>399</td>
</tr>
<tr>
<td>7</td>
<td>AANNPKC</td>
<td>N/A</td>
<td>500</td>
<td>311</td>
<td>320</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>800</td>
<td>290</td>
<td>321</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1000</td>
<td>310</td>
<td>325</td>
</tr>
</tbody>
</table>

This following chat shows the encryption time of different cryptographic algorithms on different data size. It’s clear from the chat that the AANNPKC algorithm take the less time by any other algorithm for the encryption of any size of data.

This following chat shows the decryption time of different cryptographic algorithms on different data size. It’s clear from the chat that the AANNPKC algorithm take the less time by any other algorithm for the decryption of any size of data.
The above graph show the key value required for the different algorithms. It shows that the AANNPKC’s key depend on the plaintext only. Mostly the security depends on the key length on any algorithms but in AANNPKC key value depends on the plaintext length that’s means key value every time change according to that and security also depend on that.
The above diagrams demonstrate throughput for different algorithms. Throughput is characterized as the quantity of bytes sent over the channel per unit of time. It’s clear from the result that the AANNPKC have the best throughput in all algorithms that will make this algorithm best with respect to other algorithms.

Throughput = bytes sent / time in seconds

VIII. FLOW CHART OF PROPOSED ALGORITHM

The above flowchart show the flow of the data from sender side to receiver side with each and every step of encryption and decryption.
Autoencoder Artificial Neural Network Public Key Cryptography in Unsecure Public channel Communication

The optimality of the proposed algorithm can be judged based on various parameters. The parameters that we have chosen in our research work are as follows: Encryption Time taken to encrypt a file of fixed size. The less Encryption time is preferred as the algorithm has to be fast enough to convert the secret message into a hidden message. Decryption Time taken to decrypt a file of fixed size. The less Decryption time is preferred as the algorithm has to be fast enough to convert the hidden message back into the secret message. Time to generate keys Time to generate keys should be as less as possible as keys play the most important role in Encrypting and Decrypting data.

Througput should be as high as possible as it denotes the amount of data to be transferred from the channel.

IX. CONCLUSION

The main objective of this paper is to generate a secure symmetric key, and hence communicate the information over the wired or wireless network medium with the easy and effective way. In this work, a new cryptographic algorithm proposed and implemented using autoencoder artificial neural networks which fits to the required objective, and is efficient and effective to send and receive data securely. Also, for this algorithm low or high bandwidth does not affect the performance. Consequently it’s clear that this approach is easy, fast, and simple, if compared with other algorithms.

REFERENCES


AUTHORS PROFILE

Vikas Sagar received B.tech from MJP Rohilkhand University, Bareilly in 2010 and, Master of Engineering Degree from Punjab Engineering College Chandigarh in 2012. He is research scholar in the Department of Computer Science at Gurukula Kangri Vishwavidyalaya, Haridwar. He has published 7 research papers in several International/National Conference and Journals. His research interests are in Cryptography, Artificial Neural Networks and its applications.

Krishan Kumar is presently working as Assistant Professor in the Department of Computer Science, Faculty of Technology, Gurukula Kangri Vishwavidyalaya, Haridwar, India. He obtained his MCA and Ph.D. in Computer Science and Information Technology from Institute of Engineering and Technology, MJP Rohilkhand University, Bareilly. He has qualifiedUGC-NET in Computer Science and Application. He has published 25 research papers in several International/National Conference/Journals. He is an active member of various societies and also the Chairman of Computer Society of India, Haridwar Chapter. His principal research area is Artificial Neural Networks and its applications.