International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 / https://doi.org/10.14445/22315381/IJETT-V70I6P227

Volume 70 Issue 6, 257-271, June 2022

Original Article

Fitness Landscape Analysis of Block Ciphers for Cryptanalysis using Metaheuristics

Seeven Amic[1], K. M. Sunjiv Soyjaudah[2], Gianeshwar Ramsawock[3]

[1]*Université des Mascareignes, Rose Hill, Mauritius*
[2]*University of Mauritius, Réduit, Mauritius*
[3]*University of Mauritius, Réduit, Mauritius*

[1]samic@udm.ac.mu

Received: 28 March 2022 Revised: 14 May 2022 Accepted: 07 June 2022 Published: 29 June 2022

Abstract *- Fitness Landscape Analysis is a well-known method that has been used to evaluate or predict the performance of evolutionary or metaheuristic algorithms for tackling combinatorial optimization problems. Researchers often attempt to solve combinatorial optimization problems using metaheuristic algorithms without having proper knowledge of the underlying nature and behaviour of the problem. A similar scenario is observed in cryptanalysis using metaheuristic methods with limited or unconvincing results. There is no evidence of successful cryptanalysis of modern block ciphers such as DES or AES using metaheuristics except for toy, weakened or classical ciphers. This work aims at establishing whether metaheuristics, in general, is an efficient and promising approach to solving the problem of cryptanalysis of block ciphers. FLA has been used for this purpose. Experimental investigations reveal that the failure of cryptanalysis might be due to the high level of the ruggedness of the fitness landscape of the cryptographic keys. Furthermore, it is shown that the terrain of the fitness of cryptographic keys tends to become more rugged as the key space becomes larger, which, at this stage, tends to indicate that metaheuristic algorithms may not be very appropriate to perform rigorous cryptanalysis.*

Keywords *- Block Cipher, Cryptanalysis, Fitness Landscape Analysis, Local Optima Network, Metaheuristics*.

1. Introduction

Cryptanalysis of block ciphers can be viewed as a problem in combinatorial optimization, whereby the main goal is to search for the secret cryptographic key that has been used to encrypt the plaintext into the ciphertext. One method of black-box known-text attack of modern block ciphers is to use metaheuristic algorithms for cryptanalysis. Several works have shown that metaheuristics can be used to perform the cryptanalysis of classical ciphers [1], [2], [3], [4]. Numerous attempts to attack block ciphers with comparative results have been suggested by some authors [5], [6], [7]. However, if at all, there is very little evidence that successful cryptanalysis of modern ciphers can be achieved using metaheuristics. Most of the works evaluate and compare the efficiency of one metaheuristic algorithm with another for the problem of cryptanalysis. Metaheuristics have the natural particularity of being relatively simple algorithms and easy to implement and might be the best alternative resort when no other method works. Unfortunately, metaheuristic algorithms have inherent drawbacks: firstly, they cannot assure the optimal solution to a problem; secondly, experimental metaheuristic optimization results are irreproducible. And lastly, metaheuristics are frequently trapped in local optima leading to suboptimal solutions.

The No Free Lunch (NFL) theorem states that the overall average performance of different algorithms on a class of problems is comparably similar [8]. Consequently, comparing the performance of different metaheuristic algorithms in an attempt to tackle a combinatorial optimization problem might prove to be futile. A way must be devised to pre-establish whether a chosen algorithm can solve a given combinatorial problem. Furthermore, given the growing plethora of metaheuristic algorithms such as the Firefly algorithm or Whale optimization algorithm as chronologically enumerated in the historical survey [9], the systematic selection of the most efficient ones for a problem at hand becomes challenging.

This work performs an in-depth FLA of the key search space for two specific block cryptographic algorithms: reduced versions of DES and AES. The purpose is to determine the suitability of the algorithm for doing cryptanalysis. More precisely, the local search methodology is used to study the characteristics of the fitness landscape of cryptographic keys of block ciphers. Fitness landscape analysis can assess the complexity of an optimization problem and consequently help design algorithms with higher efficiency. Furthermore, it gives the researcher an informed insight into the hardness of the given problem based on its characteristics. It provides substance to predict the success or deceptiveness of an optimization algorithm.

The prime motivation of fitness landscape analysis is to investigate the search space from the local search perspective. Therefore, this paper is organized as follows: In Section 2, an account of the idea of fitness landscape as applied to combinatorial problems and some definitions and metrics. In Section 3, the novel complementary method of Local Optima Networks is described by examining some specific and relevant metrics for the cryptanalysis problem. The methodology section conducts an in-depth fitness landscape analysis for cryptanalysis of a few miniaturized versions of modern block ciphers. The results of the investigations are afterward reported and discussed. In the conclusion section, the findings of the current work are summarized, and the potential future research directions are proposed.

2. Background

2.1. Definitions

The definitions below will prove useful in the following subsections:

A *genotype* is a set of genes (attributes) of a given organism (object, individual, solution).

A *phenotype* is an observable characteristic (value) of a genotype of an organism. An organism's phenotype depends on its genotype, environmental (not inherited), and epigenetic (inherited) factors.

An *Individual* is an abstraction of a solution to a problem.

Fitness is a quantitative measure of the success of survival or reproduction of a phenotype.

2.2. Fitness Landscapes

The concept of fitness landscapes emanates from the field of theoretical biology. Sewall Wright [10] is the first author to introduce the idea of a fitness function defined over a collection of genotypes. The plot of all potential solutions to a given problem against their corresponding fitness is known as the fitness landscape of the problem. The term was first coined by Kauffman and Levin in [11]. Simply stated, the fitness landscape exposes the fitness behavior of the different individuals in the search space concerning each other. On a two-dimensional plot of fitness against the solutions of a search space, a line may be observed with hills, mountains, dips, and valleys, as shown in

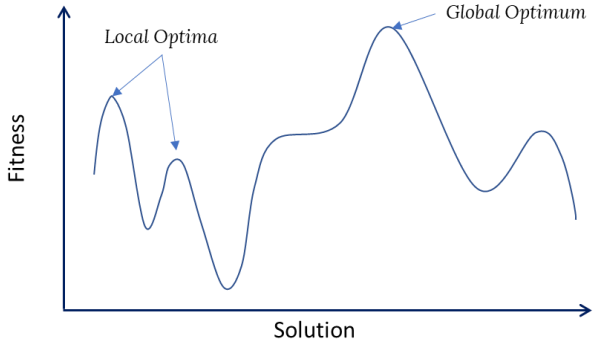


Fig. 1.

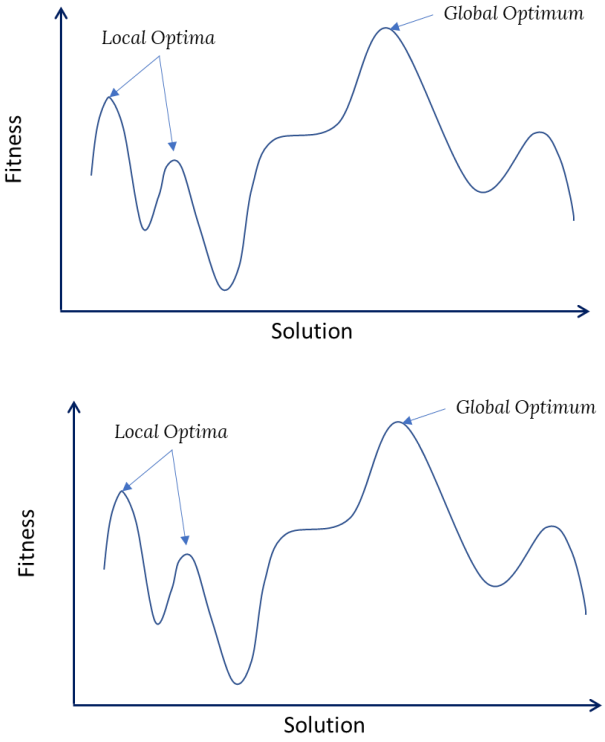


Fig. 1 Fitness Landscape

For a maximization problem, the hill peaks represent potential solutions (local optima), whereas the highest peak is the global optimum. The fitness landscape model of a problem is an abstract expression of the interrelationships among the genotype, phenotype, and fitness of solutions to the problem. Fitness landscapes represent the dynamic behaviour of evolutionary processes. Evolution is the defining factor of the success or failure of a species (a collection of individuals).

