International Journal of Science and Engineering Applications
Volume 11-Issue 06, 72 - 75, 2022, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1106.1001

DCA Method in Mineral Exploration, Example: Predict the Location of New Samples

Hamed Nazerian
University of Catania
Italy

Bahareh Hedayat
Shahrood University of Technology
Iran

Behnam Kakavand

Abstract: As the name implies, Discriminant Correspondence Analysis (DCA) is a format of Discriminant analysis (DA) and correspondence analysis (CA). Like differential analysis, the goal is to categorize observations into predefined groups and, like Correspondence analysis, to use nominal variables. The main idea of the DCA is to represent each set of observations and perform a simple Correspondence analysis on the variables (a matrix). The main observations are complementary elements, and each observation is attributed to the closest group. A comparison between the predictions and predictions of classification leads to evaluating the Discriminant correspondence. This information can be used for a similar case to classify new observations, and the validation of estimates can also be examined using cross-validation techniques such as Jack Knife or Bootstrap. For example, samples were taken from different regions. After scoring the parameters and training in this analysis, a new sample was entered, and a group (region) was determined compared to the previous samples.

Keywords: Discriminant Correspondence Analysis, DCA, CA, DA, Mineral Sampling.

1. INTRODUCTION

In mining engineering, many statistical methods have been adapted from other disciplines. These methods can be used in mining engineering with a simple idea [1-15]. As the name of this method suggests, discriminant analysis is a format of Discriminant analysis and correspondence analysis. Like differential analysis, the goal is to categorize observations into predefined groups and, like Correspondence analysis, to use nominal variables [16-18].

The main idea of the DCA is to represent each set of observations and perform a simple Correspondence analysis on the variables (a matrix). The main observations are complementary elements, and each observation is attributed to the closest group. A comparison between the predictions and predictions of classification leads to evaluating the Discriminant correspondence. This information can be used for a similar case to classify new observations, and the validation of estimates can also be examined using cross-validation techniques such as Jack Knife or Bootstrap [19-35].

2. DISCUSSION

For example, the type of mineral usually depends on its origin. For example, we sampled 12 samples from 3 different regions (4 samples from each region) and asked one person (unaware of their origin) to rate the samples on a 5-point scale.

The scores were then converted to binary code in an index matrix (which can be used in Correspondence analysis). For example, a score of 2 became the binary value of "0 1 0". The data are given in Table 1.

2.1 Subject description

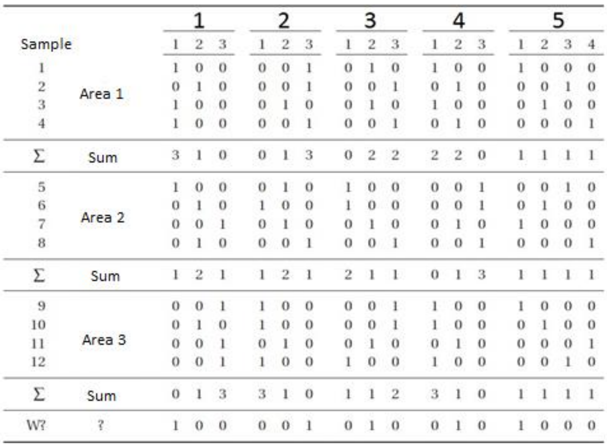
We have K groups; in each group, Ik is observed, and I represent the sum of the observations. To simplify, we assume that our observations are rows, and our variables are columns that contain the J variable, which we call the J × I matrix. The index matrix of K × I is called Y. The value of 1 indicates the belonging of the row that represents the group, and 0 indicates the non-belonging to the group in the columns. The J × K matrix, also named N, is the group matrix, which holds the number of variables for each group. For example, we found that:

$$\mathbf{N}=\mathbf{Y}^{\mathrm{T}}\mathbf{X}=\begin{bmatrix} 3 & 1 & 0 & 0 & 1 & 3 & 0 & 2 & 2 & 2 & 2 & 0 & 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 & 2 & 1 & 2 & 1 & 1 & 0 & 1 & 3 & 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 3 & 1 & 0 & 1 & 1 & 2 & 3 & 1 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (1)$$

2.2 Method

Doing CA in the group N matrix produces two points, one for groups (set F) and one for variables (set G). In general, these factor scores have general values so that their variance is equal to the specific values attributed to the factors. The above table is named N, and the first step in the analysis is to calculate the probability matrix Z = N-1N.

Table 1: Information of 3 regions: 12 samples from 3 different regions in 5 descriptive points. A value of 1 indicates that the sample is variable. The W sample is unknown as a complementary observation.



		1			2			3			4			5			
Sample		1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	4
1	Area 1	1	0	0	0	0	1	0	1	0	1	0	0	1	0	0	0
2		0	1	0	0	0	1	0	0	1	0	1	0	0	0	1	0
3		1	0	0	0	1	0	0	1	0	1	0	0	0	1	0	0
4		1	0	0	0	0	1	0	0	1	0	1	0	0	0	0	1
Σ	Sum	3	1	0	0	1	3	0	2	2	2	2	0	1	1	1	1
5	Area 2	1	0	0	0	1	0	1	0	0	0	0	1	0	0	1	0
6		0	1	0	1	0	0	1	0	0	0	0	1	0	1	0	0
7		0	0	1	0	1	0	0	1	0	0	1	0	1	0	0	0
8		0	1	0	0	0	1	0	0	1	0	0	1	0	0	0	1
Σ	Sum	1	2	1	1	2	1	2	1	1	0	1	3	1	1	1	1
9	Area 3	0	0	1	1	0	0	0	0	1	1	0	0	1	0	0	0
10		0	1	0	1	0	0	0	0	1	1	0	0	0	1	0	0
11		0	0	1	0	1	0	0	1	0	0	1	0	0	0	0	1
12		0	0	1	1	0	0	1	0	0	1	0	0	0	0	1	0
Σ	Sum	0	1	3	3	1	0	1	1	2	3	1	0	1	1	1	1
W?	?	1	0	0	0	0	1	0	1	0	0	1	0	1	0	0	0

r is the sum of the rows Z and the parameter C is the sum of the placed columns, Dc and Dr are the diagonals C and r.

The scoring factor is obtained by breaking down the following single value:

$$D_r^{-1/2}(Z-rc^T)D_c^{-1/2}=P\Delta Q^T \qquad (2)$$

Delta (Δ) is a diagonal matrix with single values, and 2Δ = Λ is a matrix of unique values.

The row and (respectively) column of the score invoice is obtained from the following formula:

$$G=D_c^{-1/2}Q\Delta \qquad \& \qquad F=D_r^{-1/2}P\Delta \qquad (3)$$

The square of the distance from the row and column to the relevant body centers is calculated as follows:

$$d_c=\mathrm{diag}\{GG^T\} \qquad \& \qquad d_r = \mathrm{diag}\{FF^T\} \qquad (4)$$

The square cosine between row i and factor l and column j and factor l is created as follows:

$$o_{i,\ell}=\frac{f_{i,\ell}^2}{d_{r,i}^2} \qquad o_{j,\ell}=\frac{g_{j,\ell}^2}{d_{c,j}^2}, \qquad (5)$$

With $d_{r,I'}^{\,2}$ and $d_{c,j'}^{\,2}$the element i from dr and the element j from d_c are created, respectively. The cosine of squares helps determine the location of important factors to observe. The share of row i to factor l and column j to factor l is obtained as follows:

$$t_{i,\ell}=\frac{f_{i,\ell}^2}{\lambda_\ell} \qquad \& \qquad t_{j,\ell}=\frac{g_{j,\ell}^2}{\lambda_\ell} \qquad (6)$$

Coexistence (or contribution) helps identify the location of essential observations for a factor.

Complementary or rational elements that can represent factors are called transfer formulas.

Specifically, i^T_{sup} as an illustrative row and j_{sup} as an illustrative column. The $f_{sup}$and g_{sup} specifications are obtained through the following:

$$\mathbf{f}_{\text{sup}}=\left(\mathbf{i}_{\text{sup}}^{\mathsf{T}}\mathbf{1}\right)^{-1}\mathbf{i}_{\text{sup}}^{\mathsf{T}}\mathbf{G}\boldsymbol{\Delta}^{-1} \qquad \& \qquad \mathbf{g}_{\text{sup}}=\left(\mathbf{j}_{\text{sup}}^{\mathsf{T}}\mathbf{1}\right)^{-1}\mathbf{j}_{\text{sup}}^{\mathsf{T}}\mathbf{F}\boldsymbol{\Delta}^{-1} \qquad (7)$$

Note that the scalar rules $^{-1)}i^T_{sup}$ (and $^{-1)})j^T_{sup}1$ (are used to ensure the sum of the elements i_{sup} or j_{sup} are equal to one. If this condition is met, there is no need for this law.

After the group's analyses were performed, the main observations were stored as complementary elements and factor scores in a matrix called F_{sup}. To calculate these scores, the first-row profile matrix is calculated:

$$R=(\mathrm{diag}\{X1\})^{-1}X \qquad (8)$$

And apply Equation 7 below:

$$F_{sup}=RG\Delta^{-1} \qquad (9)$$

The Euclidean distance between the observations and the groups calculated from the score factor is equal to the distance ϰ2 between their row profiles.

The K × I distance matrix between observations and groups is calculated as follows:

$$D=S_{sup}1^T+1S^T -2F_{sup}F^T \qquad (10)$$

with

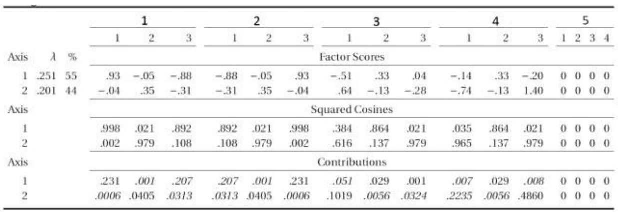
$$S=\mathrm{diag}\{FF^T\} \quad \text{و} \quad S_{sup}=\mathrm{diag}\{F_{sup}F^T_{sup}\} \qquad (11)$$

Each observation will be assigned to the nearest group.

2.2.1. Model evaluation

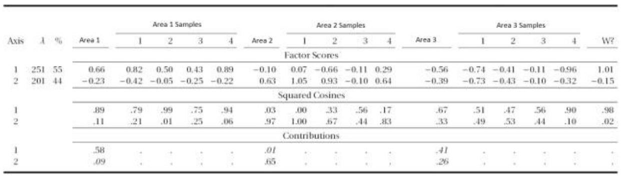
The resolution quality can be considered a model with fixed effects or a random model. For the fixed effect model, the correct classifications are compared with the answers obtained from Equation 10. The fixed-effect model evaluates the classification quality in the sample used to construct the model, and the stochastic model evaluates the classification quality based on the new observations. Typically, this step is done using cross-validation techniques such as Jack Knife or Bootstrap.

Table 2: Factor scores, cosine squares, and auxiliary variables (set J). Negative scores of the correspondence part are also shown in italics.



