A known-plaintext heuristic attack on the Fourier plane encryption algorithm

Unnikrishnan Gopinathan, David S. Monaghan, Thomas J. Naughton*, and John T. Sheridan

School of Electrical, Electronic and Mechanical Engineering, University College Dublin, Belfield, Dublin 4, Ireland

∗ *Dept. of Computer Science, National University of Ireland Maynooth, Ireland*

john.sheridan@ucd.ie

Abstract: The Fourier plane encryption algorithm is subjected to a known-plaintext attack. The simulated annealing heuristic algorithm is used to estimate the key, using a known plaintext-ciphertext pair, which decrypts the ciphertext with arbitrarily low error. The strength of the algorithm is tested by using this estimated key to decrypt a different ciphertext which was also encrypted using the same original key. We assume that the plaintext is amplitude-encoded real-valued image, and analyze only the mathematical algorithm rather than a real optical system that can be more secure. The Fourier plane encryption algorithm is found to be susceptible to a known-plaintext heuristic attack.

OCIS codes: (070.2580) Fourier optics and optical signal processing : Fourier optics (070.4560) Fourier optics and optical signal processing : Optical data processing (200.3050) Optical computing : Information processing

References and links

1. B. Javidi, ed. *Optical and Digital Techniques for Information Security*, (Springer Verlag, 2005).
2. Ph. Réfrégier and B. Javidi, "Optical image encryption based on input plane and Fourier plane random encoding," Opt. Lett. **20**, 767-769 (1995).
3. G. Unnikrishnan, J. Joseph, and K. Singh, "Optical encryption system that uses phase conjugation in a photorefractive crystal," Appl. Opt. **37**, 8181-8186 (1998).
4. O. Matoba and B. Javidi, "Encrypted optical memory system using three-dimensional keys in the Fresnel domain," Opt. Lett. **24**, 762-764 (1999).
5. B. Javidi and T. Nomura, "Securing information by use of digital holography," Opt. Lett. **25**, 28-30 (2000).
6. P. C. Mogensen and J. Glckstad, "Phase-only optical encryption," Opt. Lett. **25**, 566-568 (2000).
7. E. Tajahuerce and B. Javidi, "Encrypting three-dimensional information with digital holography," Appl. Opt. **39**, 6595-6601 (2000).
8. B. M. Hennelly and J. T. Sheridan, "Optical image encryption by random shifting in fractional Fourier domains," Opt. Lett. **28**, 269-271 (2003).
9. T.J. Naughton and B. Javidi, "Compression of encrypted three-dimensional objects using digital holography," Opt. Eng. **43**, 2233-2238 (2004).
10. A. Carnicer, M. Montes-Usategui, S. Arcos, and I. Juvells, "Vulnerability to chosen-cyphertext attacks of optical encryption schemes based on double random phase keys," Opt. Lett. **30**, 1644-1646 (2005).
11. Y. Frauel, A. Castro, T.J. Naughton, and B. Javidi, "Security analysis of optical encryption," Proc. SPIE **5986**, 25-34 (2005).
12. W. Stallings, *Cryptography and Network Security*, Third edition, (Prentice Hall, 2004).
13. S. Kirkpatrick, C. D. Gellatt and M. P. Vecchi, "Optimization by simulated annealing," Science **220**, 771-680 (1983).
14. M. Nieto-Vesperinas, R. Navarro, and J. F. Fuentes, "Performance of a simulated annealing algorithm for phase retrieval," J. Opt. Soc. Am. A **5**, 30-38 (1988).

#67670 - $15.00 USD Received 31 January 2006; revised 31 March 2006; accepted 3 April 2006