Heuristic search methods such as metaheuristics and evolutionary algorithms attempt to locate the optimal solution(s) in huge *n*-dimensional spaces of candidate solutions, where *n* is the problem's dimension. The search space may be conceptualized as a collection of points in an *n*-dimensional spatial structure. Each point represents a candidate solution and has fitness. The scatter plot of the points against fitness produces the fitness landscape surface. Interestingly, this fitness landscape structure reveals important insights into the complexity and the efficiency of the heuristic search algorithm(s) on the problem at hand.

Stadler [14] formally defines a fitness landscape composed of a triplet $(X, \mathcal{N}, f)$, where

- X – search space, the set of genotypes
- $\mathcal{N}$ – neighbourhood relation, $\mathcal{N}: X \to 2^X$
- f – objective, fitness or cost function, $f: X \to \mathbb{R}$

The fitness of a solution is a numerical measure of the quality of a potential solution to a problem. The scatter plot of phenotype (candidate solutions) against their corresponding fitness produces a fitness landscape.

It is important to emphasize the operator(s), often simple, which transform a potential solution into another, thus enabling the "discovery" of the neighbourhood of a particular solution. Two neighbours are separated by a *distance*, which a distance function can determine (e.g., Hamming distance).

The following concepts are related to fitness landscapes:

Ruggedness: A measure of roughness or unevenness of a surface. It is generally presumed that search is easier in smooth landscapes than in rugged ones.

Random Walk: A path through the search space in the neighbourhood of a particular solution, starting with an initial solution, s_0.

Hill Climbing: A procedure that traverses the search space of a problem by starting with a random solution, then iteratively improves the solution by incrementally modifying it until no further improvement is possible. There are two variants of hill-climbing: *Best Improvement* and First Improvement. The Best-Improvement hill-climbing approach authorizes a movement to the neighbour with the highest fitness in the individual's neighbourhood. The First-Improvement approach moves towards the first solution whose fitness is higher than the current one.

2.3. The essence of metaheuristic algorithms

Some combinatorial optimization problems are extremely complex to be addressed using traditional exact algorithms. In such cases, approximate algorithms provide an alternative route toward solving the problem. Metaheuristic algorithms (MA) fall under approximate algorithms that guide the search process based on some heuristic rule(s). The term "metaheuristic" was first introduced by Glover [12]. While there are several different definitions for metaheuristics, it is commonly agreed that "*a metaheuristic is a strategy that guides the search process towards a (near-)optimal solution*" [13]. MA incorporates mechanisms, namely, *intensification* and *diversification*, to avoid being trapped in local solutions. Nonetheless, MA cannot guarantee a global solution to a problem, though good-enough solutions may be obtained.

2.4. Fitness Landscape Analysis

During the last decade, the analysis of fitness landscapes has been applied extensively in performing the following [15]:

(a) comprehension of complex problems
(b) analysis and interpretation of algorithm behaviour
(c) algorithm performance prediction
(d) parameter configuration of algorithms
(e) automated algorithm selection

FLA helps to establish whether the solutions in a search space are related and whether bonds, structures, or patterns are observed among them. FLA is a huge resource-consuming activity for moderate to large complex problems. Consequently, performing a one-time FLA to gain problem insights is expensive and probably futile. However, FLA is highly relevant for obtaining a deeper understanding of problems falling into the NP-hard class [16].

3. Fitness Landscape Analysis Metrics

An FLA can be conducted using two complementary approaches: Statistical and Informational. Malan et al. [17] identified 22 techniques for analysing fitness landscapes based on ruggedness, modality, and evolvability metrics. Recently, Malan [15] extended the survey with 11 additional methods for landscape analysis ranging from Local Optima Networks to loss-gradient clouds. Some of the enumerated techniques provide metrics that may help assess the effectiveness of metaheuristics in solving a class of problems. However, the assessment of the hardness of a problem based only on one metric might be too risky. On the other hand, too many metrics for such a task could be confusing and misleading. Furthermore, the efficiency of a metaheuristic approach depends on the selection of the particular algorithm, its proper parameter calibration, and the characteristics of the problem at hand.

3.1. Statistical Metrics

In statistical FLA, the following metrics are useful to obtain a coarse view of the hardness of the problem based on its fitness landscape.

3.1.1. Fitness Distance Correlation

Jones introduced Fitness Distance Correlation (FDC) as a measure of problem difficulty [18]. It characterizes the correlation between fitness values and their corresponding distance to the optimal global solution. The FDC can be calculated by Equation (1).

$$FDC = \frac{\frac{1}{n}\sum_{i=1}^{n}(f_i-\bar{f})(d_i-\bar{d})}{\sigma_F \sigma_D} \qquad (1)$$

where n is the total number of solutions, f_i is the fitness of the ith solution and d_i Represents the distance of the ith solution to the global solution. σ_F represents the standard deviation of fitness whereas σ_D is the standard deviation of distance. The Equation may be adapted for a representative sample of random individuals. Jones [18] formulated that the difficulty of a problem could be deduced from the calculated value of its FDC according to Table 1.

Table 1. Fitness Distance Correlation

Fitness Distance Correlation	Meaning
$\text{FDC} \leq -0.15$	"Easy"
$-0.15 < \text{FDC} < 0.15$	"Difficult"
$\text{FDC} \geq 0.15$	"Deceptive"

3.1.2. Fitness Autocorrelation Coefficient

The fitness autocorrelation coefficient, ρ, is the correlation between adjacent fitnesses and measures the extent of ruggedness present within a fitness landscape [19]. Autocorrelation is the correlation coefficient that characterizes a relationship between values of the same data series at a given interval (called a *lag*) in a time series. It can be calculated by recording the fitness of visited

solutions during a random walk of arbitrary length in the search space using Equation (2).

$$\rho(lag) = \frac{\sum_{i=1}^{n-lag}(f(x_i)-\bar{f})(f(x_{i+lag})-\bar{f})}{\sigma^2(f(x_i))} \qquad (2)$$

where $f(x_i)$ is the calculated fitness of the solution x_i,

$\bar{f}$ – the mean fitness in the random walk of n steps,

lag – the number of steps ahead of the current solution,

σ^2 – the variance in fitness.

A random walk of $n \approx 1000$ steps is sufficient for a good estimate of the autocorrelation coefficient [19].

The *autocorrelation length*, τ, is the reciprocal of the autocorrelation coefficient with lag = 1. The autocorrelation length is the distance beyond which solutions become uncorrelated.

$$\tau = -\frac{1}{\rho(1)} \qquad (3)$$

A small value of τ indicates a rugged landscape, whereas a long τ means a smooth landscape.