			1			2			3			4			5			
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	4
Axis	λ	%	Factor Scores															
1	.251	55	.93	−.05	−.88	−.88	−.05	.93	−.51	.33	.04	−.14	.33	−.20	0	0	0	0
2	.201	44	−.04	.35	−.31	−.31	.35	−.04	.64	−.13	−.28	−.74	−.13	1.40	0	0	0	0
Axis			Squared Cosines															
1			.998	.021	.892	.892	.021	.998	.384	.864	.021	.035	.864	.021	0	0	0	0
2			.002	.979	.108	.108	.979	.002	.616	.137	.979	.965	.137	.979	0	0	0	0
Axis			Contributions															
1			.231	*.001*	*.207*	*.207*	*.001*	.231	*.051*	.029	.001	*.007*	.029	*.008*	0	0	0	0
2			*.0006*	.0405	*.0313*	*.0313*	.0405	*.0006*	.1019	*.0056*	*.0324*	*.2235*	*.0056*	.4860	0	0	0	0

Table 3: Invoice scores, cosine squares, region alignment, and complement row for unknown sample region (W). Negative scores of the correspondence part are also shown in italics.



				Area 1 Samples					Area 2 Samples					Area 3 Samples				
Axis	λ	%	Area 1	1	2	3	4	Area 2	1	2	3	4	Area 3	1	2	3	4	W?
			Factor Scores															
1	251	55	0.66	0.82	0.50	0.43	0.89	−0.10	0.07	−0.66	−0.11	0.29	−0.56	−0.74	−0.41	−0.11	−0.96	1.01
2	201	44	−0.23	−0.42	−0.05	−0.25	−0.22	0.63	1.05	0.93	−0.10	0.64	−0.39	−0.73	−0.43	−0.10	−0.32	−0.15
			Squared Cosines															
1			.89	.79	.99	.75	.94	.03	.00	.33	.56	.17	.67	.51	.47	.56	.90	.98
2			.11	.21	.01	.25	.06	.97	1.00	.67	.44	.83	.33	.49	.53	.44	.10	.02
			Contributions															
1			.58	.	.	.	.	*.01*	.	.	.	.	*.41*	.	.	.	.	.
2			*.09*	.	.	.	.	.65	.	.	.	.	*.26*	.	.	.	.	.

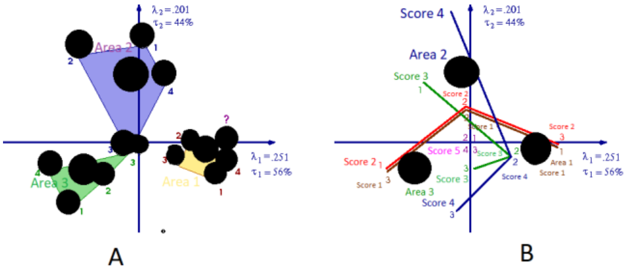


Figure 1: DCA analysis shown in two dimensions. (A) Set I: Rows (samples), sample projected as complementary elements, sample? The sample is unknown. (B) Set J: Column (number of points). Sample categories are also included for ease of interpretation. Both shapes have the same scale (some points shifted slightly for readability). (Tables 2 and 3).

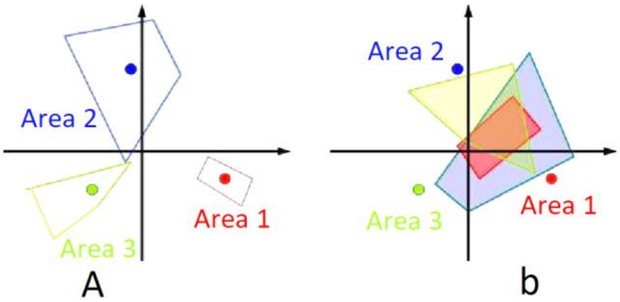


Figure 2: DCA analysis. 2D display. (A) Fixed effect model. Three areas and indicators for samples. (B) Stochastic effect model. Samples of the chipped jack behind the fixed effect solution are shown. The index shows that the stochastic effect classification is more varied and transferred.

3. CONCLUSION

Place Tables 2 and 3 show the analysis results, and Figure 1 shows them. The quality of the fixed effect of the model was evaluated with the following configuration matrix:

$$\begin{bmatrix} 4 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 1 & 4 \end{bmatrix}$$

In this matrix, rows are assigned to predicted groups, and columns are natural groups. For example, out of 5 points assigned to Zone 3 (Group 3), one is from Zone 2 (Group 2), and four samples are from Zone 3. The overall quality can be calculated from the diameter of the matrix. Here we find that (4 + 3 + 4) 11 samples of R12 are correctly classified.

Jack Knife's estimation method was used to evaluate the generalization capacity of the analysis to the new sample (for example, this relates to random effect analysis).

Each sample was placed outside the sample set, a DCA was performed on the remaining samples (11), and the sample was assigned to the nearest group. This method gives us the following configuration matrix:

$$\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

As expected, the performance of the random effect model was lower than the fixed-effect model, and only 6 (2 + 2 + 2) samples out of 12 samples were correctly identified. The difference between the fixed effect model and the stochastic effect is shown in Figure 2 when the data is jacked up (using a multidimensional metric scaling). The quality of the model can be evaluated by drawing the polygon of each category. For the fixed effect model, the centers of gravity are polygons of the categories, indicating that DCA is a least-squares estimation technique. The random model did not function properly due to the more significant variance (it has larger covered polygon areas) and rotated around the area (the polygon was not based on the center of gradation but the center of gravity). Codes of this method were also made available for further research.

4. REFERENCES

1. Khayer, K., et al., *Permeability Estimation from Stoneley Waves in Carbonate Reservoirs.* Geological Bulletin of Turkey, 2022. **65**: p. 42.

2. Khosravi, V., et al., *Hybrid Fuzzy-Analytic Hierarchy Process (AHP) Model for Porphyry Copper Prospecting in Simorgh Area, Eastern Lut Block of Iran.* Mining, 2022. **2**(1): p. 1-12.

3. Nazerian, H., et al., *Design of an Artificial Neural Network (BPNN) to Predict the Content of Silicon Oxide (SiO2) based on the Values of the Rock Main Oxides: Glass Factory Feed Case Study.* International Journal of Science and Engineering Applications (IJSEA), 2022. **2**(11): p. 41-44.

4. Shirazy, A., et al., *K-Means Clustering and General Regression Neural Network Methods for Copper Mineralization probability in Chahar-Farsakh, Iran.* Türkiye Jeoloji Bülteni, 2022. **65**(1): p. 79-92.

5. Hedayat, B., et al., *Feasibility of Simultaneous Application of Fuzzy Neural Network and TOPSIS Integrated Method in Potential Mapping of Lead and Zinc Mineralization in Isfahan-Khomein Metallogeny Zone.* Open Journal of Geology, 2022. **12**(3): p. 215-233.

6. Ahmadi, M.E., et al., *Assessment of the Influence of Sulfuric Acid/Hydrogen Peroxide Mixture on Organic Sulfur Reduction of High Sulfur Coals and Their Chemical Composition.* Open Journal of Geology, 2022. **12**(3): p. 199-214.

7. Aali, A.A., et al., *Geophysical Study to Identify Iron Mineralization Anomalies Using Terrestrial Magnetometry in the Chak-Chak Exploration Area, Iran.* Türkiye Jeoloji Bülteni, 2022. **65**(2): p. 159-170.

8. Nazerian, H., et al., *Design of an Artificial Neural Network (BPNN) to Predict the Content of Silicon Oxide (SiO2) based on the Values of the Rock Main Oxides: Glass Factory Feed Case Study.* International Journal of Science and Engineering Applications (IJSEA), 2022. **2**: p. 41-44.

9. Khayer, K., et al., *Permeability Estimation from Stoneley Waves in Carbonate Reservoirs.* Türkiye Jeoloji Bülteni, 2022. **65**(1): p. 1-8.

10. Adel, S., Z. Mansour, and H. Ardeshir, *Geochemical behavior investigation based on k-means and artificial neural network prediction for titanium and zinc, Kivi region, Iran.* Известия Томского политехнического университета. Инжиниринг георесурсов, 2021. **332**(3): p. 113-125.

11. Shirazy, A., A. Shirazi, and H. Nazerian, *Application of Remote Sensing in Earth Sciences–A Review.* International Journal of Science and Engineering Applications, 2021. **10**(5): p. 45-51.

12. Khayer, K., et al., *Determination of Archie's Tortuosity Factor from Stoneley Waves in Carbonate Reservoirs.* International Journal of Science and Engineering Applications (IJSEA), 2021. **10**: p. 107-110.

13. Shirazy, A., et al., *Geophysical study: Estimation of deposit depth using gravimetric data and Euler method (Jalalabad iron mine, kerman province of IRAN).* Open Journal of Geology, 2021. **11**(8): p. 340-355.

14. Shirazy, A., et al., *Investigation of Geochemical Sections in Exploratory Boreholes of Mesgaran Copper Deposit in Iran.* International Journal for Research in Applied Science and Engineering Technology (IJRASET), 2021. **9**(8): p. 2364-2368.

15. Nazerian, H., et al., *Predict the Amount of Cu Using the Four Ca, Al, P, S Elements by Multiple Linear Regression Method.* International Journal for Research in Applied Science and Engineering Technology (IJRASET), 2021. **9**: p. 1088-1092.

16. Shirazi, A., A. Hezarkhani, and A.B. Pour, *Fusion of Lineament Factor (LF) Map Analysis and Multifractal Technique for Massive Sulfide* Copper *Exploration: The Sahlabad Area, East Iran.* Minerals, 2022. **12**(5): p. 549.

17. Williams, L.J., et al., *A tutorial on multiblock discriminant correspondence analysis (MUDICA): a new method for analyzing discourse data from clinical populations.* 2010.

18. Abdi, H., *Discriminant correspondence analysis*. 2007, Sage Thousand Oaks, CA. p. 1-10.

19. Shirazy, A., et al., *Investigation of Magneto-/Radio-Metric Behavior in Order to Identify an Estimator Model Using K-Means Clustering and Artificial Neural Network (ANN)(Iron Ore Deposit, Yazd, IRAN).* Minerals, 2021. **11**(12): p. 1304.

20. Shirazi, A., et al., *Geochemical and Behavioral Modeling of Phosphorus and Sulfur as Deleterious Elements of Iron Ore to Be Used in Geometallurgical Studies, Sheytoor Iron Ore, Iran.* Open Journal of Geology, 2021. **11**(11): p. 596-620.

21. Shirazi, A. and A. Shirazy, *Introducing Geotourism Attractions in Toroud Village, Semnan Province, IRAN.* International Journal of Science and Engineering Applications, 2020. **9**(16): p. 79-86.

22. Khakmardan, S., et al., *Evaluation of Chromite Recovery from Shaking Table Tailings by Magnetic Separation Method.* Open Journal of Geology, 2020. **10**(12): p. 1153-1163.

23. Doodran, R.J., et al., *Minimalization of Ash from Iranian Gilsonite by Froth Flotation.* Journal of Minerals and Materials Characterization and Engineering, 2020. **9**(1): p. 1-13.

24. Shirazy, A., et al., *Cementation exponent estimate in carbonate reservoirs: A new method.* Global Journal of Computer Sciences: Theory and Research, 2020. **10**(2): p. 66-72.

