International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN: 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)
doi : https://doi.org/10.32628/IJSRSET21843

Secure File/Data Transfer Between Airgap Network

Arathi Navaneeth*[1], Vignesh P P[1], Sreehari N R[1], K. Pramilarani[2]
[1]UG Scholar, New Horizon College of Engineering, Bangalore, Karnataka, India
[2] Senior Assistant Professor, New Horizon College of Engineering, Bangalore, Karnataka, India

ABSTRACT

Article Info
Volume 8, Issue 4
Page Number : 149-155

Publication Issue :
July-August-2021

Article History
Accepted : 15 July 2021
Published: 20 July2021

Wireless Data communication is fastest growing technology era in which the research society has recently embarked. Today, Computer data can include financial transactions such as electronic payments, M- wallets and sensitive multimedia contents. The explosive volumes of computer devices personal data, bring-up more attention to securely data storage rather than consideration on data privacy and confidentiality levels. In this scenario Air Gap Data Communication, Machine Leaning (ML) and image processing brings an important role in the electronic data management. It is always expensive and hard to manage the data manually without adopting machine learning and image processing techniques using metadata. The contribution of this research article is to demonstrate a securing computer data storage secrecy and privacy in cloud communication framework in terms of automatic data classification using computer training datasets with help of Training dataset which classifies the data based on the confidentiality level of the record with higher accuracy and powerful timelines as compared to the traditional KNN algorithms and RSA algorithm securing such confidential data category afterwards by applying various existing cryptographic solutions to assuring data privacy, confidentiality levels and alerting the use of abusive contents and simulation results demonstrates that reducing the overall cost. Training dataset which classifies the data based on the confidentiality level of the record with higher accuracy and powerful timelines as compared to the traditional KNN algorithms and RSA algorithm securing such confidential data category.

Keywords : KNN, RSA, MCC, Cloud Computing, Quality of Service Issue

I. INTRODUCTION

Computer Cloud Computing may be a speedy innovation era that involves largest vary of knowledgeable technologies and applications that touches virtually each client through laptop devices. MCC away the restrictions from geographically domains and becomes capable shoppers to urge what they require to try to at anyplace and anytime from net. As a result of laptop devices challenges like low process force, restricted performance, battery life, and lack of quality of service issue (QoS), restricted vitality, shared remote network, and storage capacity. Laptop applications propel computing towards the cloud system as a result of high usage of process force and information storage needs for smartphone supporters. Another important face to that addresses in MCC is that the security and privacy of the pc information storage. MCC provides information sharing facility within the middle of information of knowledge of information operators and laptop shoppers and these data square measure saving in numerous geographical locations. Therefore, such form of laptop information is very fictile to exposing high hazard by means that of confidentiality, integrity, accessibility next to the standard process model. laptop shoppers hesitates from sharing confidential documents to the mediator's storage service suppliers on cloud as a result of obscure nature hands for backup and restore operations. Attackers can have an effect on shopper's sure information as outcome of miser information accession. laptop users hesitates from sharing confidential documents to the mediator's storage service suppliers on cloud because of unknown nature hands for backup and restore operations. Additionally, they're issues regarding their personal information being compromised because of high level attacks against user specific applications and use mechanisms like IDS (Intrusion Detection System)-based tools unfold of cloud storage systems. Compromising and exploiting these touchy information can have serious negative impression on the shoppers being as individual or a company. Therefore, we tend to should wish to forestall such valuable laptop information over the cloud environments.

The existing laptop cloud storage system frameworks utilize security algorithms to encrypted information while not having thought its confidentiality level which could be impossible. Addressing public and hid classified information by the similar fashion and at the equal security level which can hold expendable expense and increasing the time interval. Machine leering is associate degree was application of artificial is intelligence (AI) that has systems the flexibility to mechanically learn and improve from expertise while not being expressly programmed. Machine learning focuses on the event of computer programs that may access information and use it learn for themselves. Data processing is most important applications of Machine Learning. Each instance inside a dataset is developed by machine learning algorithms supported few prophetic options. Machine learning algorithms area unit typically classified as supervised or unattended. Supervised was machine learning algorithms will apply what has been learned within the past to new information victimization tagged examples to predict future events. In distinction, unattended machine with learning algorithms area unit used once the knowledge wont to train is neither classified nor tagged. K-Nearest Neighbor algorithmic program is one amongst the only classification algorithms. Provides keen accuracy and a lot of versatile nature to accumulate the new advancement as compared to different algorithms. K-NN easiness is predicated upon geometer distance and circular function similarity works that area unit typically wont to classify the information. the target of this analysis is to work out the confidentiality based mostly category of {the information the info the information into a file through machine learning algorithmic program and cut back the information coding and cryptography method by applying encoding to just for crucial confidential data. This that proposed framework can increase laptop device potency and reduces the storage quality and period of time when coding and cryptography of the pc information.