3.2. Entropy Metrics

Statistical Fitness Landscape Analysis metrics based on correlation permit the study of fitness landscapes but give only a coarse picture of the landscape's structure. Additional information is required to gain a detailed understanding of the landscape. Vassilev et al. [20] propose a set of entropy (or information) measures to define the topography of a fitness landscape by analyzing random walks over the landscape. The analysis of a random walk of n steps on a landscape $\mathcal{L}$ produces a sequence of fitness values $\{f_t\}_{t=0}^{n}$, which holds features about the landscape structure. This information can be represented as a string $S(\varepsilon) = s_1 s_2 \cdots s_n$, where $s_i \in \{\bar{1}, 0, 1\}$ as defined in Equation (4).

$$s_i = \Psi_{f_t}(i,\varepsilon) = \begin{cases} \bar{1}, & if\ f_{i-1} - f_i > \varepsilon \\ 0, & if\ |f_i - f_{i-1}| \le \varepsilon \\ 1, & if\ f_i - f_{i-1} > \varepsilon \end{cases} \qquad (4)$$

Where ε is the difference in fitness between adjacent points on a path and $\varepsilon \in \mathcal{I}$. $\mathcal{I}$ is the range of fitness values. The symbol $\bar{1}$ characterizes a downward slope (↘) from the previous fitness to the next, 0 represents a flat transition (→), whereas 1 denotes an upward slope (↗). The choice of ε affects the accuracy of the derivation of string $S(\varepsilon)$. The string $S(\varepsilon)$ can be viewed as a sequence of pairs of characters, $s_i s_{i+1}$I.e., a substring of length 2. The value of ε should be chosen less than ε^*, the *information stability*. The information measures proposed by Vassilev et al. [20] are described in the following subsections. The salient advantage of Information FLA is that it is based upon random walks, and the exhaustive analysis of the whole search space is not required. It makes the technique particularly attractive, especially for large search spaces.

However, this proposed statistical analysis has an inherent drawback as it presumes that the landscape is statistically isotropic [38]. The statistical information derived from the random walks is assumed to be independent of the walk's starting point and depends only on the genotype distance. Any adequately long walk in any direction is assumed to yield similar results. Unfortunately, this assumption does not always hold for constrained combinatorial problems.

3.2.1. Information Content

The *information content*, $H(\varepsilon)$, of a walk on a fitness landscape provides a measure of the amount of entropy in the structure of the landscape and can be calculated using Equation (5).

$$H(\varepsilon) = -\sum_{p \neq q} P_{[pq]}\, log_6\, P_{[pq]} \qquad (5)$$

$P_{[pq]}$ is the probability of the block pq in the information string. $n_{[pq]}$ is the frequency of pq in $S(\varepsilon)$, and n is the number of substrings of length 2. The logarithm in Equation (5) is off base 6 because $\#pq = 6$. Hence the information content is normalised in the range [0, 1]. Information content measures the percentage of divergence in sequential fitness values for the chosen parameter, ε. Hence, it pertains to the ruggedness, smoothness, and neutrality of the landscape's surface.

3.2.2. Partial Information Content

The partial information content, $M(\varepsilon)$, of a walk on the landscape is a characteristic that is directly associated with the number of local optima as it characterizes the slope changes in the walk. $M(\varepsilon)$ is calculated by Equation (6).

$$M(\varepsilon) = \frac{\mu}{n} \qquad (6)$$

where μ is the length of $S'(\varepsilon)$, a special string derived from $S(\varepsilon)$. String $S'(\varepsilon)$ is obtained by removing all 0's and duplicating adjacent values from $S(\varepsilon)$. Hence, $S'(\varepsilon)$ is obtained by eliminating unimportant parts of $S(\varepsilon)$. The length of $S'(\varepsilon)$ is μ, the modality of the path on the landscape.

The number of optima of a path ℓ of a landscape with partial information content, $M(\varepsilon)$, can be calculated as :

$$\#optima(\ell) = \left\lfloor \frac{nM(\varepsilon)}{2} \right\rfloor \qquad (7)$$

When a landscape path is flat, i.e., without any slopes, then the partial information content $M(\varepsilon)$ equals 0. When $M(\varepsilon)$ is equal to 1, then the landscape path contains the highest number of optima.

3.2.3. Landscape Modality

The modality of a fitness landscape is characterised by the number of local optima and their distribution in a search space. A landscape with only one optimum is called unimodal, whereas bimodal with two optima and multimodal with many optima. It is still unknown how

landscape modality affects search complexity; however, it is established that unimodal landscapes may be difficult to search [22], and multimodal landscapes may be adapted to improve search efficiency [23]. Landscape modality is also associated with the magnitude of the basins of attraction of the optima. The degree of isolation of an optimum solution is determined by the number of solutions present within its basin of attraction. Hence, a larger basin of attraction leads to a smaller degree of isolation and vice versa. Furthermore, evolutionary search may present different levels of difficulty for landscapes with equal number of optima [20].

3.2.4. Density-Basin Information

The *density-basin information* is a landscape feature that assesses the amount of neutral and smooth sections on a landscape [20]. For a landscape path expressed as a string $S(\varepsilon)$, the density-basin information is calculated using Equation (8).

$$h(\varepsilon) = -\sum_{p\in\{\bar{1},0,1\}} P_{[pp]}\, log_3\, P_{[pp]} \qquad (8)$$

The log of base 3 is taken because there are three blocks of interest, namely, $\bar{1}\bar{1}11$ and 00. The higher the density-basin information, the more neutral or smoother the landscape.

3.2.5. Information Stability

The difference, ε, between adjacent fitnesses on a path determines the accuracy of the information content and the partial information content on a landscape. The minimum value of ε, denoted by ε^*, for which the string $S_i(\varepsilon)$ Becomes all zeros is known as the *information stability* shown in Equation (9).

$$\varepsilon^* = min\{\varepsilon, S_i(\varepsilon) = 0\} \qquad (9)$$

Conceptually, the information stability is the largest difference between consecutive neighbours in a random walk. Information stability, ε^*, characterizes ruggedness, smoothness, and neutrality of fitness landscapes. Large values of ε^* Corresponds to highly rugged landscapes, whereas smaller values relate to smooth ones.

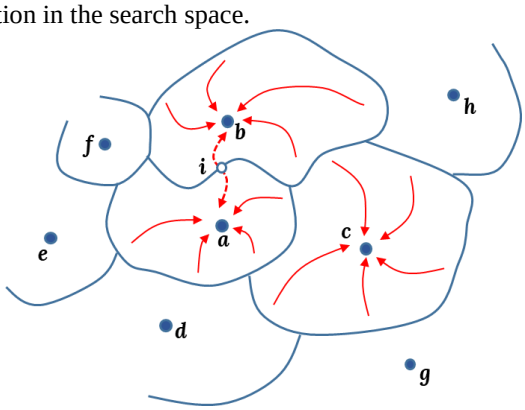
4. Local Optima Networks

Network structures show interconnections among entities in a system. Lately, much attention has been devoted to studying and analyzing networked systems such as transportation, professional or social interaction, and biological systems. Statistical features and metrics of such networks provide deeper insight and understanding of how such systems operate and reveal approaches to optimize them. This work investigates the use of network structures to study the fitness landscapes of cryptographic algorithms, particularly local optima networks.