25. Shirazy, A., A. Shirazi, and A. Hezarkhani, *Behavioral Analysis of Geochemical Elements in Mineral Exploration:-Methodology and Case Study*. 2020: LAP LAMBERT Academic Publishing.

26. Shirazy, A., et al., *Geochemical and geostatistical studies for estimating gold grade in tarq prospect area by k-means clustering method.* Open Journal of Geology, 2019. **9**(6): p. 306-326.

27. Shirazi, A., A. Hezarkhani, and A. Shirazy, *Exploration Geochemistry Data-Application for Cu Anomaly Separation Based On Classical and Modern Statistical Methods in South Khorasan, Iran.* International Journal of Science and Engineering Applications (IJSEA), 2018. **7**(4): p. 39-44.

28. Shirazi, A., A. Hezarkhani, and A. Shirazy, *Remote Sensing Studies for Mapping of Iron Oxide Regions, South of Kerman, IRAN.* International Journal of Science and Engineering Applications (IJSEA), 2018. **7**(4): p. 45-51.

29. Shirazi, A., et al., *Geostatistics studies and geochemical modeling based on core data, sheytoor iron deposit, Iran.* Journal of Geological Resource and Engineering, 2018. **6**: p. 124-133.

30. Alahgholi, S., A. Shirazy, and A. Shirazi, *Geostatistical studies and anomalous elements detection, Bardaskan Area, Iran.* Open Journal of Geology, 2018. **8**(7): p. 697-710.

31. Khakmardan, S., et al., *Copper oxide ore leaching ability and cementation behavior, mesgaran deposit in Iran.* Open Journal of Geology, 2018. **8**(09): p. 841.

32. Shirazi, A., A. Shirazy, and J. Karami, *Remote sensing to identify copper alterations and promising regions, Sarbishe, South Khorasan, Iran.* International Journal of Geology and Earth Sciences, 2018. **4**(2): p. 36-52.

33. Shirazy, A., et al., *Exploratory Remote Sensing Studies to Determine the Mineralization Zones around the Zarshuran Gold Mine.* International Journal of Science and Engineering Applications, 2018. **7**(9): p. 274-279.

34. Shirazi, A., et al., *Introducing a software for innovative neuro-fuzzy clustering method named NFCMR.* Global Journal of Computer Sciences: theory and research, 2018. **8**(2): p. 62-69.

35. Shirazy, A., A. Shirazi, and A. Hezarkhani, *Predicting gold grade in Tarq 1: 100000 geochemical map using the behavior of gold, Arsenic and Antimony by K-means method.* Journal of Mineral Resources Engineering, 2018. **2**(4): p. 11-23.

Securing Cloud Computing Contents with Cryptography and Steganography

Faluyi Bamidele Ibitayo[1*]
Department of Computer Science, Federal Polytechnic Ado-Ekiti, Ekiti State. Nigeria

Oguntuase RianatAbimbola[2*]
Department of Computer Science, Federal Polytechnic Ado-Ekiti, Ekiti State. Nigeria

Makinde Bukola Oyeladun[3*],
Department of Computer Science, Osun State College of Technology, Esa-Oke. Nigeria

Abstract: Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable and reliable computing resources (e.g., networks, servers, storage, applications, services) that can be rapidly provisioned and released with minimal consumer management effort or service provider interaction. This paper is about securing the cloud, which help to protect private information, sensitive data and can enhance the security of communication between client apps and servers. In essence, when your data is encrypted, even if an unauthorized person or entity gains access to it, they will not be able to read it. There will be a secure way to access our data and make it secure by using Cryptography and Steganography. This paper is a details work on combination of cryptography and steganography used in securing data/information. It shows how the simplest methods work and how they can be explored. It uses symmetric encryption algorithm to provide more security.

.**Keywords-Door;** Cloud Cryptography, Steganography, Encryption, Decryption, Cloud.

1. INTRODUCTION

Cloud computing refers to the ability to access and manipulate information stored on remote servers, using any Internet-enabled platform, including smartphones (Getaneh, *et. al*, 2016). Cloud computing is a relatively new technology that will have a great impact on our lives. Using this technology, it is possible to access computing resources and facilities anytime and anywhere. Cloud computing is also known as the cloud. Cloud computing serves a wide range of functions over the Internet like storage. Taking advantage of resource sharing, cloud computing is able to achieve consistency and economies of scale (Rohan and Dhanamma, 2017). Cloud computing does not have a common accepted definition yet. The National Institute of Standards and Technology (NIST) defined five essential characteristics of cloud computing, namely: on-demand self-service, broad network access, resource pooling, rapid elasticity or expansion, and measured service (Khalil, *et. al*, 2014).

In cloud computing data are growing exponentially but security of data is still questionable. Due to the transfer of data to the cloud data center, the security problem occurs and data owner loss their control on data. Security and privacy for cloud data is a major aspect of cloud computing that is still not solved. The major problem of cloud is unencrypted data, unencrypted data can very easily be accessed by unauthorized users like unauthorized internal employees and as well as external hackers. The internal employees can easily access data intentionally or accidently. External hackers may gain access to databases in such environments using hacking techniques like session hijacking and network channel eavesdropping. Virus and Trojan can be uploaded to cloud systems and can cause damage (Jamil and Zaki, 2011). Most effective technique to protect our data is Cryptography and Steganography. The data privacy is also one of the key concerns for Cloud computing.

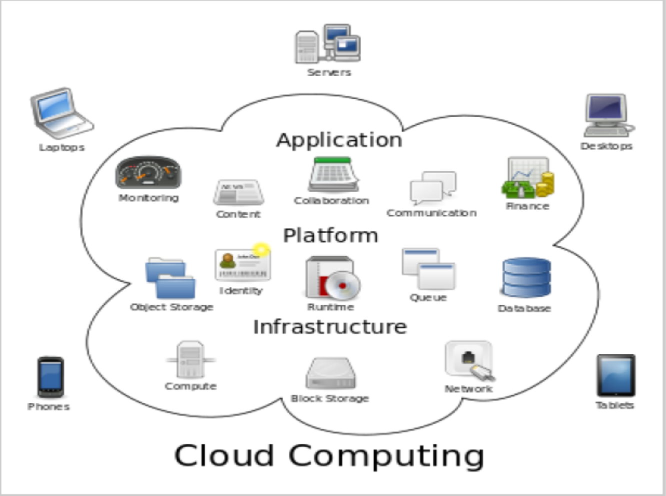


Figure 1.1: Cloud Computing Logical diagram (Guddu, 2019)

2. THEORECTICAL BACKGROUND

Cloud computing evokes different perceptions in different people. To some, it refers to accessing software and storing data in the “cloud” representation of the Internet or a network and using associated services. To others, it is seen as nothing new, but just a modernization of the time-sharing model that was widely employed in the 1960s before the advent of relatively lower-cost computing platforms. These developments eventually evolved to the client/server model and to the personal computer, which placed large amounts of computing power at people’s desktops and spelled the demise of time-sharing systems.

In 1961, John McCarthy, a professor at MIT, presented the idea of computing as a utility much like electricity. (1) Another pioneer, who later developed the basis for the ARPANET, the Department of Defense's Advanced Research Projects Agency Network, and precursor to the Internet, was J.C.R. Licklider. In the 1960s, Licklider promulgated ideas at both ARPA and Bolt, Beranek and Newman (BBN), the high-technology research and development company, that envisioned networked computers at a time when punched card, batch computing was dominant. He stated, "If such a network as I envisage nebulously could be brought into operation, we could have at least four large computers, perhaps six or eight small computers, and a great assortment of disc files and magnetic tape units—not to mention remote consoles and teletype stations—all churning away."

The conjunction of the concepts of utility computing and a ubiquitous worldwide network provided the basis for the future evolution of cloud computing. In an October, 2009 presentation titled "Effectively and Securely Using the Cloud Computing Paradigm,"3 by Peter Mell and Tim Grance of the National Institute of Standards and Technology (NIST) Information Technology Laboratory, cloud computing is defined as follows:

Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable and reliable computing resources (e.g., networks, servers, storage, applications, services) that can be rapidly provisioned and released with minimal consumer management effort or service provider interaction [Ronald and Russell, (2010)].

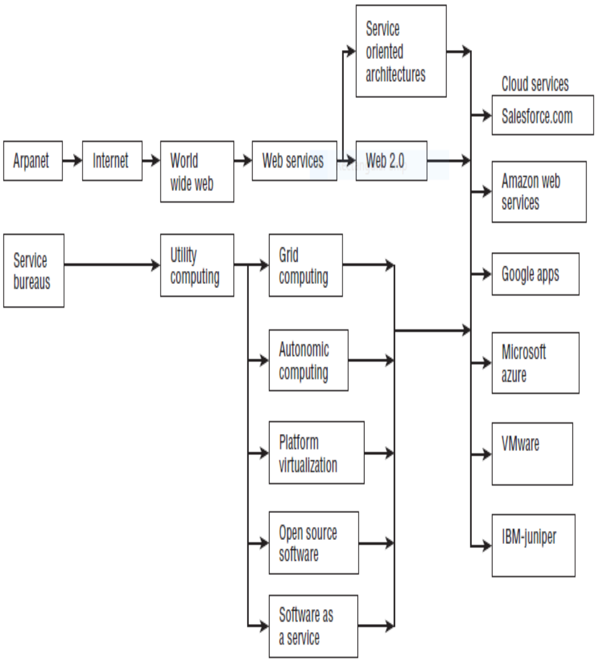


Figure 1: Origins of cloud computing (Ronald and Russell, 2010).

A brief history of cryptography, the earliest form of cryptography was simple writing of a message, as most people could not read. In fact, the very word cryptography comes from Greek words kryptos and graphein, which means hidden and writing, respectively (Tony, 2009). The first known evidence of the use of cryptography (in some form) was found in an inscription carved around 1900BC, in the main chamber of the tomb of the nobleman KhumhotepII, in Egypt. The scribe used some unusual hieroglyphic symbols here and there in place of more ordinary ones. The purpose was not to hide the message but perhaps to change its form in a way which would make it appear dignified. Though the inscription was not a form of secret writing, but incorporated some sort of transformation of original text, and to do so. Evidence of some use of cryptography has been seen in most major early civilizations. "Arthsashtra", a classic work on state craft written by Kautalya, Fast forwarding to around 100 BC, Julius Casar was known to uses a form of encryption to convey secret messages to his army generals posted in the war front. This substitution cipher, known as Caesar cipher and is perhaps the most mentioned historic cipher in academic literature (Huzaifa, 2013).

Steganography has been derived from Greek word "Stego" which means "Covered" and "Graphia" which means "writing". Steganography is an ancient technique of covert communication. Herodotus has mentioned in one of his seminal works of history, Histories during the 400B.C about the tradition of secret writing. He used to tonsure tht head of his most trusted servants with tattooed the scalps with secret message and waited for the hair to grow. The servant used to travel between the borders without carrying anything contentious freely. At the reception end his head would be tonsured again and the message will be conveyed. Another example of steganography is during the Vietman War were the captures US armed force showed hand gestures during a photo sessions to convey some military secrets. The field of Steganography is limitless and any kind of cover media can be text, image (grey, binary, color), audio, video etc (UKEssays, 2018).

