Feature Based Cryptanalytic Technique for Digital Forensics Analysis of Visual Cryptographic Digital Image Data Based on Formal Concept Analysis

Quist-Aphetsi Kester [1,2,3], Laurent Nana[1], Anca Christine Pascu[1], Sophie Gire[1],Jojo M. Eghan[3], Nii Narku Quaynor[3]

[1] Lab-STICC (UMR CNRS 6285), European University of Brittany, University of Brest, France
Kester.quist-aphetsi@univ-bret.fr / kquist@ieee.org
[2] Faculty of Informatics, Ghana Technology University College
[3]Department of Computer Science and Information Technology, University of Cape Coast

Abstract - *Lossless pixel value encrypted images still maintains the some properties of their respective original plain images. Ciphered Images that maintain the properties of their plain images of a given domain are very useful in certain applications where the conservation of pixel values but visual concealment is of a paramount concern. Medical images that have a fully reversible and recoverable process are of key importance in Medicine. Hence visually ciphered images stored or transmitted over secured or unsecured networks can also be analyzed in a forensic investigation to determine possible plain image equivalence. Digital Forensics processes have played crucial role in fighting crime both in society and cyberspace. In this paper, feature based cryptanalytic technique for digital forensics analysis of visual cryptographic digital image data based on formal concept analysis was proposed. Different techniques of visual cryptographic approaches were engaged in ciphering the plain image and our proposed approach was engaged in the cryptanalysis of the plain image after feature extractions from both the plain and the ciphered images. A lattice was generated which was then used authenticate and match the ciphered images to their respective ciphered plain images. At the end, the Galois lattice of both ciphered and plain image remained the same.*

Keywords: formal concept analysis, cryptanalysis, digital forensics, lattices, digital image, feature extraction, pixels

1 Introduction

The high increase in multimedia image usage for data communications over secured and unsecured network was due to the digitization of processes such as digital filing of documents, video conferencing, social media activities etc[1-3]. Secret communications between two parties using multimedia can also involve communications of encrypted image [4]. The rise in crime and availability of approaches to securing data to the general public has created avenues for people to implement cryptographic approaches in securing and concealing image contents. These approaches hinder criminal investigative procedures and prevent easy analysis of digital evidences [5-7]. Cryptanalysis is an effective way of analyzing ciphers and encrypted data with the high hopes of decrypting the data or breaking the cipher. These approaches are very crucial in solving a range of issues in military communication applications and digital forensics toolkits. Cryptographers overtime have device the means of securing messages as well breaking codes [8-9].

Security in multimedia supplications is critical for the future. In this paper, we proposed a feature based cryptanalytic technique for digital forensics analysis of visual cryptographic digital image data based on formal concept analysis was proposed. Features were extracted from both plain and ciphered images and then lattices were built to help match plain images to their respected ciphered images. At the end, the Galois lattice of both ciphered and plain image remained the same. The paper has the following structure; section II Related works, section III is Methodology, section IV Results and analysis, and section V concluded the paper.

2 Literature Review

Forensics approaches cannot be effective in the presence of anti forensics procedures such as altering of content data during recovery process, incomplete evidence, encrypted data etc And as society has become increasingly reliant upon digital images to communicate visual information, a number of forensic techniques have been developed to verify the authenticity of digital images. Hence the digital forensics community requires new tools and strategies for the rapid turnaround of large forensic targets [10-13]. Alin C in their work described several statistical techniques for detecting traces of digital tampering in the absence of any digital watermark or signature. They quantify statistical correlations that result from specific forms of digital tampering, and devise detection schemes to reveal these correlations [14]. Dehnie, S proposed a digital image forensics for identifying computer generated and digital camera images [15]. Formal Concept analysis Formal is a field of applied mathematics based on the mathematization of concept and conceptual hierarchy. It thereby activates mathematical thinking for conceptual data analysis and knowledge processing [16]. Its applications in forensics are normally in the domain of computer aided investigations. Where the data collected on crime re being analyzed using the approach[17-18]. In our approach we engaged feature based cryptanalytic technique for digital forensics analysis of visual cryptographic digital image data based on formal concept analysis was proposed. Different techniques of visual cryptographic approaches were engaged in ciphering the plain image and our proposed approach was engaged in the cryptanalysis of the plain image after feature extractions from both the plain and the ciphered images.