The Local Optima Networks (LONs) techniques offer an alternative model of fitness landscapes based on graphs [24] and are inspired by the analysis of landscapes of energy in theoretical chemistry [25]. LONs model a search space as a weighted directed network whose nodes represent local optima, whereas edges display potential transitions among optima. The size of the nodes corresponds to the size of the basin of attraction. The weight of the directed edges gives a quantitative measure of the probability of a potential transition between two connected optima. This approach is useful but requires an exhaustive treatment of the basins of attraction and produces a densely connected network. Another approach is to consider escape edges [26], in which the edges' weight accounts for the probability of escaping a local optimum by hill-climbing following a controlled mutation. One achieves a controlled mutation by flipping one or two bits in the current individual in the search space.

It is now a fact that the analysis of local optima networks of certain combinatorial problems shows striking correlations between the LON features and the search difficulty of the problem.

A basin of attraction is a region where all the solutions lead towards the same local optimum (the attractor) and are subjected to a hill-climbing operation. The number of solutions, including the optimal solution, that fall into this area is known as the size of the basin of attraction. The basin of attraction surrounding optimum local yields a partition in the search space.



shows a partial search space with local optima surrounded by their respective basin of attraction. The boundary of a basin of attraction is a set of solutions that have at least one neighbour in another basin. During landscape traversal, boundary elements allow the transition between basins of attraction with a certain probability within a search space. It is known as basin transition.

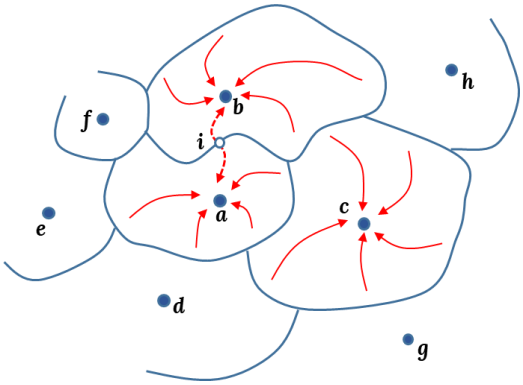


Fig. 2 Basin of Attraction with Local Optima

Another approach is to treat the weight of the edges as representing the probability of escaping the basin of attraction of a local optimum after a mutative operation (of one or two-bit flips in a binary space) followed by a local search step (hill-climbing). These edges are known as escape edges and do not require an exhaustive evaluation of the basins of attraction [26].

Consequently, the LON model consisting of a weighted and directed graph permits the analysis and evaluation of the global terrain of a fitness landscape by identifying the local optima in the search space and examining the transitions among them.

4.1. General Characteristics of Weighted Networks

A network (or graph) has N nodes (also known as vertices) and M edges (links) between pairs of nodes. Each node is numbered, $j = 1, \dots, N$. The structure of a network is represented by an $N \times N$ matrix $A = \{a_{ij}\}$ and an edge that connects node i to j has weight w_{ij}. Sometimes, the weights are normalized so that they fall in the range [0, 1] by dividing all the weights by the largest weight; i.e., the normalized weights are $w_{ij}/\max(w_{ij})$. If the direction of the link between nodes is important, then a directed graph is formed. The LON metrics of interest are summarized in Table 2.

Table 2. Local Optima Network Metrics

Notation	Description
fit	The average fitness of local optima
$zout$	Average out-degree
wii	The average weight of self-loops
$y2$	The average disparity for out-going edges
knn	Weighted assortativity
wcc	Average weighted clustering coefficient
fnn	Fitness-to-fitness correlation

The definitions of the main and auxiliary LON metrics are given below.

The *vertex strength* s_i is a measure of the importance, or centrality, of the vertex v_i is calculated as the total weight of its immediate neighbours using Equation (10).

$$s_i = \sum_{j=1}^{N} a_{ij} w_{ij} \qquad (10)$$

The *vertex degree*, $k_i I$s the number of connected vertices to the vertex. For a directed graph, the number of incoming connections is known as in-degree, k_i^{in}And the number of out-going connections is known as the out-degree, k_i^{out}. Intuitively, $k_i = k_i^{in} + k_i^{out}$.

The *average fitness of the local optima* gives an estimate of the overall quality of the optima. The higher the average fitness, the larger the basins of attractions, and the higher the chance for a successful search process.

A self-loop to a basin indicates that starting with a local optimum followed by a series of random moves ends within the same basin. If the weight of the self-loop is high, then the probability that a random move ends within the same basin is high. If the *average self-loop weight* is high, then the search process will likely be trapped in local optima basins.

The *disparity* measures the diversity of the weight proportions of the edges of node *i* to the total weight and is calculated using Equation (11).

$$Y_2(i) = \sum_{j \in V(i)} \left[\frac{w_{ij}}{s_i}\right]^2 \qquad (11)$$

where $s_i = \sum_{j \neq i} w_{ij}$ Termed as the strength of the i^{th} vertex. This parameter reveals whether there are preferential directions when a walk leaves a given node in the network.

The *weighted assortativity* is a measure of preference for a node in a network to attach to others based on a particular feature(s). The assortativity coefficient of a network is defined as the Pearson correlation coefficient of vertex degree between pairs of connected vertices, as shown in Equation (12). It measures the tendencies of nodes of a network to connect with other nodes [27].

$$k_{nn,i}^{w} = \frac{1}{s_i} \sum_{j=1}^{N} a_{ij} w_{ij} k_j \qquad (12)$$

where s_i – the strength of ith node

The assortativity of a weighted and directed graph, *G(V, E)*, can be calculated using the Equation (13) based on the sample Pearson correlation coefficient

$$\rho_{X,Y}(G) = \frac{\sum_{i,j \in V} w_{ij} (X_i - \bar{X}_{source})(Y_j - \bar{Y}_{target})}{W \sigma_X \sigma_Y} \qquad (13)$$

where

$$W = \sum_{i,j \in V} w_{ij} \qquad (14)$$

When a network mixes assortatively (i.e., with a positive assortativity), then its high-degree nodes tend to group, forming a subnetwork with a higher mean degree than the rest of the network [28] hence enhancing higher connectivity and reachability among vertices of the whole network. The assortativity coefficient of a network ranges from -1 to +1. When closer to -1, it indicates that two nodes of similar properties might not be related. When its value approaches +1, there exists a high likelihood that two nodes of similar properties are connected.

The *average weighted clustering coefficient*, *wcc*, of a network is a measure of the tendency of nodes of the graph to a group. The overall average weighted clustering coefficient can be calculated using Equation (15).

$$wcc = \frac{1}{n} \sum_i c_i^w \qquad (15)$$

The *weighted clustering coefficient*, c_i^w The node in a network is a measure of local cohesiveness ("cliquishness") of a neighbourhood based on the degree

and strength of the node formed by a triplet (triangle) of nodes i, j, and h. It is calculated by Equation (16).

$$c_i^w = \frac{1}{s_i(k_i-1)} \sum_{j,k} \frac{(w_{ij}+w_{ih})}{2} a_{ij} a_{ih} a_{jh} \qquad (16)$$

The *fitness-to-fitness correlation metric*, fnn, gives the average correlation between fitness values of connected local optima of the whole network. The fitness-to-fitness correlation of a given vertex can be calculated using the Spearman rank correlation coefficient between the fitness of the i^{th} vertex and the weighted average of its neighbourhood vertices.

$$f_{nn,i}^w = \frac{1}{s_i} \sum_{j \neq i}^{n_i} w_{ij} f_j \qquad (17)$$

where n_i is the neighbours of the i^{th} vertex, and s_i.is the strength of the i^{th} node.