This project is about securing the cloud, which help to protect private information, sensitive data and can enhance the security of communication between client apps and servers. In essence, when your data is encrypted, even if an unauthorized person or entity gains access to it, they will not be able to read it. There will be a secure way to access our data and make it secure by using Cryptography and Steganography.

3. LITERATURE REVIEW

Cloud Computing and Security Issues despite the hype about the cloud, customers are reluctant to deploy their business in the cloud. Security issues is one of the biggest concerns that has been affecting the growth of cloud computing. It adds complications with data privacy and data protection continues to affect the market. Users need to understand the risk of data breaches in the cloud environment. The paper highlights issues related to cloud computing (Rohan, 2017). Enhancing Cloud Computing Security using Cryptography & Steganography to increase the security of data in data centres of cloud to ensure data security in cloud computing by encoding secret data using two levels of encryption are DES & RSA algorithm and to enhancing the security we use LSB algorithm to hide these encrypted data inside edges of colour images which is called steganography (Dheyah, 2019). Hybrid cryptography and steganography method to embed encrypted text message within image cryptography still has many drawbacks such as stole and decrypts the original texts using automatic decryption counter. The main aim of this research is to improve the cryptography securing level using supportive method which is Steganography. The Steganography is the processes of hide the data or information in media files such as video, images and audio files (Khider *at. al.,* 2019). The need for data security techniques has facilitated the evolution of different techniques necessary for ensuring the safety of data and information through the platform hence ensuring effective communication channels. Therefore, Steganography techniques have been identified as one of the cloud computing techniques necessary for enhancing cloud data security. In return, this paper analyzes the application of stenography technique in cloud computing as an approach to enhance cloud data security (Jacob *at. al* 2019).

Advanced Encryption Standard (AES) combined with Bit-Level Embedding for securing cloud data, in this paper, a combination of these to provide a high level of security to cloud data, AES as symmetric algorithm in cryptography combined with Bit-Level embedding as text Steganography to serve as a cover medium considering that AES works on both software and hardware while text steganography has less data redundant (Ishaq *at. al* 2021). Securing cloud data using blowfish algorithm combined with text steganography, in this paper a combination of these algorithms to provide a high level of security to cloud data, Blowfish symmetric algorithm as cryptography combined with text steganography as a cover medium considering that both have proven to have good performance and less data redundant in literature (Ishaq *at. al.,* (2020).

3.1 HISTORY OF CLOUD COMPUTING

Cloud computing evokes different perceptions in different people. To some, it refers to accessing software and storing data in the “cloud” representation of the Internet or a network and using associated services. To others, it is seen as nothing new, but just a modernization of the time-sharing model that was widely employed in the 1960s before the advent of relatively lower-cost computing platforms. These developments eventually evolved to the client/server model and to the personal computer, which placed large amounts of computing power at people’s desktops and spelled the demise of time-sharing systems.

In 1961, John McCarthy, a professor at MIT, presented the idea of computing as a utility much like electricity. Another pioneer, who later developed the basis for the ARPANET, the Department of Defense’s Advanced Research Projects Agency Network, and precursor to the Internet, was J.C.R. Licklider. In the 1960s, Licklider promulgated ideas at both ARPA and Bolt, Beranek and Newman (BBN), the high-technology research and development company, that envisioned networked computers at a time when punched card, batch computing was dominant. He stated, “If such a network as I envisage nebulously could be brought into operation, we could have at least four large computers, perhaps six or eight small computers, and a great assortment of disc files and magnetic tape units not to mention remote consoles and teletype stations all churning away.

The conjunction of the concepts of utility computing and a ubiquitous worldwide network provided the basis for the future evolution of cloud computing (Ronald and Russel, 2010).

3.2 CLOUD COMPUTING

Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable and reliable computing resources (e.g., networks, servers, storage, applications, services) that can be rapidly provisioned and released with minimal consumer management effort or service provider interaction (Peter and Tim, 2002).

Software Developers describe Cloud in a different way than a System Administrator, while a Database Administrator may have different definition. Cloud means a wide range of scalable services that users can access via an Internet connection. Providers like Microsoft, Amazon, Google and many more provide various cloud-based services for which users can pay on the basis of service subscription and consumption. Many providers offer a wide range of Cloud services like Messaging, Social Computing, Storage, CRM, Identity management, Content Management etc (Rohan and Dhanamma, 2017).

Cloud Computing Characteristics

According to the official definition, cloud computing has five main characteristics: resource pooling, broad network access, rapid elasticity, on-demand self-service, and measured service.

i. Shared resources: clients can share resources like networks, servers, storage, software, memory, and processing simultaneously. Providers can dynamically allocate resources according to the fluctuations in demand, and the client is completely unaware of the physical locations of these services.
ii. Broad network access: the cloud allows a broad access to the network using the Internet from any device.
iii. Elasticity: the cloud is flexible and configurable. Clients feel that resources are unlimited.
iv. On-demand self-service: if needed, any customer can automatically configure the cloud without the interference of service technicians. Customers perform scheduling and decides the required storage and computing power.
v. Measured service: different cloud services can be measured using different metrics. Detailed usage reports are generated to preserve the rights of customers and providers (Yazan *at. al*., 2019).

Advantages of cloud computing

Cloud computing offers the following major advantages to the users.

- The 3rd party provider owns and manages all the computing resources (servers, software, storage, and networking) and electricity needed for the services. The users only need to "plug into" the cloud. The users do not need to make a large upfront investment on computing resources; the space needed to house them; electricity needed to run the computing resources; and the cost of maintaining staff for administering the system, network, and database.
- The users can increase or decrease the level of use of the computing resources and services flexibly and easily.
- The users pay most likely much less for the services, because they pay only for the computing resources and services they use, and the subscription-based or payper-use charges are likely much lower than the cost of maintaining on-premises computing resources. If the users are to maintain on-premises computing resources, they

also need to make the worst-case plan to account for the occasional or seasonal peak needs.

- The users can in practice access the cloud for services anytime from anywhere (Getaneh *at. al.*, 2016).

Types of Cloud

Service Models. Cloud computing has four different service models:

(i) Software as a service (SaaS): it is the most popular cloud service, and the software resides on the provider platform. -e consumer can access the software using a web browser or an application programming interface (API). It follows a pay-peruse business model. Consumers do not need to worry about the software upgrades and maintenance; some limited application configuration capability might be available to consumers. Salesforce and Office 365 are popular examples.

(ii) Platform as a service (PaaS): it provides development and testing environments. -e consumer develops his/her own application on a virtual server and has some control over the application hosting environment, particularly the application and data, making it faster to develop, test, and deploy applications. Cloud Foundry is a good example.

(iii) Infrastructure as a service (IaaS): it provides the infrastructure, operating systems, and applications. It is the service of choice for companies that do not have the necessary capital to buy hardware. Customers pay according to consumption.

(iv). Infrastructure is scalable depending on processing and storage needs. -e consumer has control over applications, data, middleware, and operating systems but not over the underlying cloud infrastructure. Amazon EC2 is a good.

(v) Anything as a service (XaaS): it offers a variety of services ranging from personal services to large resources over the Internet (Yazan *at. al.*, 2019).

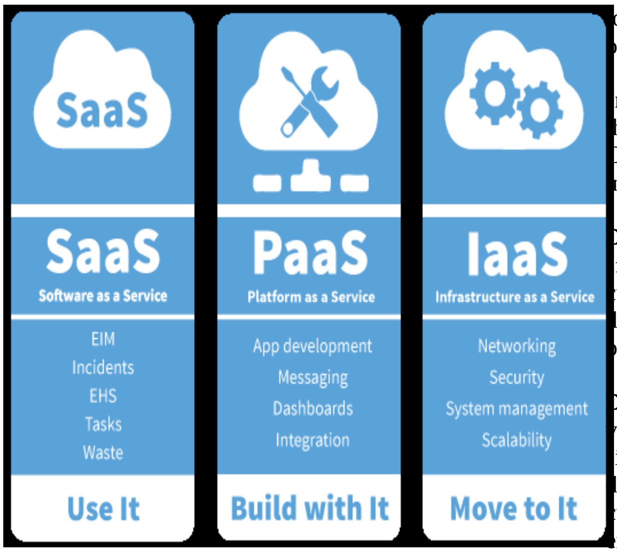


Figure 2.1: Cloud Computing Service Model (Rohan and Dhanamma,, 2017)

Delivery Models. Cloud computing has five different delivery models:

(i) **Private cloud**: it is located on premises, over the intranet, behind the firewall, and usually managed by the same organization that uses it. -eir services are offered to the organization employees. Security issues are limited; a good example is VMware.

(ii) **Public cloud**: it is located off premises, over the Internet, and usually managed by a cloud service provider. -the services are offered to the public. It is less secure than the private cloud, some popular public clouds are Dropbox Amazon EC2, and Microsoft Azure.

(iii) **Hybrid cloud**: it combines private and public clouds, and it has trust and confidentiality issues because of the public part. A good example is Rackspace.

(iv) **Community cloud**: it is a group of entities with a common goal, share the cloud; universities usually share a single cloud. A good example is NYSE Capital Markets Community Platform (Yazan *at. al.*, 2019).

3.3 THREATS IN A CLOUD ENVIRONMENT

In cloud environment threats are experienced at two levels i.e. at the providers' level and at the clients' level. Providers face threats that may occur at the provider side. Threats could be like Physical Security, Disgruntle employees, a natural disaster for data centre and etc. Many of the threats faced by Providers are described in the next section (Avizienis *at. al.* 2020). On the other hand, clients have their own threats that may occur because of their own or provider (Chen and Zhao, 2012). The client may lose their data completely or partially. Their data may get corrupted. If anyhow client attacker gets user id and password, they may modify the data or completely delete the data. If the attacker accesses the cloud account, they may increase the bill for cloud usage. This section explores various threats that can occur in a cloud environment.

Physical Security: Physical security concerns the IT hardware of the data center. The IT hardware for the data center is important as if it is damaged anyhow, all data will be ost. The security is against any intruder, flood attack or any ther natural disaster which may cause harm to data.

nformation Security/Data Loss/Leakage: This concerns he level of security of data given by a service provider. Confidentiality, Integrity and Availability (CIA) is included in nformation security (Khoshkholghi, *at. al.* 2014).

Data Location: Data location is important as, in illegal ituations; the respective country's laws are applicable. If a ustomer has sensitive data, then the customer must go for a ata centre in their country so as to use the same country legal rocess.

Data Segregation: A poor data segregation may give rise to ulnerability as there can be many different customers on a ingle cloud. Thus, an attacker may get easy success to steal ata. To get high segregation, complete isolation from another ustomer (i.e. private cloud) can be employed. Further, strong ncryption can be applied to the data, however it can be expensive.

Data Recovery: Data recovery is vital considering the loss of data in cases such as manmade disaster or natural disaster. However, it poses the challenge in maintaining the regular back-ups on the same machine where data is located as damage of the machine results in loss of data as well.