II. TECHNOLOGIES USED

- Intended to have an operating model which will help users in air gap network transfer their files and data between different systems.

- This will help in classifying the contents and recognizing the confidentiality of the data which are transferred between these networks.
- Implementing KNN classifier which use data and classify new data points based on similarity measures for the image processing also implementing the RSA algorithm which is used for encryption and decryptions of the data we send between the systems.
- Implementing Image processing for hiding data and information in an image or a picture and transferring it with high security encryption.
- Implementing air gap communication, through an application which will be connected to a network and not connected to an internet which can be Wi-Fi as well as wired connection. Also other USB, CD devices will not be connected.

This application implements the cryptography to send data in an effective way that helps the end users to communicate in a free and secure way without any internet connections.

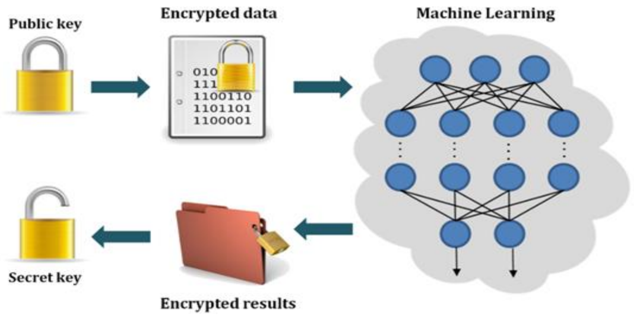


Fig 2.1 Encryption of data

➢ ENCRYPTION AND DECRYPTION

Encryption is that the method of translating plain text knowledge into one thing that is random and mindless. Decipherment is that the method of changing that back to plaintext. The goal of each encoding algorithmic rule is to create it as troublesome as potential to decipher the encrypted knowledge while not mistreatment the key. If a extremely smart encoding algorithmic rule is employed, there's no technique considerably higher than methodically making an attempt each potential ways that. For Associate in algorithmic rule like that, the longer the key, the harder it's to decipher the message while not possessing the key. Here during this project the encoding and decipherment is completed mistreatment the RSA algorithmic rule.

➢ IMAGE PROCESSING

Image process could be a technique to perform some operations in a picture to induce AN increased image or to extract some helpful info from it. It's a kind of signal process wherever the input is a picture and output is also image, any data, info, personal information and confidential information. Nowadays, image process is among chop-chop growing to technologies. It forms core analysis space among engineering and technology branch conjointly. Here the image process is finished mistreatment KNN formula wherever we tend to hide information in pictures and transfer it.

➢ AIRGAP NETWORK

The Air gap data communication and MCC standard allows consumer to approaches and manages their applications data improving the capabilities of moving the storage and compute intensive tasks of computer to the computers. The major safety worries in the computer application data security, client authentication and privacy. The smartphone may be the hot source of location tracking partially as location based services. Because of poor computing capability of computer devices, encryption algorithms with large keys are not feasible to be kept running at the computer. MCC requires the fastest methodology of encryption that requires least storage, processing and communication overhead MCC offers great advantages to their clients for better approachability of data from anywhere by any gadget connected to the internet. Also, it tends to be aptitude for hold support at any time because of accidental damage or loss of data for recuperation purposes.