5. Methodology

In broad terms, this endeavor aims to apply FLA to investigate the pertinence and performance of metaheuristic algorithms in the cryptanalysis of block ciphers. The procedure is split into two steps. Firstly, the general Fitness Landscape Analytic metrics are evaluated to assess the difficulty of the block-cipher cryptanalysis problem. The second step consists of modelling the fitness landscape using the Local Optima Network approach to extract additional features and information about the fitness landscape. Consequently, it is anticipated that sufficient insights will be gained from the fitness landscape of cryptographic keys of a few selected block ciphers to understand, assess and improve cryptanalysis by metaheuristics. The general framework for the fitness landscape analysis is shown in Fig. 3.

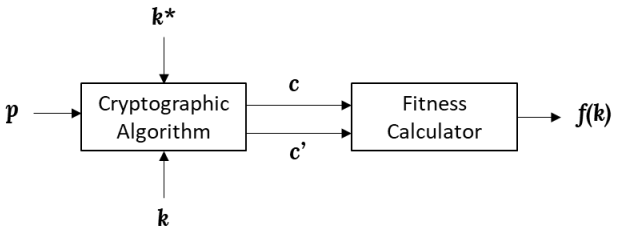


Fig. 3 Fitness Landscape Framework Setup

This work analyzes simple block ciphers with small keys (8 to 16 bits). Such tiny-key ciphers are cryptographically weak but are complex enough to mimic the architecture of their stronger versions. Consequently, the exhaustive set of keys may be generated and their fitness calculated using the fitness calculator, $f(k)$:

$$f(k) = \frac{1}{d \cdot n} \sum_{i=1}^{n} \#(e(k^*, p_i) \odot e(k, p_i)) \qquad (18)$$

where d is the data block size and n is the number of plaintext/ciphertext pairs available for known text cryptanalysis using k^*It is the actual key. $e(k, p)$ is the ciphertext bitstring obtained after the encryption of plaintext p using key k. $\odot$ are the $XNOR$ operator, and # counts the number of bits set to 1. The data and key can be represented as bit strings per their cryptographic specifications.

It is important to note that for a given block cipher, the fitness landscape is expected to be static, i.e., it does not change each time it is generated. However, in our case, because the fitness of a cryptographic key is based on random plaintext-ciphertext pairs, it could be different, though with an insignificant difference, if calculated with a different collection of plaintext-ciphertext pairs.

The known-text cryptanalytic attack strategy is used whereby, for a selected block cipher, a finite collection of n plaintext/ciphertext pairs are created by making use of *the* (selected) key k^* Being cryptanalysed. The following cryptographic schemes were analysed:

Table 3. Cryptographic Schemes under study

Cipher	Acronym	Key Size	Block Size
Simplified AES [29]	AES8	8	8
Simplified DES [30]	DES10	10	8
Simplified AES [29]	AES12	12	12
Mini AES [31]	AES16	16	16

5.1. Exhaustive Fitness Landscape metrics

For the Simplified DES and AES versions listed in Table 3, the exhaustive set of keys, k, in the search space is generated, and for a specifically selected key, k^*The fitness for each key, $f(k)$, is calculated and stored in an array. The fitness of $k^*$is the 1.0. First, an arbitrary key is selected as the cryptographic key somewhere midway in the search space, and the remaining keys are plotted against their calculated fitness. Furthermore, the frequency distribution of the fitness among the keys can also be plotted. These two plots shall give a general visualization of the distribution of the fitness of the keys over the entire search space.

The above experiment is executed several times with different chosen cryptographic keys. Ideally, every key could be designated as the encryption key and systematically generate the exhaustive array of corresponding fitness values. However, only a few randomly selected keys are used as the cryptographic key, one at a time, while others act as potential keys (solutions). The average fitness-distance correlation and the average fitness autocorrelation coefficients are calculated for a sample of selected keys.

5.2. Local Optima Network Analysis

The LON analysis starts with constructing the LON, which consists of identifying the local optima (the vertices) and establishing the connections (the edges with corresponding weights) among them. The edges can be either escape edges or basin transition edges. An escape edge exists between two local optima if there is some solution s, which is at most at a distance D steps away from one optimum. Still, it can reach the other optimum if subjected to a controlled mutation (one or two-bit flips) followed by a hill-climb. Two local optima have a basin

transition if at least one solution belongs to both basins of attraction. The weight of the edge between the two optima is the number of solutions that satisfy the criteria mentioned. Implementing the connections with escape edges is privileged instead of basin transition edges, as the former yields a less dense LON [32]. J. E. Fieldsend [33] has proposed a computationally efficient LON construction Java package, which has proved to be exceptionally fast with large LONs. This package has been adapted for the LON construction for the cryptanalysis problem. Additionally, for each identified local optimum, the fitness value and corresponding basin size are computed and recorded in a CSV file for further processing in the network analysis software.

Since the search space of this problem is relatively small, the exhaustive set of local optima (vertices) is identified using the algorithm given in Fig. 4. The algorithm for extracting the escape edges is given in Fig. 6, and the auxiliary hill-climb operation is shown in Fig. 5.

Algorithm: Extract LON Vertices

def $V \leftarrow \{\}$ **as** the list of vertices
foreach $x \in X$ do
  $isOptimum \leftarrow true$
  $i \leftarrow 1$
  while $x_i \in \mathcal{N}(x)\ AND\ isOptimum$ **do**
    if $f(x) < f(x_i)$ **then**
      $isOptimum \leftarrow false$
    endif
    $i \leftarrow i + 1$
  end while
  if $isOptimum$ **then**
    $V \leftarrow V \cup \{x\}$
  end if
end for
return V

Fig. 4 LON Vertices Extraction Algorithm

Algorithm: $hillclimb(x, X, f, \mathcal{N})$

def v **as** local optimum
$v \leftarrow x$
foreach $u \in \mathcal{N}(x)$ **do**
  if $f(u) > f(v)$ **then**
    $v \leftarrow u$
  endif
end for
if $v \neq x$, **then**
  $v \leftarrow hillclimb(v, X, f, \mathcal{N})$
end if
return v

Fig. 5 Hill-climb Algorithm

Algorithm: Extract Escape Edges

def V **as** a set of local optima
def n **as** $\#V$
def m **as** neighbourhood Hamming distance
def $A[1..n][1..n] \leftarrow 0$ as an adjacency matrix
def $B[1..n] \leftarrow 0$ as basin size vector
foreach $x \in X$ **do**
  $v \leftarrow hillclimb(x, X, f, \mathcal{N})$
  $B[v] \leftarrow B[v] + 1$
end for
foreach $v \in V$ do
  $u \leftarrow hillclimb(v, X, f, \mathcal{N}_m)$
  $A[v][u] \leftarrow A[v][u] + 1$
the end for
return A, B

Fig. 6 Escape Edges Extraction Algorithm

Once the vertices are identified, an adjacency list of the resulting network is constructed. The edges and corresponding weights are calculated and stored in a CSV file. The list of optima (nodes) and edges can be imported into network analysis environments for metrics extraction and visualisation. In this work, a Java program was developed to calculate some LON statistics, and the *R* environment with the network analysis library *igraph* was used for visualisation. For some metrics like *disparity* and *assortativity*, the standard functions in *igraph* and *R* language were used.

6. Results and Discussion

6.1. Statistical FLA Analysis

The scatter plots of the cryptographic key against fitness for the selected ciphers are given in Fig. 7. The chosen cryptographic key with fitness 1.0 is marked with a star. It can be observed that the fitness is spread around the average fitness (0.5). The standard deviation of the fitness decreases with the increase of the size of the keys, as shown in the top left corner of Fig. 9. Hence, the fitness of the keys tends to be closer to the overall mean for larger search spaces. It follows that the evolvability of the keys is greatly compromised in the band of fitness that does not contain any key. It can be observed more clearly in the scatter plots for AES12 and AES16 in the fitness region between 0.7 and 0.99 in Fig. 7 (c) and Fig. 7 (d).