Secure Data Transfer: Data transfer must be secure as the data can be theft on the network itself. If any leak occurs, it may trouble the customer. Customer must use cloud where security is also available on the network, like the one who uses https which works on SSL, not on Http which gives unreliable connection.

User Access Control: It is essential for a cloud service provider to incorporate the best access management system. Poor management may mix multiple customer data which may again trouble the customer. Authentication level must be increased in case of multiple customers such as, username and fingerprint whitelisting of IP address for remote actions.

Cloud Portability: Cloud portability is necessary considering the desire of customer to move to another service provider. However, this may arise the situation of LOCK-IN, where portability is not possible because of lack of some standard data formats, procedure or tools that aids in maintaining the portability of application, data, or service.

Denial of Service (DOS): As there are multiple customers, the data center must be that much efficient enough to handle the maximum requests, otherwise it can face Denial of Service (DOS). Further, there can be DOS attack for which Servers must have algorithms to handle.

3.4 MITIGATION STRATEGIES: This section suggests various mitigation strategies for the various attacks in cloud environment.

Disgruntled Employees: Cloud Service providers should make a check on disgruntled employees who may try to leak secret cloud information or may harm the data centre in any way and etc.

Monitor Data Center: Data centers must be monitored always as these are the key places where all data is stored. Monitoring must include filtering from any suspicious packet. Data center must be monitored to make a check from any environmental threat too.

Ensuring CIA: Confidentiality, Integrity and Authority mitigate the threat as it includes the security information. Further, Data Isolation must be ensured at the provider end.

Service Level Agreement: Service Level Agreement may help to mitigate the threat level. In SLA, the client may tell to include their own clauses that can help in future to deal with any situation. SLA's must be properly studied by both the parties, this will for sure decrease the threat level.

Find best cloud provider: There are many cloud service provides (CSPs) in the market, if any client wants to hire any cloud they must ensure how there CSP must be and analyze which one would be the best in terms of experience, standards, regulation.

Use HTTPS (SSL) instead of HTTP (Plaint text): CSPs must use secured protocols for networking, as they protect data from any intruder, which may try to read or corrupt data. The secured protocol encrypts the data too which is today's demand for data protection.

Intrusion Detection: CSPs' must use algorithms for Intrusion Prevention, Intrusion Detection System. These algorithms may secure the cloud is a much better way which will, in turn, benefit the clients' data too.

Security Testing: To mitigate the effect of security loopholes on cloud, security testing applicable on cloud must be done. These testing will build a rigid cloud system.

Encryption: Encryption Techniques must be applied at each level of storage whether it is back up space or real data location. Encryption Techniques must also be present at network levels as well.

Standard API: This is a major problem that client face if any client using one cloud of any company and want to switch to other and lock-in problem occurs. If the API of one cloud is not compatible with another one to switch with then this lock-in problem occurs. Mitigation to this problem is to use of standard API to build the Cloud (Himanshi, *at. al.,* 2020).

3.5 DATA SECURITY

Data security is the practice of protecting digital information from unauthorized access, corruption, or theft throughout its entire lifecycle. It's a concept that encompasses every aspect of information security from the physical security of hardware and storage devices to administrative and access controls, as well as the logical security of software applications.

Data Security Capabilities and Solutions

Data security and tools and technologies should address the growing challenges inherent in securing today's complex, distributed, hybrid, and/or multi-cloud computing environments. We have vulnerability assessment and risk analysis tools, automated compliance reporting, data discovery and classification tools and data and file activity monitoring.

TYPE OF DATA SECURITY

- Encryption
- Data Erasure
- Data Masking
- Data Resiliency

ENCRYPTION

Data encryption is a process that helps to solve various external and malicious threats. Unencrypted data is very vulnerable for susceptible data, as it does not provide any security mechanism. Unencrypted data can very easily be accessed by unauthorized users. Unencrypted data risks the user data which leads to cloud server to escape various data information to unauthorized users. For example, the famous file sharing service Drop box was accused for using a single encryption key for all user data the company stored. These unencrypted, insecure data encourage the malicious users to misuse the data one or the other way (Rohan and Dhanamma 2017).

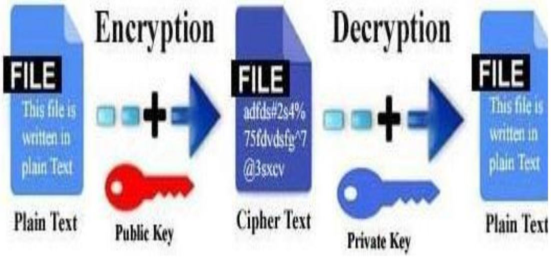

Encryption
Decryption
FILE
This file is written in plain Text
Plain Text
Public Key
FILE
adfds#2s4% 75fdvdsfg^7 @3sxcv
Cipher Text
Private Key
FILE
This file is written in plain Text
Plain Text

Figure 2.2: Encryption and Decryption (Jaydip, 2019)

3.6 CRYPTOGRAPHY

Cryptography is the practice and study of techniques for securing communication and data in the presence of opponents. How cryptography can help you secure your messages. So, to protect your text, first, you have to convert your readable message to an unreadable form. Here, you change the message to a bunch of random numbers. After that, you encrypt it with a key in Cryptography we call this **cipher text**.

TYPES OF CRYPTOGRAPHY

Cryptography is broadly classified into two categories: Symmetric key Cryptography and Asymmetric key Cryptography (popularly known as public-key cryptography). Now Symmetric-key Cryptography is further categorized as Classical Cryptography and Modern Cryptography. The entire Cryptography type breakdown looks something like the image below, based on different use of keys and encryptions.

1. **Symmetric Cryptography:** With symmetric encryption, normal readable data, known as plain text, is encoded (encrypted), so that it is unreadable. The sender scrambles the message using a key, the receiver has to decode the message using the same key used for encoding it. Thus, this key known as a “secret key” is the most important part of symmetric encryption since anyone who has access to it can decrypt private data. Examples of symmetric key algorithm are DES (digital encryption standard), AES (advanced encryption standard), Twofish, Blowfish, 3DES etc.
    a. Transposition Ciphers: This is a method of encryption in which the letters of the plaintext are systematically rearranged into another sequence. So that the ciphertext constitutes a transformation of the plaintext. That is, the order of the units is changed (the plaintext is reordered). Mathematically, a bijective function is used on the characters’ positions to encrypt and an inverse function to decrypt.
    b. Substitution Cipher: Method of encryption by which units of plaintext are replaced with ciphertext, according to a fixed system; the “units” may be single letters (the most common), pairs of letters, triplets of letters, mixtures of the above, and so forth.
2. **Asymmetric Key Encryption (or Public Key Cryptography):** The asymmetric encryption key works by encoding the transmitted messages. However, instead of using the same key, it uses a different one to decrypt the message. Examples of Asymmetric are RSA (ravest-shamir-Addeman)
   The key for encryption is available to all users of the network. As such, it is known as a public key. On the other hand, the key used for decryption is kept secret and only used privately by the user. Hence it is known as a private key.

   This Modern Cryptography is divided into Stream Cipher and Block Cipher.
   - Stream Cipher: Symmetric or secret-key encryption algorithm that encrypts a single bit at a time. With a Stream Cipher, the same plaintext bit or byte will encrypt to a different bit or byte every time it is encrypted.
   - Block Cipher: An encryption method that applies a deterministic algorithm along with a symmetric key to encrypt a block of text, rather than encrypting one bit at a time as in stream ciphers.

AES Algorithm: The Advanced Encryption Standard, AES, is a symmetric encryption algorithm and one of the most secure. This method uses a block cipher, which encrypts data one fixed-size block at a time, unlike other types of encryption, such as stream ciphers, which encrypt data bit by bit.

AES is comprised of AES-128, AES-192 and AES-256. The key bit you choose encrypts and decrypts blocks in 128 bits, 192 bits and so on. There are different rounds for each bit key. A round is the process of turning plaintext into cipher text. For 128-bit, there are 10 rounds; 192-bit has 12 rounds; and 256-bit has 14 rounds.

Since AES is a symmetric key encryption, you must share the key with other individuals for them to access the encrypted data. Furthermore, if you don’t have a secure way to share that key and unauthorized individuals gain access to it, they can decrypt everything encrypted with that specific key (Yang *et. al.*, 2018).

Blowfish Algorithm: -Blowfish is a variable-length, symmetric, 64-bit block cipher. Designed by Bruce Schneier in 1993 as a “general-purpose algorithm,” it was intended to provide a fast, free, drop-in alternative to the aging Data Encryption Standard (DES) and International Data Encryption Algorithm (IDEA) encryption algorithms. Blowfish features a 64-bit block size and takes a variable-length key, from 32 bits to 448 bits. It consists of 16 Feistel-like iteration operates on a 64-bit block that’s split into two 32-bit words. Blowfish uses a single encryption key to both encrypt and decrypt data (Isaq *at. al.*, 2020).

3.7 STEGANOGRAPHY

Steganography is the technique of hidden communication. Using steganography a secret message is embedded in a medium, such as an image or a sound clip and sent. The existence of the hidden message is not known except by the sender and receiver. The word is derived from the greek words stegos meaning covered and graphia meaning writing.

How it works: The "cover" is the medium which is used to hide the secret information. The information to be hidden can be a plain text message, a cipher text, another image, or anything that can be represented in binary.

Cover media can be a lot of things: text, images, audio and video. As images are the most commonly used medium, let's look at that closer [Isaq *at. al.* (2020).

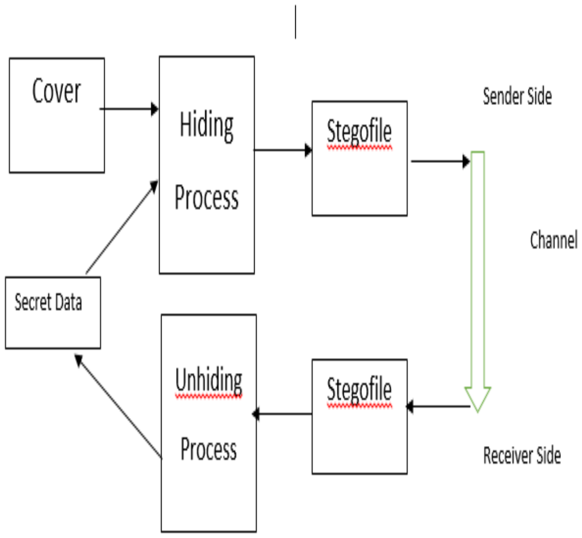


Figure 2.2: Steganography Process (Dheyah, 2019)

Common methods of concealing data in digital images include:

Least significant bit (LSB) insertion, Masking and filtering and Transformations

Usage: Steganography is a powerful tool that enables people to communicate without eavesdroppers even being aware that the communication is taking place. It can even be combined with cryptography so that the message is not just hidden but also scrambled so that even on being discovered it cannot be read.

LSB (Least significant bit insertion): This is a very popular method because of its simplicity. In this method, the LSB of each byte in the image is used to store the secret data. The resulting changes are too small to be recognized by the human eye. The disadvantage of this technique is that since it uses each pixel in an image, a lossless compression format like bmp or gif has to be used for the image. If lossy compression is used, some of the hidden information might be lost.