1. Introduction

Many image encryption algorithms have been proposed in recent years Refs. [1-9], to cite just a few. Many of these algorithms can be implemented using optical techniques taking advantage of both the natural two-dimensional (2D) imaging capabilities of optics and the parallelism achievable with optical processing. Optical systems are also capable of encrypting real-world 3D objects [7, 9]. Optical encryption algorithms have yet to undergo the rigorous cryptanalysis which all conventional cryptographic algorithms are subjected to. There are instances in the literature when an optical encryption mechanism is shown to be robust to blind decryption for selected keys in the key space. However, this is not sufficient to evaluate the strength of an encryption algorithm. Two previous studies have already been performed on the strength of optical encryption [10, 11], specifically on the well-known the Fourier plane encoding algorithm [2] also analyzed in this paper. Carnicer et al. [10] and Frauel et al. [11] examined exact solutions to pixels in the decryption key, with the former concentrating only on chosen-plaintext attacks [12] (where the attacker had the advantage of being able to choose whatever plaintext-ciphertext pair they want) and the latter considering both chosen-plaintext and known-plaintext attacks [12]. In Ref [10], a Dirac delta function is used as the chosen plaintext to find the Fourier plane key. The method proposed in Ref [11] is based on the principle that Fourier plane encoding algorithm is linear. In this paper, we take the first steps of a cryptanalysis using heuristics to estimate the decryption key and describe a known-plaintext attack on the Fourier plane-encoding algorithm. The advantage of using a heuristic to estimate decryption key pixels rather than an analytical technique to determine exact solutions for the pixels is that heuristics can take considerable less time to run. Furthermore, since the data routinely encrypted by optical encryption is image data, slight errors in the decrypted data can often be tolerated, and so an exact solution is not generally required.

In a known-plaintext attack, discussed in this paper, a single arbitrary (unchosen) plaintext-ciphertext pair and the encryption method are known by the attacker. Furthermore, we assume that the plaintext is amplitude-encoded real-valued image. With this a priori information, our approach is to use a simulated annealing (SA) algorithm [13] to find a key which decrypts the ciphertext (encrypted image) with some chosen error threshold. We choose this error to be sufficiently low so that the entire information in the input image can be recovered. The so obtained key is used to decrypt a different“ unseen plaintext, encrypted using the same set of original keys. The encryption algorithm is evaluated based: (i) on its ability (in terms of length of time on a particular computing platform) to withstand the SA heuristic decryption attack using the known-plaintext, and (ii) based on the resulting error in the decryption of an unseen plaintext.

2. Fourier plane encoding algorithm

The Fourier plane algorithm encodes an input image f to a stationary white noise by using two statistically independent random phase codes in the input plane and Fourier plane. The image is multiplied by the first random phase code R_1. A Fourier transform is performed on this product and multiplied by the second random phase code R_2. A second Fourier transform gives the encrypted image. The encoded image ψ can be expressed mathematically as

$$\psi = [f(.)R_1] * \hat{R}_2(.) \tag{1}$$

where $\hat{X}$ denotes the Fourier transform of X, and $*$ denotes a convolution. The intensity of the approximated input image $\tilde{f}$ is decoded as

$$\left|\tilde{f}\right|^2 = \left|\psi(.) * \hat{R}_3(.)\right|^2 \tag{2}$$

#67670 - $15.00 USD Received 31 January 2006; revised 31 March 2006; accepted 3 April 2006

where R_3 is the SA algorithm's estimation of the complex conjugate of R_2. If we assume that the input image is real-valued, then we are only interested in $|f|$ and can effectively ignore R_1. Therefore in our analysis, when we refer to the decryption key we mean mask R_2.

3. Known-plaintext attack using SA algorithm

One way to test the strength of an encryption algorithm can be stated as follows: with particular known a priori information, how difficult is it for an attacker to find the key to a ciphertext which would make it possible to retrieve the plaintext? In known-plaintext cryptanalysis, the attacker has a priori knowledge of the encryption mechanism as well as a plaintext and ciphertext pair. If the attacker is able to find the key used for a given plaintext-ciphertext pair, then the security of all the past and future ciphertexts, which used the same key, arecompromised.

Let us assume that the attacker tries to decrypt a ciphertext encrypted using Fourier plane encoding by the blind decryption method. In this method, (s)he tries to decrypt the ciphertext by randomly picking a key from the key space, and compares the resulting 'decrypted' plaintext to the original plaintext.

The probability of finding the correct mask in t searches would be approximately tK^{-1} where K is the size of the key space. For an $N \times N$ pixel encryption phase mask with m phase levels, the key space is as large as $K = m^{N \times N}$. If one considers that some fraction $r(\varepsilon) \in [0,1]$ of the keys could give a decryption with some acceptable error ε, then the probability of finding one of these (estimated) keys increases to $t[r(\varepsilon)K]^{-1}$ for a particular ε. If the attacker finds any one of these estimated keys (s)he would decrypt the ciphertext with some error. The important question, however, is whether or not this estimated decryption key can also be used to decrypt another (unseen) image, encrypted with the same original encryption key. If a single unseen image is decrypted, then one could consider the encryption key as having been broken. If no unseen image can be found that is adequately decrypted, then one could consider the encryption algorithm as having withstood a SA heuristic decryption attack using that particular computing platform for that amount of time.