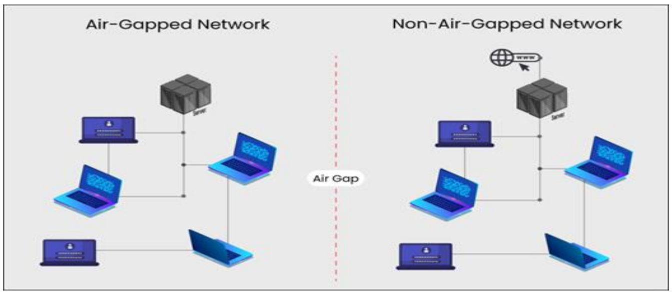


Fig 2.2 Systems in Air Gap network

➢ K- NEAREST NEIGHBOUR ALGORITHM

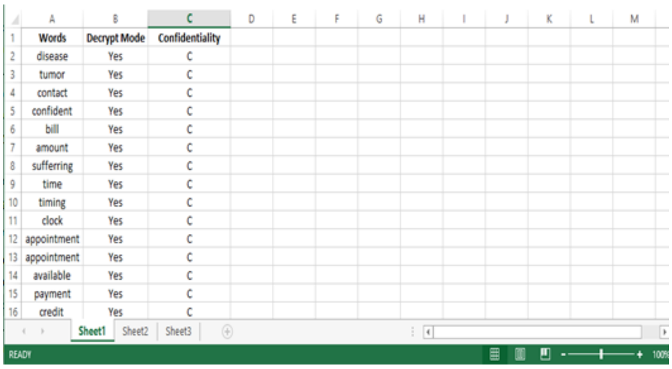
The k-nearest neighbour's algorithm is a technique for classifying objects based on the next training data in the feature s pace. It is among simplest of all mechanism learning algorithms. The algorithm operates on a set of d-dimensional vectors, D = {xi | i = 1. . . N}, where xi ∈ kid denotes the "with data point. The algorithm is initialized by selection k points in kid as the initial k cluster representatives or "centroids". Techniques for select these primary seeds include sampling at random from the dataset, setting them as the solution of clustering a small subset of the data or perturbing the global mean of the data k times. Then the algorithm iterates between two steps till junction. This results in a partitioning of the data. Next is Relocation of "means". Each group representative is relocating to the centre (mean) of all data points assign to it. If the data points come with a possibility measure (Weights), then the relocation is to the expectations (weighted mean) of the data partitions. The k-medic algorithm is similar to k-means except that the centroids have to belong to the data set being clustered.

➢ CONFIDENTIAL AND ABUSIVE DATA

Here the data is classified using the confidential data set. The encryption of the confidential data is done using RSA algorithm. The confidentiality of the data is analyzed and decided according to the percentage of the confidentiality of that particular data.

The user have options like write to file, text file, audio, image and steganography. Through steganography option, the user can hide the data inside an image for a better privacy and security of the information. When the user selects the image option, the image is analyzed and the text inside the image is been converted to text format and the confidentiality of that text is also analyzed.

This is the data set of confidential data as well as non-confidential data which is to be analyzed for the encryption and decryption. This data set have 3 columns they are, words, the column for the data which is analyzed, Decrypt Mode, which says whether the word must be encrypted or not, and the third column, confidentiality, which have two values C and NC , C- Confidential and NC – Non Confidential. Through this data set we are analyzing the confidentiality of the information.



	A	B	C
1	Words	Decrypt Mode	Confidentiality
2	disease	Yes	C
3	tumor	Yes	C
4	contact	Yes	C
5	confident	Yes	C
6	bill	Yes	C
7	amount	Yes	C
8	sufferring	Yes	C
9	time	Yes	C
10	timing	Yes	C
11	clock	Yes	C
12	appointment	Yes	C
13	appointment	Yes	C
14	available	Yes	C
15	payment	Yes	C
16	credit	Yes	C

Fig 3.1 Data set of confidential and Abusive data

We are calculating the percentage of the confidentiality of the word by adding the total confidential words in the message divided by the word length and multiplying it by hundred. By taking the count of the confidential data, we can get the total percentage of the confidentiality. Also the message is compared to the KNN dataset which have the confidential words and abusive words.

➢ ENCRYPTION USING STEGANOGRAPHY

Steganography is a technique use to hide a secret information in such a way that someone unable to find the presence of the information. It is more secured than the method called Cryptography because with cryptography only the scrambling of message is possible, whereas in steganography we use some media to encrypt the data. The main goal of steganography is to hide the information using some covered media. In case of cryptography the user can able to see the contents of message but can't comprehend the information. There are a few niche cases as well, where they are used by researchers or enthusiasts. It is regarded as the ultimate cyber defence strategy for those who absolutely cannot afford a cyber-attack or risk malware infection because the device or network is physically isolated from the outside world.