Tools: S-Tools and EzStego are tools that use LSB method for hiding information. These tools, in addition to hiding the information in the LSBs, also do some additional processing to make the hiding less detectable. For example, the EzStego tool arranges the palette to reduce the occurrence of adjacent index colors that contrast too much before it inserts the message. This ensures that there is not too much change in the color of the pixel once the LSB is modified.

4. METHODOLOGY

4.1 ANALYSIS OF THE EXISTING METHOD

The present method used in encrypting and decrypting of data/information is the most method used in some organisations to secure or guard against unauthorized or intercessor from attempting the data in the cloud. Data in cloud models can be easily accessed by unauthorized internal employees, as well as external hackers. The internal employees can easily access data intentionally or accidently. External hackers may gain access to databases in such environments using hacking techniques like session hijacking and network channel eavesdropping. Virus and Trojan can be uploaded to cloud systems and can cause damage thereby constitute problems to the main data in the cloud. The existing method pose major problem due to the slow method adopted when encrypting the data and fast when decrypting data and the system is to encrypt the text file only. When dealing with large data to be encrypted the algorithm is so slow that it takes much time before the encryption is completed therefore, the method DES and RSA is not recommendable for large data encryption. This lead to the new adoption of cryptography and steganography algorithm for security that allow data to be converted into cypher text through the use of AES then Blowfish and later hide the cypher text into an image called cover image.

4.2 PROBLEM OF THE EXISTING SYSTEM

Many algorithms have been propagated and written by some scientists, still and still, major problem associated with these algorithms are as follows:

i. Time consuming in generating a cipher text during Code Block Cipher (CBC) conversion.
ii. Easily broke single level encryption
iii. One to one encryption model when dealing with text encryption.
iv. Less secure of personal credentials due to the lack of constant bits generating for cipher text.

4.3 ANALYSIS OF THE DEVELOPED SYSTEM

Cryptography model is the act of securing communications techniques that allow only the sender and intended recipient of the information to view its contents. Here, plain-file is encrypting using a secret key, and then both the encoded plain-file and secret key are sent to the recipient for decryption and read the plain-file using cryptography algorithms. In this type of algorithm used in cipher the plain-file to be conveyed, AES is the Advanced Encryption Standard for cryptography which is used to encrypt data to keep it private. It is popular cypher used for many purposes, use to perform the first level of encryption in this system. AES is a symmetric, block cipher which means that blocks of file of a certain size (256 bits) are encrypted one at a time. And also using Blowfish algorithm for the second level encryption in this system, Blowfish is an encryption algorithm, or cipher, specifically a block cipher, it convert cipher text to plain-file and back. The Block Cipher is also known as Code Block Cipher which is highly secure and it is generated into series of binary code.

Steganography can hide any file into an image file; here any file can be used as a secret file and LSB Algorithm is the most common method used in steganography. And the colour images are the best cover which is used in steganography. The system first takes an image file as carrier file, and encrypt the image that cannot be able to play until it decrypt. It uses Advanced Encryption Standard (AES) and

Blowfish algorithms to encrypt secret file before hiding image.

Advanced Encryption Standard (AES) became effective as a federal government standard in 2002. And In June 2003, U.S. government announced that AES could be used to protect classified information and Blowfish is a variable-length, symmetric, 64-bit block cipher. Designed by Bruce Schneier in 1993 as a "general-purpose algorithm," it was intended to provide a fast, free, drop-in alternative to the aging Data Encryption Standard (DES) and International Data Encryption Algorithm (IDEA) encryption algorithms. The Steganography scheme is a symmetric block cipher that encrypts and decrypts 128-bits blocks of data using LSB of each byte in the image is used to store the secret data. The resulting changes are too small to be recognized by the human eye. In AES and Blowfish algorithms, plain-file refers to the data to be encrypted. Ciphertext refers to the data after going through the cipher as well as the data that will be going into the decipher.

Data/information is very important in our day to day life, when dealing with data, people believe that their personal information must be secure and private, in Private Cloud Storage significant data, files and records are entrusted to a third party, which enables Data Security to become the main security issue in cloud computing. Due to the confidentiality of data, organization must put in place a way of safeguarding data and record of their employees.

Personal or Organizational record in the cloud could not be secure without the use of cryptography and steganography algorithm which is used in securing and protecting data into image from intruders and intercessors in the cloud. Data is encrypted as a plain file and later converted into series of byte called cipher text which is unreadable to human. This cipher text is generated by the AES and Blowfish algorithm and the text can be seen by the intercessor but they could not be able to decrypt it due to the lack of encryption key used when encrypting the information to be conveyed. The converted text will later be hidden into an image file for further usage.

4.4 ADVANTAGES OF THE DEVELOPED SYSTEM

The new method proposed have many advantages over the existing method used in encrypting and decrypting data and information. Some of the benefits are:

i. Fast encrypting and decrypting of data
ii. Less complex algorithms
iii. Large size of bytes generated during encryptions.
iv. Very high secure standard of data used due to two ways data encryption called Cryto-Stego algorithm.
v. Multilevel method used in encrypting and decrypting files.

4.5 JUSTIFICATION OF THE DEVELOPED SYSTEM

The proposed system will work in order to eradicate the associated problems occurs in the existing method of securing data in the cloud. The system will work perfectly and efficiently as expected by any organization or users that adopted it as their method of encrypting and decrypting organization or user data in the cloud. The system is not time consuming and it is less complex in terms of algorithm and training of users.

4.6 ARCHITECTURE OF THE DEVELOPED SYSTEM

Figure 3. Describes the architecture of the proposed tool's implement. The modules of the steganography tool are also included in the architecture. The user can be either the sender or the receiver. The cover file in below architecture is image. The in-depth flow of this architecture is described in the design approach.

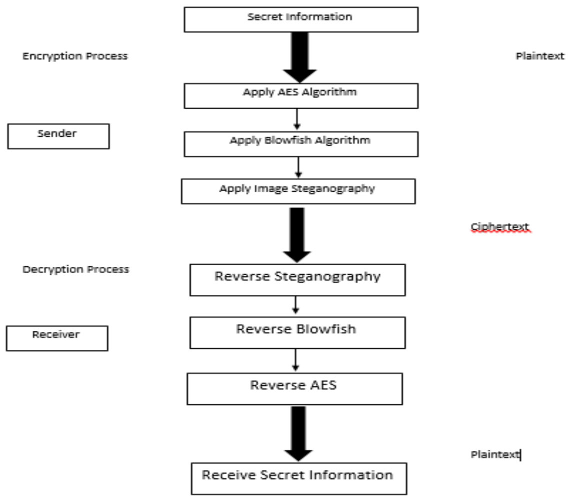


Figure 3: Developed architecture of Overall system

4.7 DESIGN OF THE CURRENT TOOL

S-Tools and EzStego are tools that use LSB method for hiding information. These tools, in addition to hiding the information in the LSBs, also do some additional processing to make the hiding less detectable. For example, the EzStego tool arranges the palette to reduce the occurrence of adjacent index colors that contrast too much before it inserts the message. This ensures that there is not too much change in the color of the pixel once the LSB is modified.

It is important to note that the convert file size can be no greater than one eight of the size of the JPEG Image file. Steganography does not even touch the header part of the JPEG Image file, the properties remain intact no matter how important as Steganography create new JPEG Image files that cannot be distinguished from the original JPEG Image file. The two important distinguished in Steganography tool are encrypt and hide process and decrypt and retrieve process.

4.8. ENCRYPT AND HIDE PROCESS

Encryption and hiding of the covert into a JPEG file is achieved in steganography in following steps. All the following steps have been implemented correctly to achieve perfect JPEG image steganography.

The entire process can be broken down into few important steps.

(i) Convert the JPEG file and convert file into binary representation.
(ii) Encrypt the convert file using AES algorithm.
(iii) Encrypt the AES encrypted file using Blowfish algorithm.
(iv) Create a random unique number from password given.
(v) Use LSB substitution to replace JPEG file bits with encrypted data bits.
(vi) After substitution is done, create new crypto-stego JPEG file.

4.9. CREATING NEW STEGO FILE

All the above steps are used in creating a new stego JPEG file. As explained above, the header part of the new stego JPEG image file will be exactly same as the original JPEG image file. The remaining data part after replacement of least significant bits is written into the data chunk part of the JPEG file.

5. IMPLEMENTATION AND TESTING

This section explains various choices made while carrying out the implementation and the programming language employed. The section also contains the discussion on experiment as well as result obtained and analysis.

5.1 PROGRAMMING LANGUAGE USED

Python is the programming language of choice. The facts presented here, adopted from the work by Microsoft & Associates (2011), underlie our decision.

Key Benefits of Python

Why use Python at all? Python is an open source scripting language which is high-level, interpreted, interactive and object-oriented. It is designed to be highly readable. The syntax of Python language is easy to understand and uses English Keywords frequently.

Features of Python Language

Python provides the following major features –

i. Interpreted: - Python is processed at runtime using the interpreter. There is no need to compile a program before execution. It is similar to PERL and PHP
ii. Object- Oriented: - Python follows object-oriented style and design patterns. It includes class definition with various features like encapsulation and polymorphism.

Key Points of Python Language

The key points of Python programming language are as follows –

- It includes functional and structured programming and methods as well as object oriented programming methods.
- It can be used as a scripting language or as a programming language.
- It includes automatic garbage collection.
- It includes high-level dynamic data types and supports various dynamic type checking.
- Python includes a feature of integration with C, C++ and languages like Java.

5.2. SYSTEM TESTING

A software life cycle often includes a separate testing phase, which is after integration and before maintenance. However, testing integral component of the software lifecycle. During the requirement must be checked and tested ok. During planning, the software production involves ensuring that the software:

i. Meet the requirement specification
ii. Meet quality standard
iii. Free from error
iv. Is reliable
v. And produces correct data in all conditions

The testing of the new software system was done with ease because of the modular programming technique employed. This allows each module in the system to be tested independently, while ensure that the required specification is met by the system.

5.3. EXPERIMENTAL WORK AND RESULTS

This section presents and discusses the experimental work and results of the developed Cryptography and Steganography. It may be helpful to look at some screenshots of the program at various points in its progress. This is the main form of Cryptography and Steganography software. It is used for both “encryption and decryption” of any files and “hiding and Unhiding” the encrypted file i.e. secret data. The cover images tested in the present experiment are shown in figures below.