We apply a SA algorithm [13] to find a phase mask which would approximately decrypt the ciphertext $\psi(.)$ to give an estimated plaintext $\tilde{f}(.)$ such that the normalized root mean squared (NRMS) error is equal to or less than some threshold ε. The NRMS error is calculated as

$$NRMS = \sqrt{\frac{\sum_{i=1}^{N}\sum_{j=1}^{N}|I_d(i,j) - I(i,j)|^2}{\sum_{i=1}^{N}\sum_{j=1}^{N}|I(i,j)|^2}} \tag{3}$$

where $I_d(.) = \left|\tilde{f}(.)\right|^2$ and $I(.) = |f(.)|^2$

Our SA algorithm involved the following steps:

Step 1: An initial guess for the random phase mask R_3 is made by assigning the phase of the Fourier transform of the encrypted image $\psi(.)$ to every other pixels in both dimensions (i.e. half the number of pixels) [14]. The other half was chosen randomly from a uniform probability distribution in the range $[0\ 2\pi)$. The step counter n is initialized to zero and the error threshold ε is set to the desired value. The initial temperature T is chosen sufficiently high so that the perturbation probability in Step 4 will be large.

Step 2: The cost value E is calculated as the NRMS error between the decrypted image and the original plaintext image.

Step 3: One pixel of R_3 is randomly selected and perturbed by $\alpha_n \pi$ where α_n is the scale of perturbation [14] at the n^{th} step, chosen as $[B\,log(A+E_n)/C]^p$ where E_n denotes the cost function calculated at the n^{th} step as in Step 2. The parameters A, B, C and p will have been fixed at the beginning of the algorithm so that $\alpha_0 \approx 1$. p determines the rate of decrease of α .The new cost function, E^{new}, is calculated using the perturbed phase mask.

#67670 - $15.00 USD Received 31 January 2006; revised 31 March 2006; accepted 3 April 2006

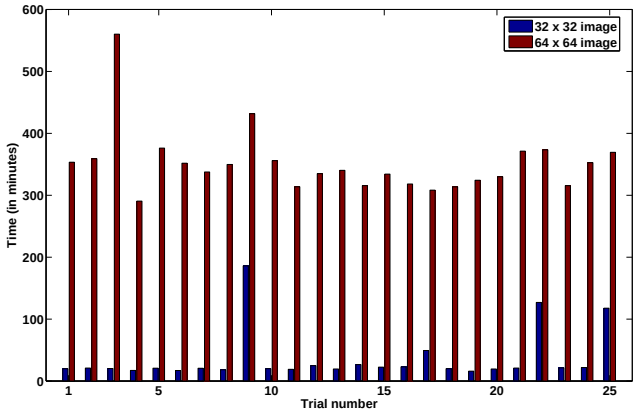

Fig. 1. Time taken to estimate the key which would decrypt the encrypted image of A with an NRMS error of 0.1. Results of 50 trials, 25 each for cases when plaintext is a 32×32 pixel image and 64×64 pixel image.

Step 4: The change in cost function due to the perturbation is calculated as $\Delta E = E^{new} - E$. The newly perturbed mask is accepted if $\Delta E < 0$. Otherwise, the mask is accepted with a probability $p(\Delta E) = exp(-\Delta E/T)$ where T is the temperature parameter.
Step 5: Steps 2, 3, and 4 are repeated until the system converges for a given temperature T. The system is considered to have converged if $|\Delta E|$ is less than 5% of the initial value for each iteration.
Step 6: The temperature is decreased according to the annealing schedule $T = T_0/(1+n)$, and the step number n is incremented.
Steps 2 to 6 are repeated until the NRMS error between the decrypted image and the original plaintext image is reduced to below ε.