The results of the frequency distribution of fitness of keys are given in Fig. 8. As observed from the bar charts, the plots are skewed to the right (lower fitness), with the peak at the 0.5-0.6 band as the key size increases. It reveals that the larger the key search space, the percentage of higher quality keys decreases. In other words, the gap between the average solutions and the global optimum contain fewer or no solutions. Hence, there is less likelihood of average solutions evolving into the global optimum, thus making the search more difficult.

The entropic metrics obtained from the experiments are summarized in Table 4 and depicted in the plots in Fig. 7.

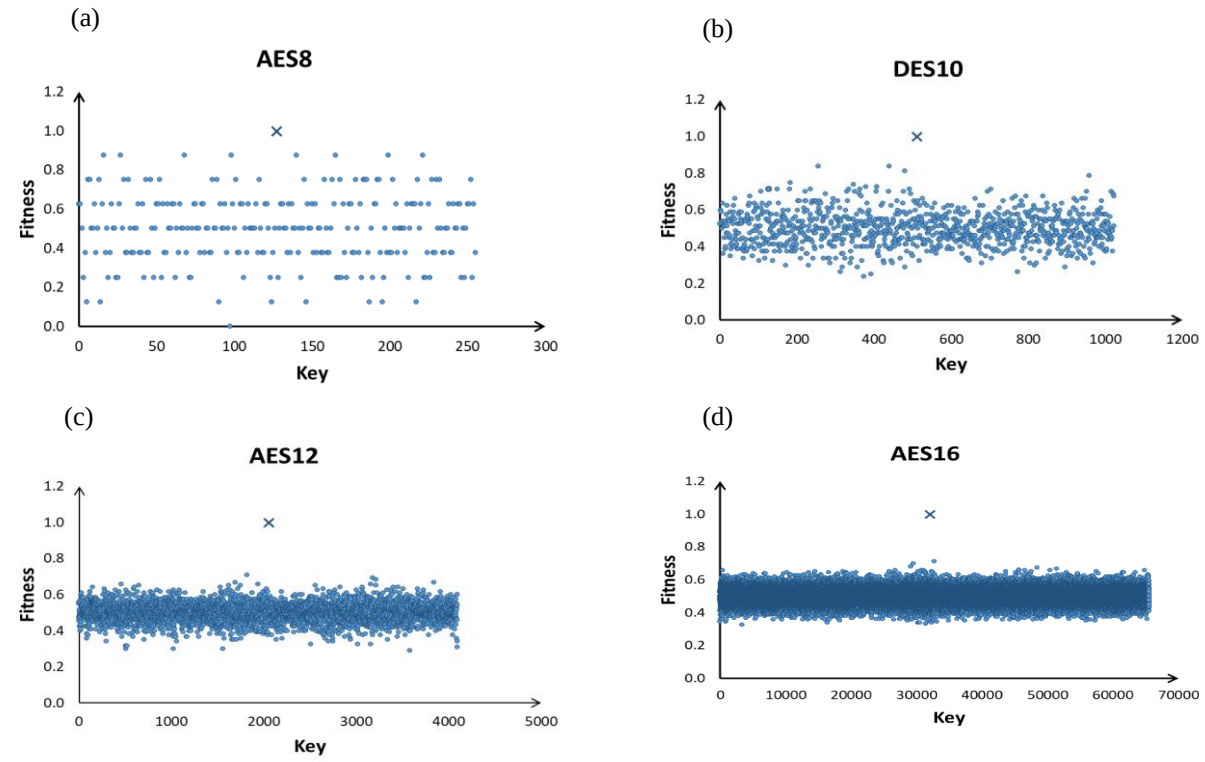


Fig. 7 Scatter plots of Fitness vs. Key value

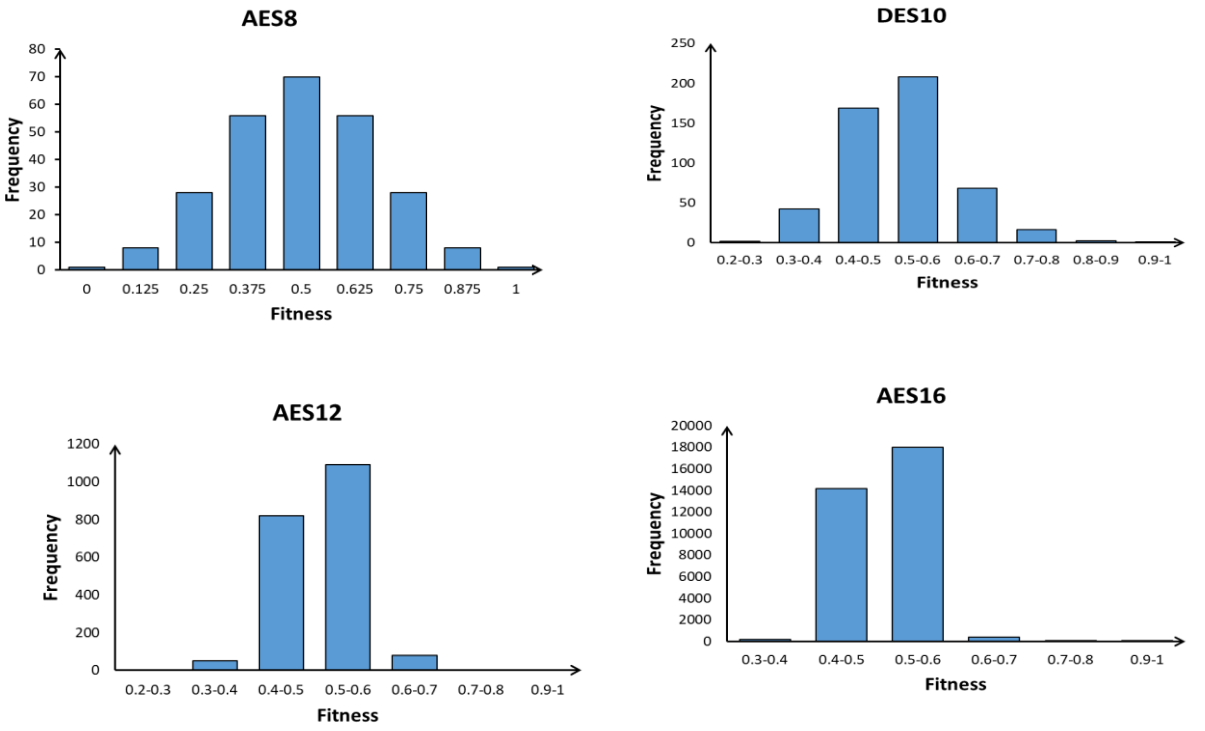
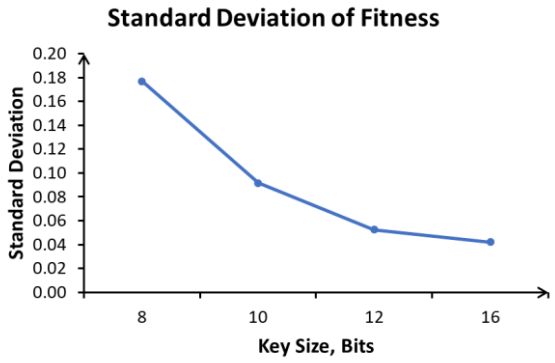