3 Methodology

Our method employed a cryptanalytic procedure by using features generated from digital images which were then used to construct a Galois lattice. The features were extracted in such a way that a change in pixel value can cause a change in concept of the lattice. This means that if there is no pixel expansion in the ciphering process of the image, a perfect match of its plain image can be obtained by using our proposed method. The overall process is indicated in figure 1 below.

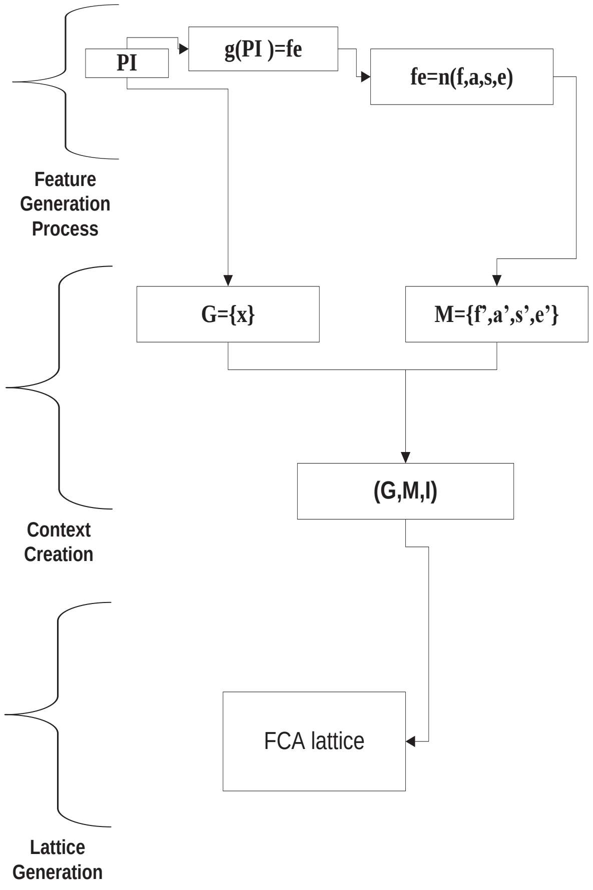


Figure 1: Summary of the Entire process

From figure 1:
PI=Plain image
g(PI)=function that operated on the plain image to produce the features
n(f,a,s,e)=function of the features
fe= the feature results
f=sum of all frequency of each pixel in the image
a=arithmetic mean of all the pixel values in the image
s=standard deviation all the pixel value in the image
e=entropy of all the pixel value the image
x=a distinct chosen pixel value number
x'=frequency of x
f'=x'/f, a'=x'/a, s'=x'/s and e'=x'/e
G =set objects extracted from the image
M=sets attributes obtained from the image
Concepts obtained are (G,M,I)
r(G,M,I)=the function that operated on G,M an d I concept to produce K
ImC=the image encryption algorithm that operated on K and Pi to produce CI

3.1 The Feature Extraction

Let I= an image=f (R, G, B)
I is a color image of m x n x 3 arrays

$$\begin{pmatrix} R & G & B \\ r_{i1} & g_{i2} & b_{i3} \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ r_{n1} & g_{n2} & b_{n3} \end{pmatrix}$$

(R, G, B) = m x n and R, G, B $\in$ I
(R o G) i j = (R) ij. (G) ij
where r_11 = first value of R
r= [ri1] (i=1, 2… m) and
x $\in$ r_i1 : [a, b]= {x $\in$ I: a $\leq$ x $\geq$ b}
a=0, b=255 and R= r= I (m, n, 1)
where g_12 = first value of G
g= [gi2] (i=1, 2... m) and
x $\in$ g_i1: [a, b]= {x $\in$ I: a $\leq$ x $\geq$ b}
a=0 , b=255 and G= g= I (m, n, 1)
and b_13 = first value of B
g= [bi3] (i=1, 2... m) and
x $\in$ b_i1 : [a, b]= {x $\in$ I: a $\leq$ x $\geq$ b}
a=0, b=255 and B=b= I (m, n, 1)
Such that R= r= I (m, n, 1)