WELCOME SCREEN OF ENCRYPT

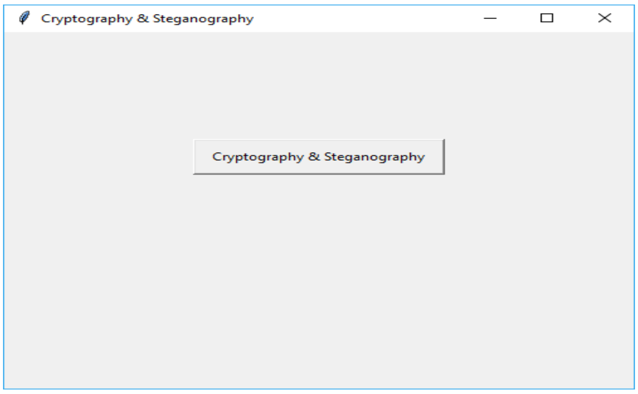


Figure 4.1: Cryptography and Steganography Welcome Interface

The Figure 4.1 above shows the first screen of the program indicating its name, when click on the button “Cryptography & Steganography” take you to the Second stage

SELECTION OF ENCRYPTION OR DECRYPTION PROCESS

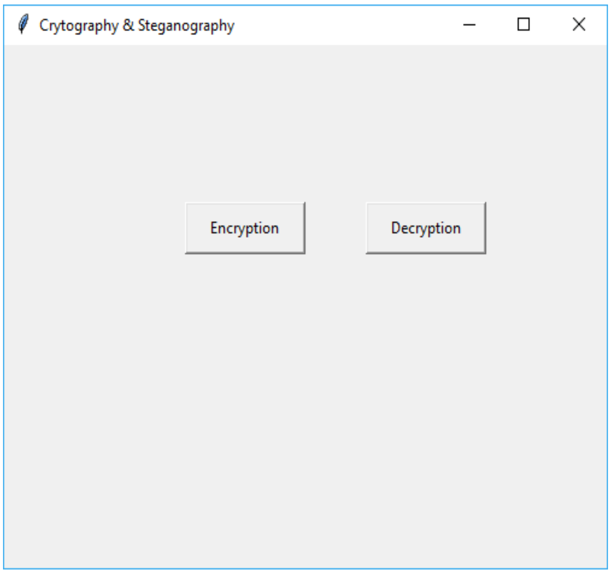
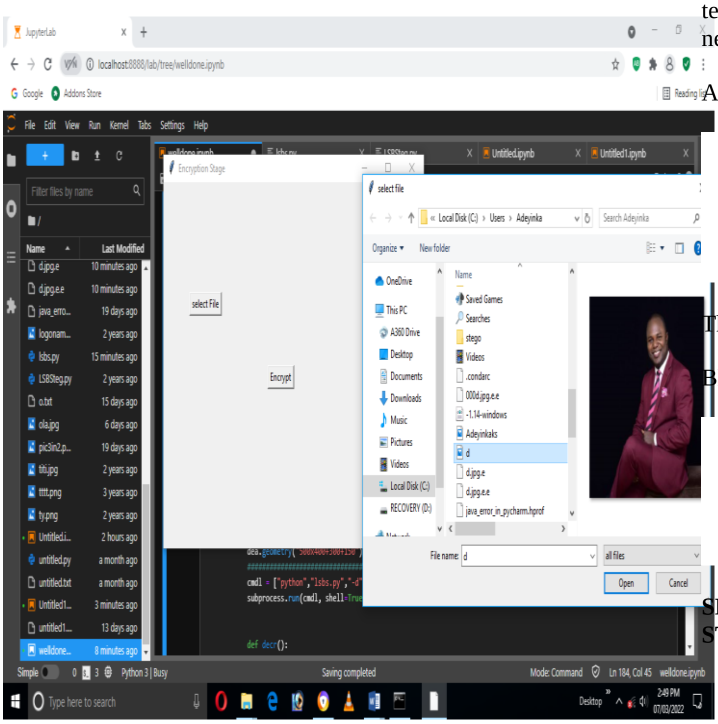


Figure 4.2: Selection of either to Encrypt or Decrypt file

The Figure 4.2 above shows the second screen of the program indicating its name, when click on the button "Encryption" or "Decryption" take you to the next stage.

SELECTION OF ENCRYPTION BUTTON TO ENCRYPT FILE


JupyterLab
localhost:8888/lab/tree/welldone.ipynb
File Edit View Run Kernel Tabs Settings Help
Encryption Stage
select File
Encrypt
select file
Local Disk (C:) › Users › Adeyinka
Organize New folder
File name d
all files
Open
Cancel
Saving completed
Mode: Command
Ln 184, Col 45 welldone.ipynb

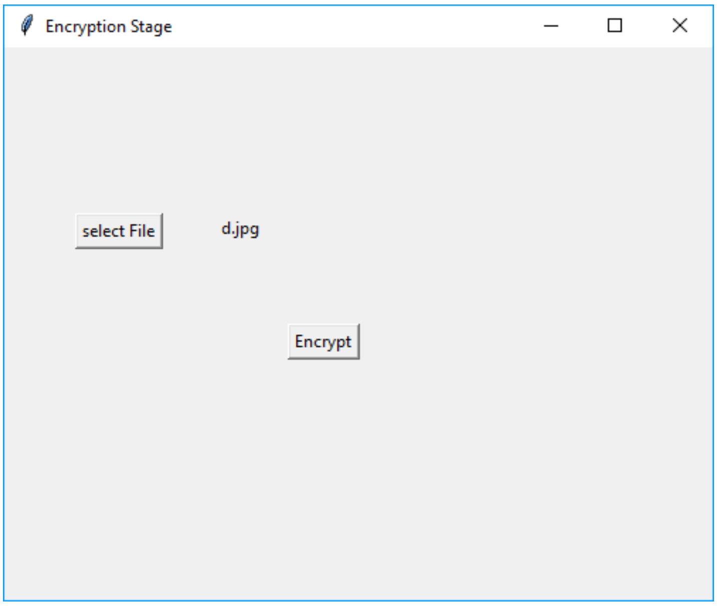
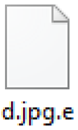


Figure 4.3: Encryption Process of File

The Figure 4.3 above shows the interface of the encryption process after clicked on "Encryption" button in figure 4.2, this stage the file is selected by clicking "select File" in other to choose the file to be encrypt then click the "Encrypt" button to encrypt the file with AES to give cipher text and the cipher text is now encrypted with Blowfish then this take us to the next stage.

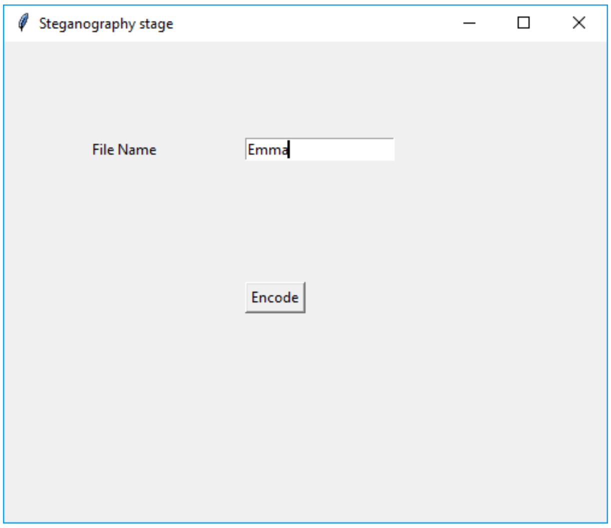
AES File Encryption



Then follow by

Blowfish Encryption

SELECTION OF ENCRYPT BUTTON LEAD TO STEGANOGRAPHY STAGE



Steganography stage

File Name: Emma

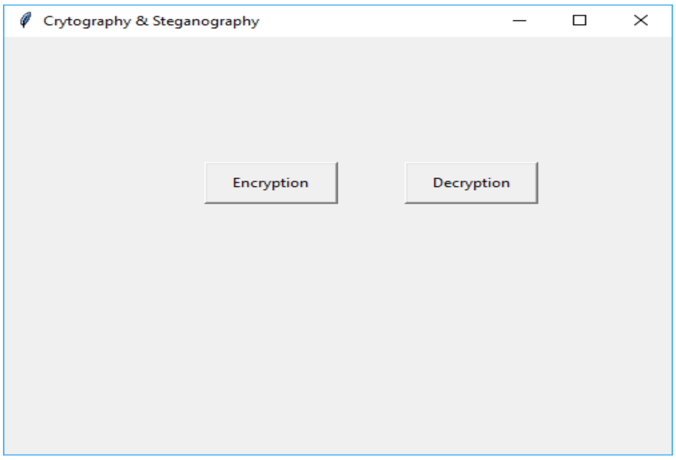
Figure 4.4: Hiding the Encrypted file into an Image

The Figure 4.4 above shows the interface of the Steganography stage when clicked on “Encrypt” button in figure 4.3, then enter the file name of the image encode into and the format of this file will change from ‘jpg’ to ‘bmp’. Now Click the Encode button to hide the file in an existing image in the code and create a new file image where the file is hidden.

The Image that hide the file

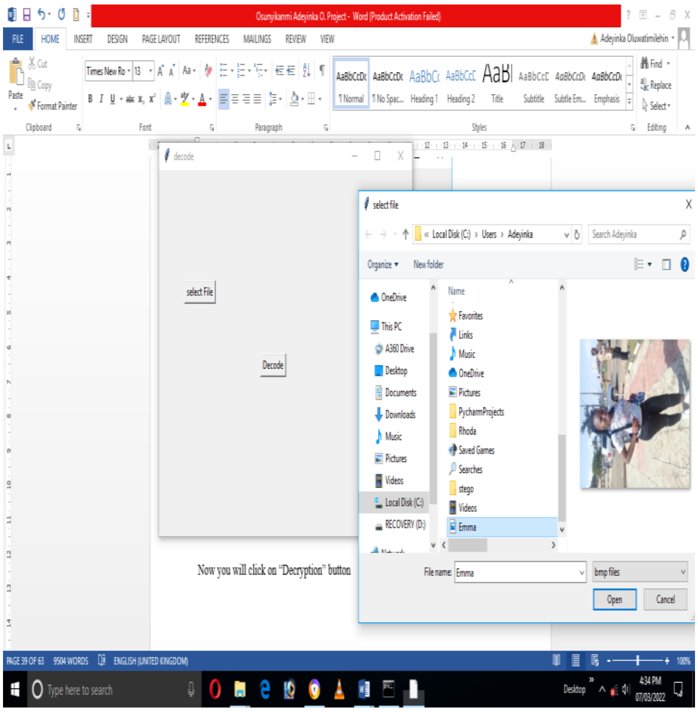
Then Close the program again

And rerun it in order to decryption the hidden file



Crytography & Steganography

Now you will click on “Decryption” button


Osunyikanmi Adeyinka O. Project - Word (Product Activation Failed)
decode
select File
Decode
select file
Local Disk (C:) › Users › Adeyinka
Organize
New folder
Now you will click on “Decryption” button
Open
Cancel

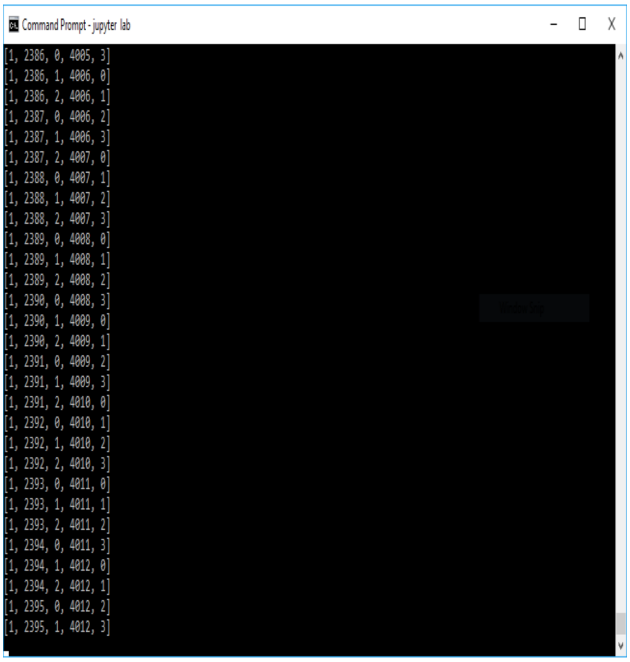
File name: Emma

bmp files

This image above show how the image is been selected, by clicking “select File” and choose the file image to Open. Then click Decode button to extract the file out of the image