4. Results and discussion

We started with an image A and its encrypted ciphertext ψ_A, encrypted using the Fourier plane encoding algorithm. The SA algorithm was used to estimate the key that would decrypt ψ_A with an error (NRMS error) of 0.1. The estimated key R_3 is used to decrypt a different ciphertext ψ_B corresponding to a plaintext B. The NRMS error in the decrypted image is measured. The error in the estimated image B is expected to be greater than that of A. We performed 25 trials each for two cases when image A has 32×32 and 64×64 pixels. For each trial we chose a different starting point for the SA algorithm. We used a Dell Optiplex GX280 Intel Pentium 4 CPU 2.8 GHz PC with 504 MB of RAM memory and MATLAB version 7 for our trials. The time taken for the algorithm to converge to an NRMS error of 0.1 in the decrypted image A for 25 trials is shown in Fig. 1. The NRMS error in the decrypted image B for 25 trials is shown in Fig. 2. Images A and B are shown in Fig. 3(a) and Fig. 3(e), respectively.

The average time taken in 22 out of 25 trials when A is a 32×32 pixel image is 22 minutes. The average error in decrypted image B for these 22 trials is 0.44. However trials 9, 22, and 25 have an average of 144 minutes and their average error is 0.86. When A is a 64×64 pixel image, the average time taken in 24 out of 25 trials is 343 minutes. The error in decrypted image B for these 24 trials is 0.4. The time taken for the remaining trails (trial 3 in Figs. 1 and 2) is 560 minutes and the corresponding error for the decrypted image B is 0.87.

The plaintext images A and B are shown in Fig. 3(a) and (e), respectively. The real and

#67670 - $15.00 USD Received 31 January 2006; revised 31 March 2006; accepted 3 April 2006

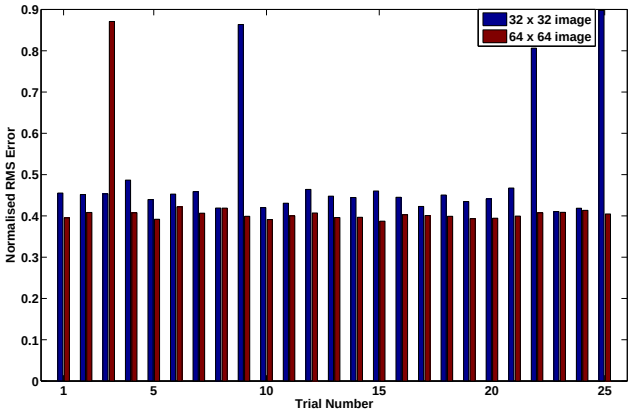

Fig. 2. NRMS error to decrypt the encrypted image of B. The key used was the one found to decrypt the encrypted image of A with NRMS error of 0.1. A and B are encrypted using the same set of keys. Results of 50 trials, 25 each for cases when plaintext is a 32×32 pixel image and 64×64 pixel image.

imaginary parts of the complex-valued encrypted image ψ_A are shown in Fig. 3(b) and (c), respectively. Figs. 3(f) and (g) show the real and imaginary part of the complex valued encrypted image ψ_B. The decrypted image of ψ_A with an error of 0.1 is shown in Fig. 3(d) and that of ψ_B with an error of 0.4 is shown in Fig. 3(h).

Of the 50 trials performed, we note that the SA algorithm converged to a solution in all cases, within 560 minutes on our particular computing platform, and found a key to decrypt ψ_A with an error of 0.1. Of these 50 trials, in 46 cases, the key successfully decrypts another unseen encrypted image ψ_B with an error of 0.4 [that is still sufficient to read the information, see Fig. 3(h)]. Only in four cases, was the error in decrypting ψ_B twice this value (approx. twice as large) and thus we regard these cases as having failed to decrypt. The existence of these four failed cases potentially adds to the security of the Fourier plane encryption technique and poses a problem for any attacker. If an attacker cannot identify such failed cases, since plaintext B is unseen, (s)he might never to able to tell, given a plaintext-ciphertext pair (A, ψ_A), whether or not a key that correctly decrypts ψ_B has been identified. However, we note that the algorithm also takes much longer to converge in these four cases. This provides a clear indication as to whether a key, which successfully decrypt ψ_B, has been found or not. Therefore our proposed use of the SA technique is as follows:

Step (a). Given the plaintext-ciphertext pair (A, ψ_A) and an unseen ciphertext ψ_B, run in parallel $s = 3$ trials of the SA algorithm.