Let X=freq(x) which is the number of times x occurred in r,g and b

$$f = \sum_{i=m}^{n} X_i$$

$$a = \frac{\sum_{n=1}^{k} x_n}{k}$$

Where x ∈ b_i1 : [a, b]= {x ∈ I: a ≤ x ≥ b}

$$s = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2}, \quad \text{where} \quad \mu = \frac{1}{N}\sum_{i=1}^{N} x_i.$$

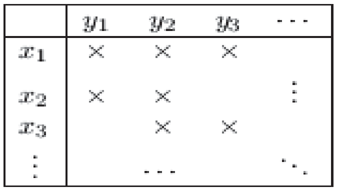
Entropy is defined as

$e = -\sum_{\eta=0}^{\varepsilon\text{-}1} \Psi$ (xi). log2 (Ψ (xi))

Where:
δ= Entropy of image
ε = Gray value of an input image (0-255).
Ψ (η) = Probability of the occurrence of symbol η

3.2 The Feature Classification using FCA

Formal concept analysis (FCA) is a method of data analysis with growing popularity across various domains. FCA analyzes data which describe relationship between a particular set of objects and a particular set of attributes.
If g ∈ A and m ∈ B then (g,m) ∈ I ,or gIm.
A formal context is a triple (G,M,I), where
•G is a set of objects,
• M is a set of attributes
•and I is a relation between G and M.
• (g,m) ∈ I is read as „object g has attribute m.
For A ⊆ G, we define
A´:= {m ∈ M | ∀g ∈ A:(g,m) ∈ I }.
For B ⊆ M, we define dually
B´:= {g ∈ G | ∀m ∈ B:(g,m) ∈I }.
For A, A1, A2 ⊆ G holds:
• A1 ⊆ A2 ⇒ A`2 ⊆A`1
• A 1 ⊆A``
• A`= A```
For B, B1, B2 ⊆ M holds:
• B1 ⊆ B2 ⇒ B_2 ⊆ B_1
• B ⊆ B``
• B`= B```
A formal concept is a pair (A, B) where
• A is a set of objects (the extent of the concept),
• B is a set of attributes (the intent of the concept),
•A`= B and B`= A.
The concept lattice of a formal context (G, M, I) is the set of all formal concepts of (G, M, I), together with the partial Order (A1, B1) ≤ (A2, B2): ⇔ A1 ⊆ A2 (⇔ B1 ⊇ B2) (Priss, U, 1997).
The concept lattice is denoted by $\mathfrak{B}$(G,M,I) .
• Theorem: The concept lattice is a lattice, i.e. for two concepts
(A1, B1) and (A2, B2), there is always
•a greatest common sub-concept: (A1⋂A2, (B1⋃ B2) ´´)
•and a least common super-concept: ((A1 ⋃ A2) ´´, B1⋂B2)
More general, it is even a complete lattice, i.e. the greatest common sub-concept and the least common super-concept exist for all (finite and infinite) sets of concepts.
Corollary: The set of all concept intents of a formal context is a closure system. The corresponding closure operator is h(X):= X``.
An implication X→Y holds in a context, if every object having all attributes in X also has all attributes in Y.
Def.: Let X ⊆M. The attributes in X are independent, if there are no trivial dependencies between them



	y_1	y_2	y_3	...
x_1	×	×	×	
x_2	×	×		⋮
x_3		×	×	
⋮		...		⋱

Figure 2: A table of attributes and properties

The table above represents logical attributes represented by a triplet (X, Y, I), where I is a binary relation between X and Y. The elements of X are called objects and correspond to table rows, elements of Y are called attributes and correspond to table columns, and for x ∈ X and y ∈ Y , (x, y) ∈ I indicates that object x has attribute y while (x, y) ∈ I.
From the image we chose our objects as a classified range of values of x, where x ∈ b_i1 : [a, b]= {x ∈ I: a ≤ x ≥ b} and a=0, b=255. G= {0-25, 26-50, 51-75, 76-100,101-125, 126-150,151-175, 176-200,201-225, 256-255}.
f'=j/f
a'=j/a
s'=j/s
e'=j/e
where j is the sum of all the frequencies of all numbers that fall within the range of each object.
Where B= {f,,a',s',e'} Major attributes and f' has {B_1, B_2, B_3 and B_4} as sub attributes. Therefore, a', s' and e' have the same sub attributes as f'. But {B_1, B_2, B_3, B_4} maps directly and exactly on at least one element of {0-0.25, 0.26-0.50, 0.51-0.75, 0.76-1.0}.