Fig. 8 Frequency Distribution of Fitness of

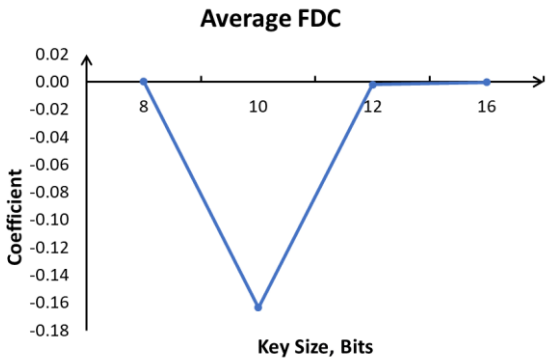
Table 4. Entropic Fitness Landscape Metrics

	Cipher			
Fitness Landscape Metric	**AES8**	**DES10**	**AES12**	**AES16**
Standard Deviation of Fitness, σ_f	0.1768	0.0915	0.0524	0.0420
Average FDC, r	0.0	-0.163627	-0.002126	-0.000386
Average Autocorrelation, ρ	-0.186586	0.423415	0.176619	0.001227
Information Content, $H(\varepsilon)$	0.551901	0.426946	0.427924	0.430992
Partial Information Content, $M(\varepsilon)$	0.659381	0.686394	0.730106	0.714205
Density-Basin Information, h	0.198301	0.268162	0.243319	0.252595
Information Stability, ε^*	0.657267	0.309933	0.254533	0.174633

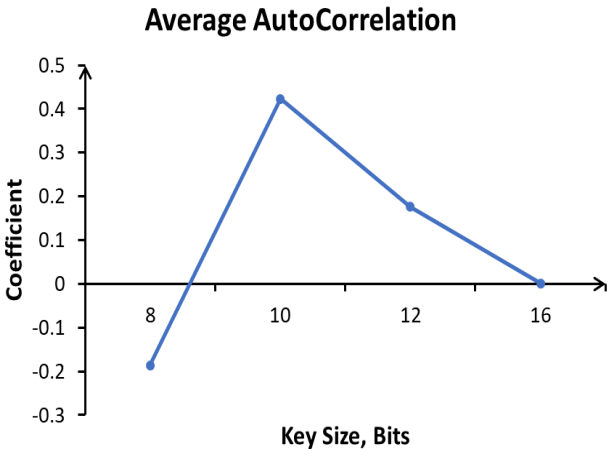
(a)



(b)



(c)



(d)

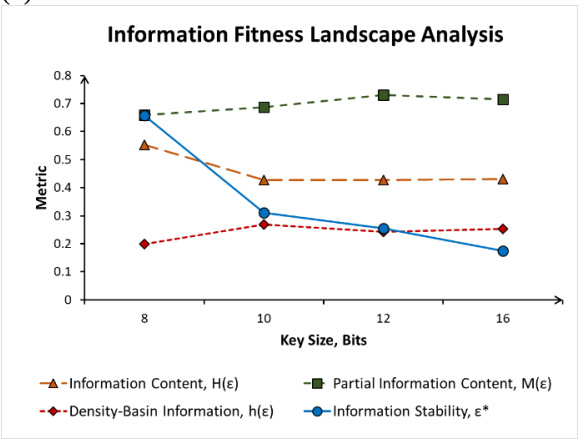


Fig. 7 Fitness Landscape Analysis Metrics

From Fig. 9(a), it is observed that the standard deviation of fitness gradually decreases with key size. Hence, as the key size of the ciphers increases, the tendency of the fitness of the keys to wrap closer around the mean. The average FDC remains approximately zero for all ciphers except for DES10, as shown in Fig. 9(b). According to the rough classification based on FDC proposed by Jones [18], the performance prediction of metaheuristics on cryptanalysis of ciphers is summarised in Table 5. Hence, the cryptanalysis of DES10 is predicted to be "Easy," while the other ciphers are categorised as "difficult."

Table 5. Fitness Distance Correlation Coefficient

Cipher	**FDC, *r***	**Range**	**Difficulty Class**
AES8	0.0	$-0.15 < r < 0.15$	Difficult
DES10	-0.163627	$r \leq -0.15$	Easy
AES12	-0.002126	$-0.15 < r < 0.15$	Difficult
AES16	-0.000386	$-0.15 < r < 0.15$	Difficult

The results observed while performing cryptanalysis using brute force (BF) attack and a few selected metaheuristic algorithms; namely, Simulated Annealing

(SA), Firefly Algorithm (FA), Tabu Search (TS), and Genetic Algorithm (GA), are summarised in ."

Table 6. The number of generations/loops required for cryptanalysis was collected and averaged over 1000 runs for each cipher and metaheuristic method. The chosen ciphers were cracked relatively rapidly using the selected metaheuristic algorithms, although the performance prediction for reduced AES was categorised as "difficult."

Table 6. Average number of generations for cryptanalysis

Cipher	BF	SA	FA	TS	GA
AES8	128.0	321.13	16.62	7.65	3.72
DES10	512.0	1134.35	35.55	27.42	3.47
AES12	2048.0	4787.24	261.04	106.48	42.84
AES16	32768.0	70292.17	3325.45	1705.90	686.31

Here, " difficult " performance prediction would mean "difficult, but not impossible." This outcome could be explained by the relatively small size of the search space of the cryptographic keys being studied, whereby the global optimum could have been visited by mere luck.

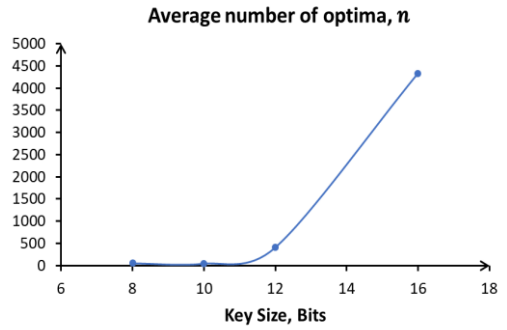
The average autocorrelation plot in Fig. 9(c) shows no observed trend with key size. The Information Fitness Landscape Analysis metrics are displayed in Fig. 9(d). The information content, partial information content, and density-basin information metrics are approximately constant with an increase in key size. However, the information stability decreases gradually with key size. It implies that the fitness landscape would paradoxically tend to be smoother with the increase in key size.

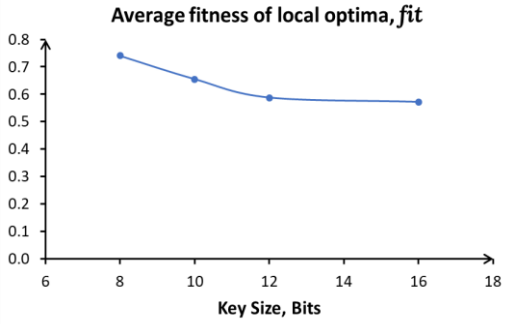
6.2. FLA using LONs

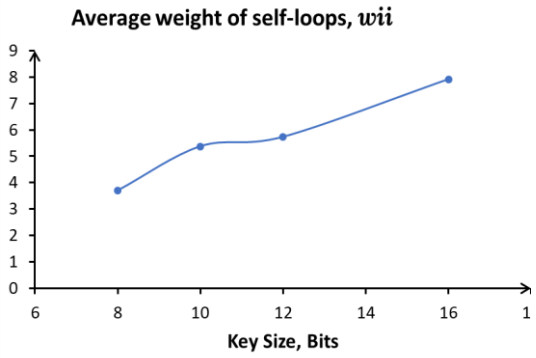
The results of FLA based on LONs are listed in Table 7, and their corresponding plots are displayed in Fig. 8. The trends displayed by the LON metrics concerning key size can be visually observed from the graphs in Fig. 10.