Decode the image to decrypt the file

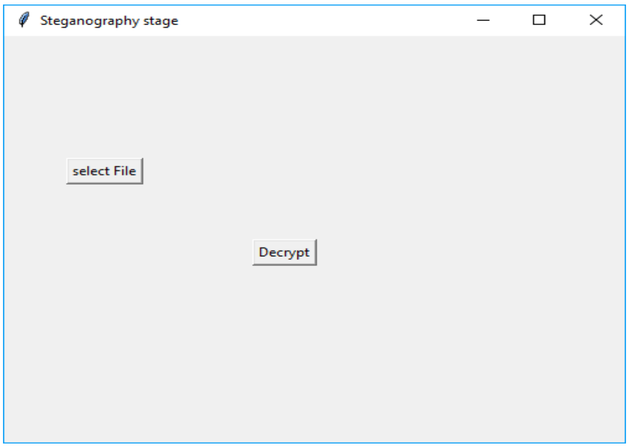
Decode File Process



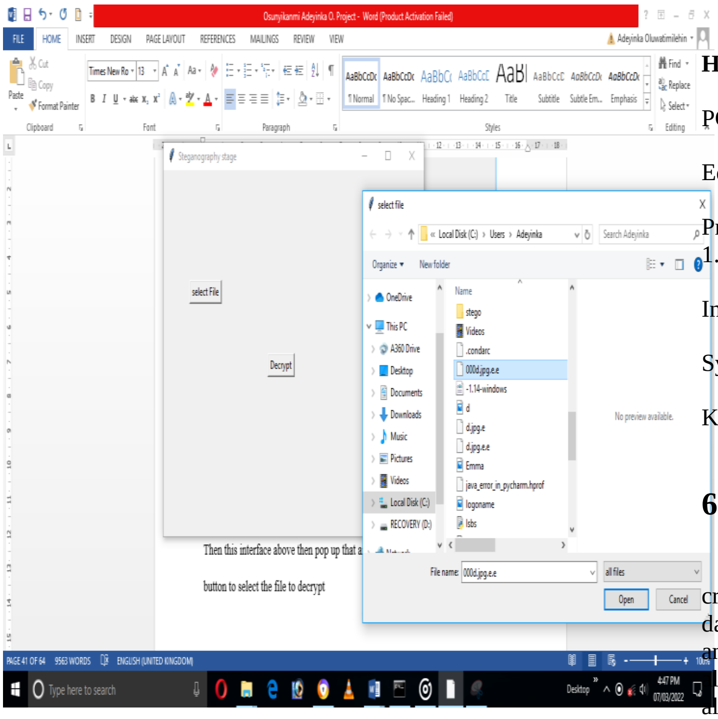
Command Prompt - jupyter lab

```
[1, 2386, 0, 4005, 3]
[1, 2386, 1, 4006, 0]
[1, 2386, 2, 4006, 1]
[1, 2387, 0, 4006, 2]
[1, 2387, 1, 4006, 3]
[1, 2387, 2, 4007, 0]
[1, 2388, 0, 4007, 1]
[1, 2388, 1, 4007, 2]
[1, 2388, 2, 4007, 3]
[1, 2389, 0, 4008, 0]
[1, 2389, 1, 4008, 1]
[1, 2389, 2, 4008, 2]
[1, 2390, 0, 4008, 3]
[1, 2390, 1, 4009, 0]
[1, 2390, 2, 4009, 1]
[1, 2391, 0, 4009, 2]
[1, 2391, 1, 4009, 3]
[1, 2391, 2, 4010, 0]
[1, 2392, 0, 4010, 1]
[1, 2392, 1, 4010, 2]
[1, 2392, 2, 4010, 3]
[1, 2393, 0, 4011, 0]
[1, 2393, 1, 4011, 1]
[1, 2393, 2, 4011, 2]
[1, 2394, 0, 4011, 3]
[1, 2394, 1, 4012, 0]
[1, 2394, 2, 4012, 1]
[1, 2395, 0, 4012, 2]
[1, 2395, 1, 4012, 3]
```

Extract File from Image below



Then this interface above then pop up that am to decrypt my file by click on “select file” button to select the file to decrypt



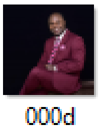
Then this interface above then pop up that a...
button to select the file to decrypt

File name: 000d.jpg.e.e

After select the file then click the “Decrypt” button to decrypt the file

Blowfish decryption outcome

AES decryption outcome



5.4 SYSTEM REQUIREMENTS

The system requirement are categorizes into two and which are software and hardware requirements:

Software Requirements

Programming Language:	Python 3.8, and Jupyter IDE
Operating System:	Microsoft Window 7, 8, 8.1, 10 Microsoft Windows XP, Windows Vista, etc.

Hardware Requirements

PC Name:	DESKTOP-031OLIG
Edition:	Window 10 Pro
Processor:	Intel® Celeron® CPUN3060 @ 1.60GHz 1.60GHz
Installed RAM:	4.00GB
System Type:	64-bit Operating System
Keyboard and Mouse	

6. CONCLUSION

This work is a details work on combination of cryptography and steganography used in securing data/information. It shows how the simplest methods work and how they can be explored. It uses symmetric encryption algorithm to provide more security. Research in this field has already begun. Next to Cryptography, one of the most active fields of research is mass detection tools for hidden contents. This research project has exposed a lot, especially about bit operations and different encryption technique. This work is interesting from the start and only got more interesting as it went on developing. It became more interested in the subject the more interesting as it went on developing. It became more interested in the more we researched it. It learnt that while implementing image Crypto-Stego algorithm is important, thinking of how to detect and attack it and the methods to do so are far more complex than actually doing the cryptography itself. There is a lot of research that is beginning to discover new ways to detect image Crypto-Stego algorithm, most of which involves some variation of statistical analysis. It is interesting to see what other methods will be developed and how accurate they will be at detecting cryptography and steganography.

7. REFERENCES

1, Chen, D. and Zhao, H., (2012). Data security and privacy protection issues in cloud computing.

2. In 2012 International Conference on Computer Science and Electronics Engineering (Vol. 1, pp. 647-651). IEEE.

3. D. Jamil and H. Zaki, (2011) “Security Issues in Cloud Computing and ountermeasures,” International Journal of Engineering Science and Technology, Vol. 3 No. 4, pp. 2672-2676.

4 Getaneh B.T., Gebreiziabher A.M. and Habtamu Z.L. (2016 Department of Information Technology, Collage of Computing and Informatics, Assosa University, Assosa, Ethiopia International Journal of Current Research Vol. 8, Issue, 07, pp.34894-34898, July.

5. Himanshi .C, Shivani .B, and Madhu .G, (2019-2020). Implementing a secure cloud environment: an explorative study, computer science & engineering, kiet group of institutions, delhi-ncr,ghaziabad, india himanshi.c20@gmail.com shivani batra*, india madhunain@gmail.com

6. Huzaifa S. (2013), a brief History of Cryptography, is a principal Product Security Engineer, working for Red Hat Product Security Team.

7. Jacob A.A., Sanika S., Sudeshna C. and Saurabh M. (2019), Application of Steganography Technique in Cloud Computing, Sharda University, Greater Noida, India & Banasthali Vidyapith, Jaipur, India International Conference on Computational Intelligence and Knowledge Economy (ICCIKE) December 11{12, 2019, Amity University Dubai, UAE. See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/339403942

8. Khider .N, Ahmed .K, Asama .K, Hamidy .H, Bagus .P, Emil .N, Mardhiah .M, Inge .H, and Zico Pratama .P (2012), Hybrid cryptography and steganography method to embed encrypted text message within image 1Department of Statistics. Faculty of Management and Economics, Wasit University, Al-Kut, Iraq 2Department of Electrical, Electronic & Systems Engineering, Faculty of Engineering & Built Environment, Universiti Kebangsaan Malaysia, Malaysia 3Institute of Visual Informatics (IVI), Universiti Kebangsaan Malaysia, Malaysia 4Information System, Faculty of Computer Science, Universitas Mercu Buana Jl. Raya Meruya Selatan, Kembangan, Jakarta 11650 5Universitas Putra Indonesia YPTK, Padang, 25221, Indonesia 6School of Electronic Engineering and Computer Science, Queen Mary University of London *khnsaif@uowasit.edu.iq1

9. Khoshkholghi, M.A., Abdullah, A., Latip, R. and Subramaniam, S., (2014). Disaster recovery in cloud computing: A survey.

10. Kumar. G, (2019), Computer Science & Engineering, VIT, RKDF University, Bhopal, India (UIJRT) United International Journal for Research and Technology | Volume 01, Issue 02, 2019 | ISSN: 2582-6832.

11. Peter M. and Tim G. in an October, (2009) presentation titled “Effectively and Securely Using the Cloud Computing Paradigm,”

12. Rohan Jathanna and Dhanamma Jagli. Int. Journal of Engineering Research and Application www.ijera.com ISSN: 2248-9622, Vol. 7, Issue 6, (Part -5) June 2017, pp.31-38 (Department of Mca, VESIT, Mumbai Email: rohan.jathanna@ves.ac.in), (Department of Mca, VESIT, Mumbai Email: dsjagli.vesit@gmail.com)

13. Ronald L.K., and Russell D.V. (2010), Cloud Security A Comprehensive Guide to SecureCloud Computing Cloud Security: A Comprehensive Guide to Secure Cloud Computing Published by Wiley Publishing, Inc. 10475 Crosspoint Boulevard Indianapolis, IN 46256 www.wiley.comCopyright © 2010 by Wiley Publishing, Inc., Indianapolis, Indiana Published simultaneously in Canada ISBN: 978-0-470-58987-8

14. Tony M.D.(2019) A brief history of cryptography, vol. 1 NO. 11| pg. 1/1 http://www.inquiries journal.com/amp/1698/a-brief-history-of-cryptography

15. U. K. Essays, (2018), the History and Background of Steganography. Retrieved from https://www.ukessays/english-language/background-of-steganography.php?vref=1

16. Yazan Al-Issa, Mohammad A.O, and Ahmed T.H.,(2019) Journal of Healthcare Engineering Volume, Article ID 7516035, 15 pages https://doi.org/10.1155/2019/7516035

17. Yang Li, Xiaoling Tao, Wei Wu; Joseph K. Liu,(2018) GO-CP-ABE: group-oriented ciphertext-policy attribute-based encryption, Int. J. of Embedded System, 2018 Vol.10, No. 1, pp.62-70