Step (b). As soon as the first of the s trials converges, accept that key and decrypt ψ_B.

With probability 0.9995 there will be at most two failed cases out of three trials, therefore with this probability the above approach will, on average, successfully decrypt ψ_B within 3×343 minutes (given a 64×64 pixel image and our particular computing platform). We have assumed that the probability of a failed case is 4/50 = 0.08, and that our statistical sample of 50 trials in this paper is sufficient to determine this fact. Regardless of the sufficiency of our sample, if the probability of a failed case is less than 0.5 (as it certainly appears to be) running SA for $s > 3$ trials and picking the majority answer will increase even further the probability of successfully attacking the Fourier plane encryption algorithm using the known-plaintext SA heuristic attack. We have found that the average time for attack is $O(n^2)$ where 'n' is the number of pixels in the plaintext and ciphertext.

#67670 - $15.00 USD Received 31 January 2006; revised 31 March 2006; accepted 3 April 2006

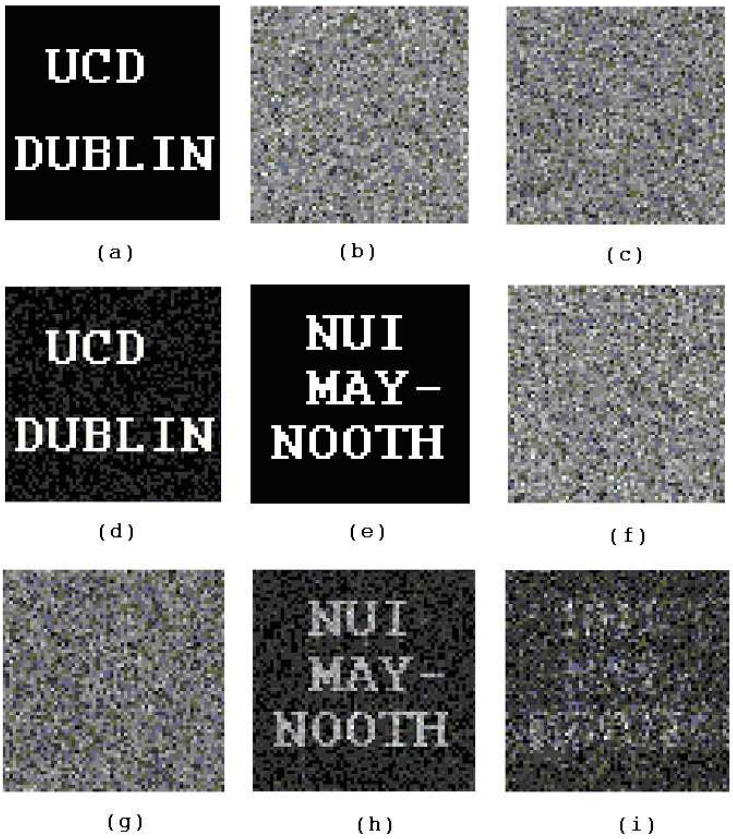

Fig. 3. (a) Images from the 64×64 pixel trials: (a) the plaintext A, (b) the real part and (c) the imaginary part of the complex-valued encrypted image of A, (d) the decrypted image with an NRMS error of 0.1, (e) the plaintext B, (f) the real part and (g) the imaginary part of the complex-valued encrypted image of B, (h) the decrypted image B with an NRMS error of 0.4, and (i) the decrypted image B in trial 3 with an NRMS error of 0.8.

5. Conclusion

We tested the strength of Fourier plane encryption algorithm with respect to a known-plaintext attack. We used an SA algorithm to estimate the key that would decrypt a ciphertext corresponding to a plaintext with a predetermined arbitrary low error. In 46 of the 50 trials, the so estimated key decrypted a different unseen ciphertext encrypted using the same original key with reasonably low error. The results from these experiments show that the Fourier plane encryption algorithm is susceptible to a known-plaintext attack with a SA heuristic. In our analysis we have assumed that the images to be encrypted are amplitude-encoded images. Furthermore, our analysis is of a mathematical formulation of the encryption algorithm itself, and does not take into account properties of the optical hardware that add to the security of the technique.

Acknowledgments

We acknowledge the support from Science Foundation Ireland.

#67670 - $15.00 USD Received 31 January 2006; revised 31 March 2006; accepted 3 April 2006