4 Analysis and Results

We chose a 24 bit depth image jpg of dimension 960 pixels by 720 pixels with a horizontal resolution of 96 dpi and a vertical resolution of 96 dpi.

Figure 3: Plain image

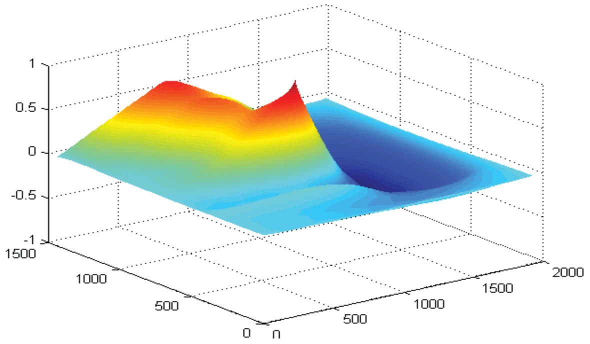


Figure 4: The graph of the normalized cross-correlation of the matrices of the plain image

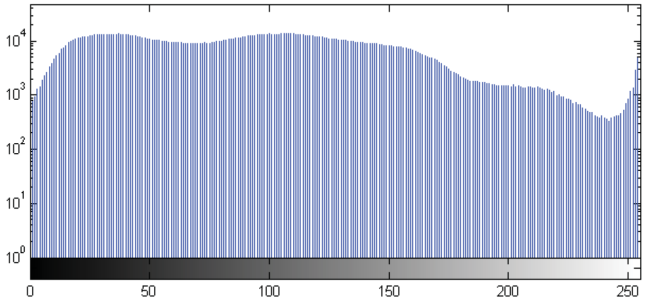


Figure 5: A graph of frequency of pixel values

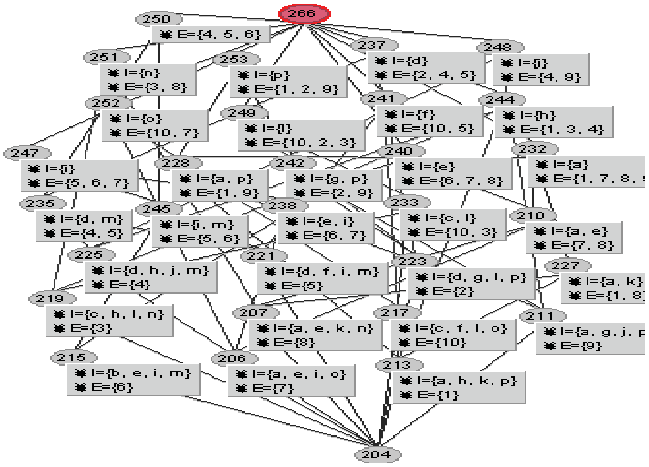


Figure 6: A Galois lattice generated from the plain image

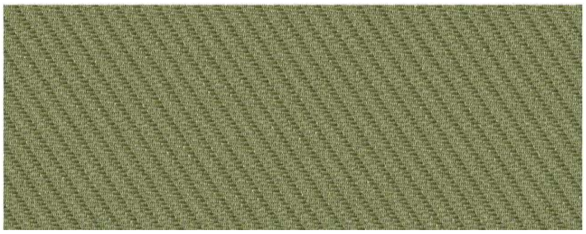


Figure 7: The ciphered image

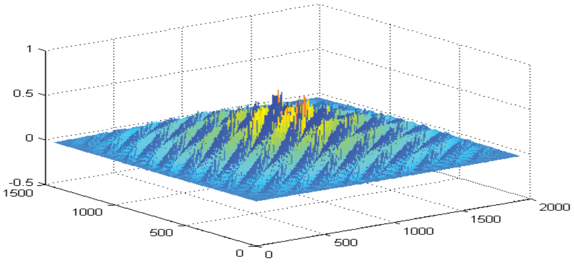


Figure 8: The graph of the normalized cross-correlation of the matrices of the ciphered image