On the other hand, in steganography the existence of information will not be noticed by viewer because it is embedded inside some medium. It may be an image, video, texts, sound or any music file.

Fig 2.3 Steganography Image

III.RESULTS AND DISCUSSION

An air-gapped network is one that is physically isolated and is not connected to any other network. The only way data can be transferred into an air-gapped network is by physically inserting some sort of removable media, such as a USB or removable disk, or by connecting a transient device like a laptop. Over the years, networks in a variety of verticals, including government, military, financial services, nuclear power plants and industrial manufacturing, have been so-called "air-gapped."

Air gapped systems are typically required in IT ecosystems which manage critical infrastructure that must never fail, or which store confidential information that must never be leaked. There are a few niche cases as well, where they are used by researchers or enthusiasts. It is regarded as the ultimate cyber defence strategy for those who absolutely cannot afford a cyber-attack or risk malware infection because the device or network is physically isolated from the outside world.

Think of it as social distancing to stop computer viruses. Here the data is classified using the confidential data set. The encryption of the confidential data is done using RSA algorithm. The confidentiality of the data is analysed and decided according to the percentage of the confidentiality of that particular data. The user have options like write to file, text file, audio, image and steganography.

Through steganography option, the user can hide the data inside an image for a better privacy and security of the information. When the user selects the image option, the image is analysed and the text inside the image is been converted to text format and the confidentiality of that text is also analysed.

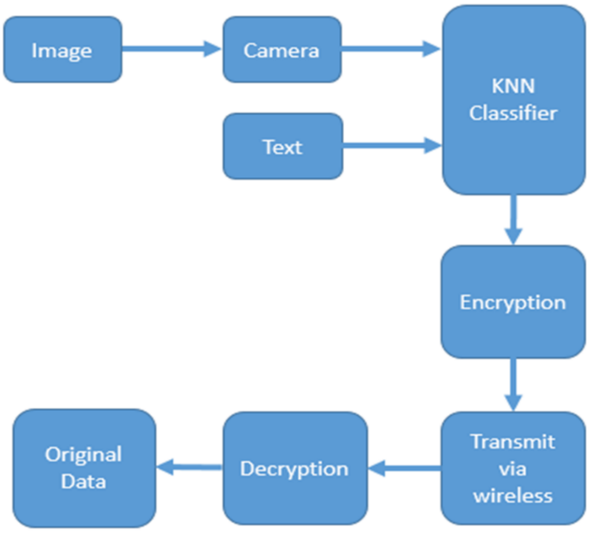


Fig 3.1: Block diagram of the data transmission

The given five options are for users to write text, text file, audio, image and steganography. Using write text option the user can write a text and send, using the text file option, the data in the text file is fetched and send. Same goes with the image and audio option. The steganography option helps the user to hide any data inside an image and send. An encrypted file, which contains a confidential data.

The confidentiality of the data is checked and if it is above 50%, the data is encrypted and send. So here the encrypted file will be decrypted using the RSA algorithm, and the user can get the required confidential data. The user have two options, RSA Algorithm and Steganography. RSA Algorithm helps the user to decrypt the encrypted text files. And the steganography option helps to decrypt the data from a stegno image. After decryption, the decrypted data can be saved as users wish. They can save it as text file, audio, image etc.

The user have options for to decide, how to save their data in their system. The encrypted file can be selected and then using the RSA Algorithm, the user can decrypt. But in case of steganography, the picture exact picture or image which contains the confidential data must be selected. Or else the data which the user needs will not be delivered properly. The user must select appropriate option or else the user will not be able to decrypt the file and get the data. Many must eventually interface with the internet indirectly or send and receive data from internet connected systems where they are more vulnerable to a range of attacks. In this case a computer that's not directly connected to the Internet, with some secure way of moving files on and off.

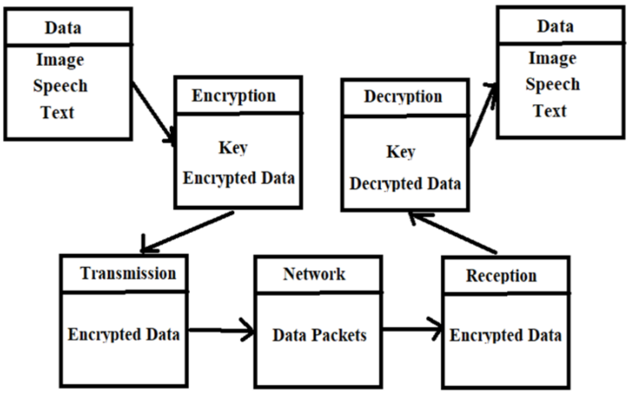


Fig 3.2: Class Diagram of the overall system

IV.CONCLUSION

In this project, we have proposed a way of air gap data communication and a very capable privacy based secure computer data repository model with help of machine classified computer training datasets through TsF-KNN algorithms and public key cryptographic algorithms that decreases the computation time and promises privacy and integrity of critical data categorization. The TsF- KNN is an augmentation of the traditional K-NN algorithm which classified the data attributes into two classes, i.e., confidential and non-confidential with high precision and low computational complexity. The proficiency of our proposed model has been demonstrated by performing simulations results. Modern computing systems are incapable of creating sufficient security protections such that they can be trusted with the most sensitive data while simultaneously being exposed to untrusted data streams. While an air-gapped system can protect data at rest, a completely isolated system or computer often can be of limited value. Many must eventually interface with the internet indirectly or send and receive data from internet connected systems where they are more vulnerable to a range of attacks. In this case a computer that's not directly connected to the Internet, with some secure way of moving files on and off. But every time a file moves back or forth, there's the potential for attack. So a technology should be proposed where the users can easily transfer files and data through an air gap network. The goal of this project is to find a technology for isolating a computer from other networks and making an air gap network where the data is transferred without the help of internet, any hardware devices and wired connections.

V. REFERENCES

[1]. Mordechai Guri, Gabi Kedma and Assaf Kachlon, "AirHopper: Bridging the air-gap between

isolated networks and mobile phones using radio frequencies" in Electronic, IEEE, Oct. 2014, ISBN 978-1-4799-7329-3.

[2]. Anunaya Inani , Manoj Singh Ravish Saxena A Secure Mobile Cloud Computing Framework Based on Data Classification Using Asymmetric Key Cryptography Elsevier SSRN ICToCT 2018

[3]. Lo'ai Tawalbeh1,*, Nour S. Darwazeh2, Raad S. Al-Qassas2 and Fahd AlDosari1 : A Secure Cloud Computing Model based on Data Classification. First International Workshop on Mobile Cloud Computing Systems, Management, and Security (MCSMS-2015.

[4]. Hoang T. Dinh, Chonho Lee, Dusit Niyato* and Ping Wang : A survey of mobile cloud computing: architecture, applications, and approaches School of Computer Engineering, Nanyang Technological University (NTU), Singapore Wireless. Communication. Mobile. Computing. 2013; 13:1587–1611 © 2011 John Wiley & Sons, Ltd. DOI: 10.1002/wcm

[5]. Weiguang SONG, Xiaolong SU, Review of Mobile cloud computing DCST CUMT City XuZhou, JiangSu, China IEEE©2011 978-1-61284-486-2/111

[6]. Ogigau-Neamtiu F. Cloud Computing Security Issues. Journal of Defense Resources Management 2012; 3(2):141-148.

[7]. Mazhar Ali , Samee U. Khan a, Athanasios V. Vasilakos bSecurity in cloud computing: Opportunities and challenges Information Sciences 305 (2015) 357–3830020-0255/ Elsevier2015

[8]. Wu J, Ping L, Ge X, Wang Y, Fu J. Cloud Storage as the Infrastructure of Cloud Computing. International Conference on Intelligent Computing and Cognitive Informatics (ICICCI), 22-23 June 2010; 380-383

[9]. Kotsiantis, S.B., Zaharakis, I.D., Pintelas, P.E., Machine learning: a review of classification and combining techniques, Artif Intell Rev, pp. 159-190 (2006).

[10].Jain, A,K., Murty, M.N., Flynn, P., Data Clustering: a review, ACM Computer Surveys vol. 31, 264-323 (1999).

Cite this article as :

Arathi Navaneeth, Vignesh P P, Sreehari N R, K. Pramilarani, "Secure File/Data Transfer Between Airgap Network", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, Print ISSN : 2395-1990, Volume 8 Issue 4, pp. 149-155, July-August 2021. Available at
doi : https://doi.org/10.32628/IJSRSET21843
Journal URL : https://ijsrset.com/IJSRSET21843