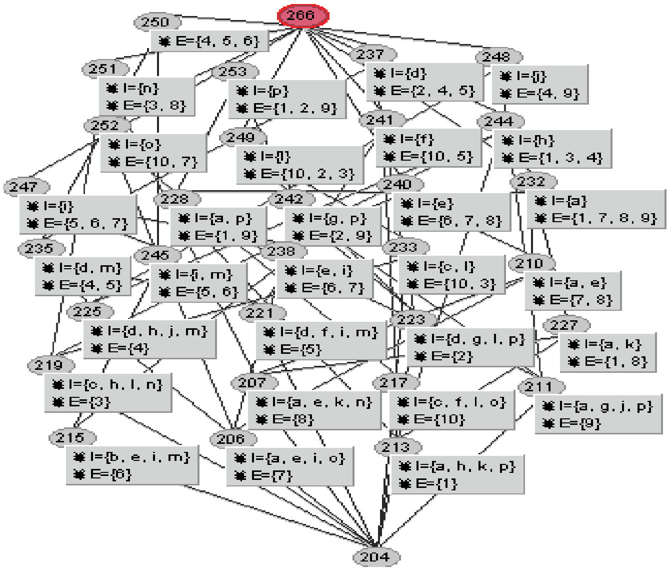


Figure 9: A Galois lattice generated from the ciphered image

Table 1: Table Objects X and attributes

G	j	f'	a'	s'
0-25	169938	0.081953	1502.998	2343.901
26-50	316517	0.152641	2799.4	4365.619
51-75	235764	0.113698	2085.189	3251.819
76-100	289149	0.139443	2557.347	3988.141
101-125	319446	0.154054	2825.306	4406.018
126-150	235758	0.113695	2085.136	3251.736
151-175	146001	0.070409	1291.29	2013.746
176-200	46113	0.022238	407.8414	636.0221
201-225	31294	0.015092	276.7764	431.6283
226-255	283620	0.136777	2508.446	3911.881

The graph of the normalized cross-correlation of the matrices of the plain image in figure 3 was plotted as shown in figure 4. The features f=2073600, a=113.066, s=72.5022 and e=7.1945 were extracted from the plain images. The frequencies of the pixel values of the plain image were plotted as shown in figure 5 and a Galois lattice was generated from the features extracted from the plain image based on table 1. Table two below showed six different techniques of visual cryptography applies to the image. The results for the features were constant even though the visual states of the encrypted images differ for the various approaches engaged.

Table 2: Table Objects X and attributes

Approaches	f	a	s	e
1	2073600	113.066	72.5022	7.1945
2	2073600	113.066	72.5022	7.1945
3	2073600	113.066	72.5022	7.1945
5	2073600	113.066	72.5022	7.1945
6	2073600	113.066	72.5022	7.1945

At the end all the graphs plotted and the concept lattices were the same for both the ciphered and the plain images. A set of encrypted images were tested against their corresponding ciphered images and the results were effective. This means that a plain image can be mapped directly to its corresponding ciphered image without decrypting the ciphered image for a given data set.

5 Conclusion

Based on the extracted features from both the plain and the ciphered images, a Galois lattice was constructed. We have realized that the Gallois lattice generated from the plain image as well as the features extracted from the plain image was the same as that of the ciphered image irrespective of pixel displacement that occurred. This was as a result of conservation of pixel values. This makes our approach a suitable forensics analysis of encrypted images based on visual cryptography or pixel displacement. Our results were very effective for different kind of approaches that engaged a non pixel expansion technique in ciphering the image. Our proposed method can help also in the indexing of images based on the extracted features and can help in evidence analysis of ciphered images based on visual cryptographic techniques that conserves pixel values.

. **Acknowledgments.** This work was supported by Lab-STICC (UMR CNRS 6285) at UBO France, AWBC Canada, Ambassade de France-Institut Français-Ghana and the DCSIT-UCC, and also Dominique Sotteau (formerly directeur de recherche, Centre national de la recherche scientifique (CNRS) in France and head of international relations, Institut national de recherche en informatique et automatique, INRIA) and currently the Scientific counselor of AWBC.

References

[1] Yadav, R., Gupta, R. K., & Singh, A. P. (2015). A Survey of Image Compression using Neural Network and Wavelet Transformation. International Journal of Research, 2(1), 301-305.

[2] Yen, N. Y., Zhang, C., Waluyo, A. B., & Park, J. J. (2015). Social Media Services and Technologies Towards Web 3.0. Multimedia Tools and Applications, 1-7.

[3] Roy, S. D., & Zeng, W. (2015). Revelations from Social Multimedia Data. In Social Multimedia Signals (pp. 135-142). Springer International Publishing.

[4] Nguyen, K. T., Laurent, M., & Oualha, N. (2015). Survey on secure communication protocols for the Internet of Things. Ad Hoc Networks.

[5] Hashem, F. S., & Sulong, G. (2015). Passive Aproaches for Detecting Image Tampering: A Review. Jurnal Teknologi, 73(2).

[6] Cao, Y., Gao, T., Sheng, G., Fan, L., & Gao, L. (2015). A New Anti‐forensic Scheme— Hiding the Single JPEG Compression Trace for Digital Image. Journal of forensic sciences, 60(1), 197-205.

[7] Wang, X. G. (2015, January). Research on digital forensics and its relevant problems. In Electronics, Information Technology and Intellectualization: Proceedings of the International Conference EITI 2014, Shenzhen, 16-17 August 2014 (p. 43). CRC Press.

[8] Yap, W. S., Phan, R. C. W., Yau, W. C., & Heng, S. H. (2015). Cryptanalysis of a new image alternate encryption algorithm based on chaotic map. Nonlinear Dynamics, 80(3), 1483-1491.

[9] Ahmad, M., Khan, I. R., & Alam, S. (2015, January). Cryptanalysis of Image Encryption Algorithm Based on Fractional-Order Lorenz-Like Chaotic System. In Emerging ICT for Bridging the Future-Proceedings of the 49th Annual Convention of the Computer Society of India CSI Volume 2 (pp. 381-388). Springer International Publishing.

[10] Kessler, G. C. (2007, March). Anti-forensics and the digital investigator. In Australian Digital Forensics Conference (p. 1).

[11] Harris, R. (2006). Arriving at an anti-forensics consensus: Examining how to define and control the anti-forensics problem. digital investigation, 3, 44-49.

[12] Stamm, M. C., & Liu, K. R. (2011). Anti-forensics of digital image compression. Information Forensics and Security, IEEE Transactions on, 6(3), 1050-1065.

[13] Richard III, G. G., & Roussev, V. (2006). Next-generation digital forensics. Communications of the ACM, 49(2), 76-80.

[14] Popescu, A. C., & Farid, H. (2005, January). Statistical tools for digital forensics. In Information Hiding (pp. 128-147). Springer Berlin Heidelberg.

[15] Dehnie, S. (2006, October). Digital image forensics for identifying computer generated and digital camera images. In Image Processing, 2006 IEEE International Conference on (pp. 2313-2316). IEEE.

[16] Ganter, B., & Wille, R. (2012). Formal concept analysis: mathematical foundations. Springer Science & Business Media.

[17] Kester, Q. A. (2013). Criminal Geographical Profiling: Using FCA for Visualization and Analysis of Crime Data. arXiv preprint arXiv:1310.0864.

[18] Kester, Q. A. (2013). Visualization and analysis of geographical crime patterns using formal concept analysis. arXiv preprint arXiv:1307.8112.

[19] Poelmans, J., Elzinga, P., Dedene, G., Viaene, S., & Kuznetsov, S. O. (2011). A concept discovery approach for fighting human trafficking and forced prostitution. In Conceptual Structures for Discovering Knowledge (pp. 201-214). Springer Berlin Heidelberg.