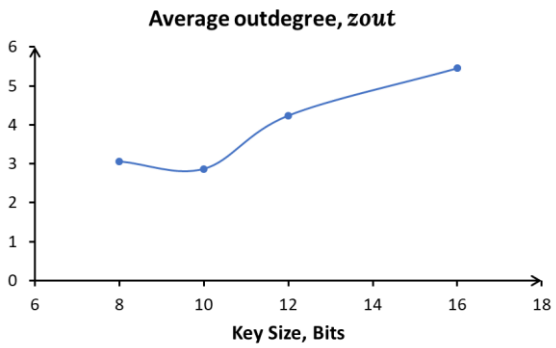
Table 7. LON Metrics

	Cipher			
LON Metric	**AES8**	**DES10**	**AES12**	**AES16**
Average number of optima, n	51	45	409.04	4333.5
Average fitness of local optima, fit	0.740196	0.655	0.587541	0.571542
The average weight of self-loops, wii	3.705882	5.377778	5.735611	7.916896
Average out-degree, $zout$	3.058824	2.866667	4.238763	5.457782
Average disparity for outgoing edges, $y2$	0.019608	0.111111	0.114047	0.087281
Average weighted assortativity, knn	5.941176	6.133333	8.343162	10.60166
Average weighted clustering coefficient, wcc	0.169406	0.082901	0.026459	0.027991
Average fitness-fitness correlation, fnn	0.657286	0.719186	0.562187	0.54449









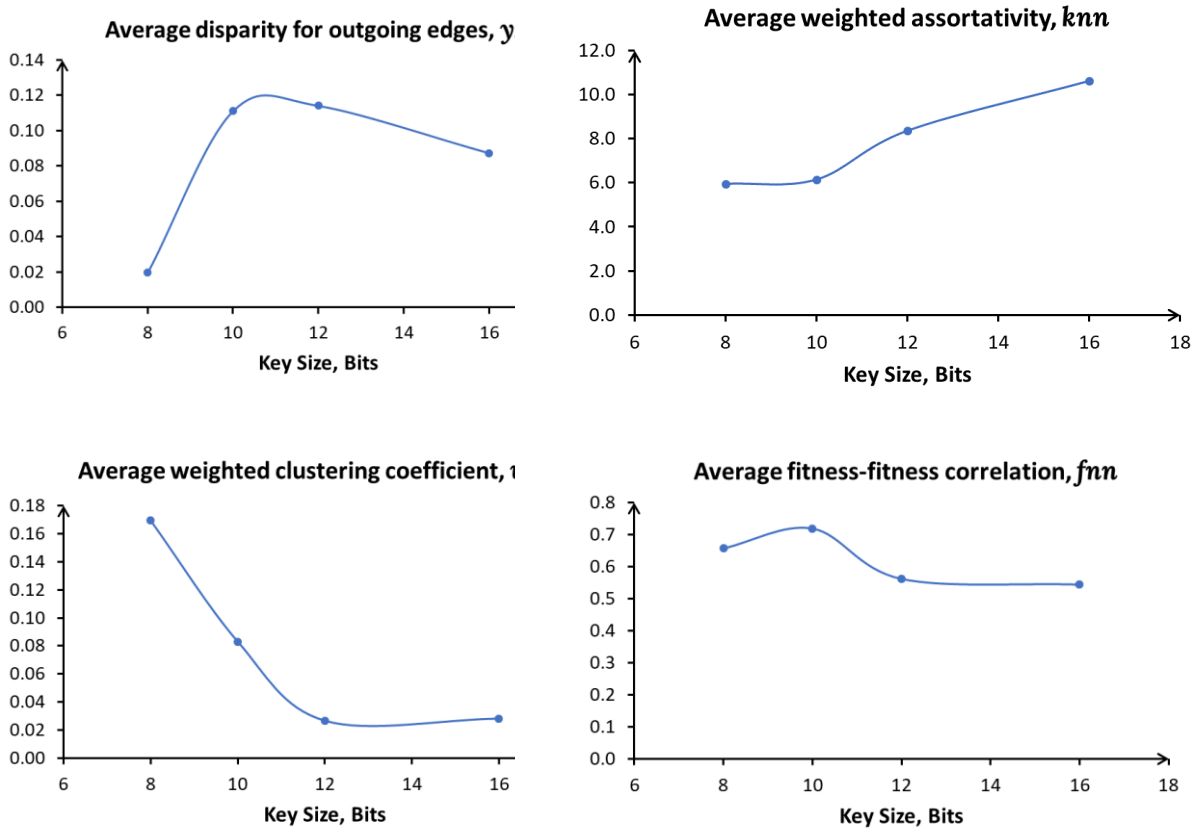


Fig. 8 LON metrics vs. Key Size

From the results of the experiments, it is observed that the average number of optima, n, increases exponentially as the key size becomes larger. A larger number of local optima renders the search space more rugged.

Furthermore, the average fitness, fit, of the local optima decreases with larger key sizes resulting in lower quality solutions during hill-climbing operations. It follows that for ciphers with larger key sizes, the fitness of the keys shall tend to be lower. The same result has been observed in the fitness distribution of keys in Fig. 7.

Larger key size ciphers result in a larger average weight of self-loops, wii. The growing weight of self-loops entails that local search shall have a higher probability of being trapped in the local optima basin, which hinders the convergence of a metaheuristic search algorithm towards the global optimum.

The increasing average out-degree, $zout$, of nodes in the LONs indicates that the number of possible escape edges from each node increases with key size. The average disparity for out-going edges, $y2$, does not show a distinctive trend. Its value is observed to be positive for all key sizes, which means that the fitness of local optima in a chain of local optima tends to increase. However, the disparity of each cipher is of low magnitude (<0.12).

The average weighted assortativity, knn, is an increasing function concerning the key size. It means network vertices with similar properties are more likely to be highly connected. The average weighted clustering coefficient, wcc, also known as transitivity, of a weighted and directed graph, measure thus the average strength among all the vertex-triplets forming a network triangle [34]. It hence provides another measure of the connectedness of vertices in a network. The average fitness to fitness correlation, fnn, shows little variation as the key size grows, meaning that the key size does not affect any correlation among local optima.

The resulting LONs for each cipher using a selected key (for each) are depicted in Fig. 9. The bubbles represent the nodes (local and global optima) of the LON, while the edges show the possible transitions between the nodes. For clarity of the diagrams, the weights of the edges have not been displayed. The diameter of a node represents the size of its basin of attraction, i.e., the number of neighbourhood elements it contains. The colour palette of the nodes ranges from dark red (dark grey in grayscale) for nodes with the highest fitness to bright yellow (light grey) for nodes with lower fitness. Incidentally, the size of the basin of attraction of a node is directly proportional to its fitness.

The visualisation of LONs with a small number of local optima (Fig. 9 (a) and Fig. 9 (b)) allows the manual

tracing of the paths from remote nodes to nodes with higher fitness (or higher size of the basin of attraction). The sources and sinks can also be seen on the graphs in Fig. 9 (a) and Fig. 9 (b). However, the "hairball" LONs of larger search spaces shown in Fig. 9 (c) and Fig. 9 (d) is denser and more complex. The global optimum is less visible, if at all, and the colour gradation among local optima is less distinctive. It follows that the local optima have approximately similar fitness.

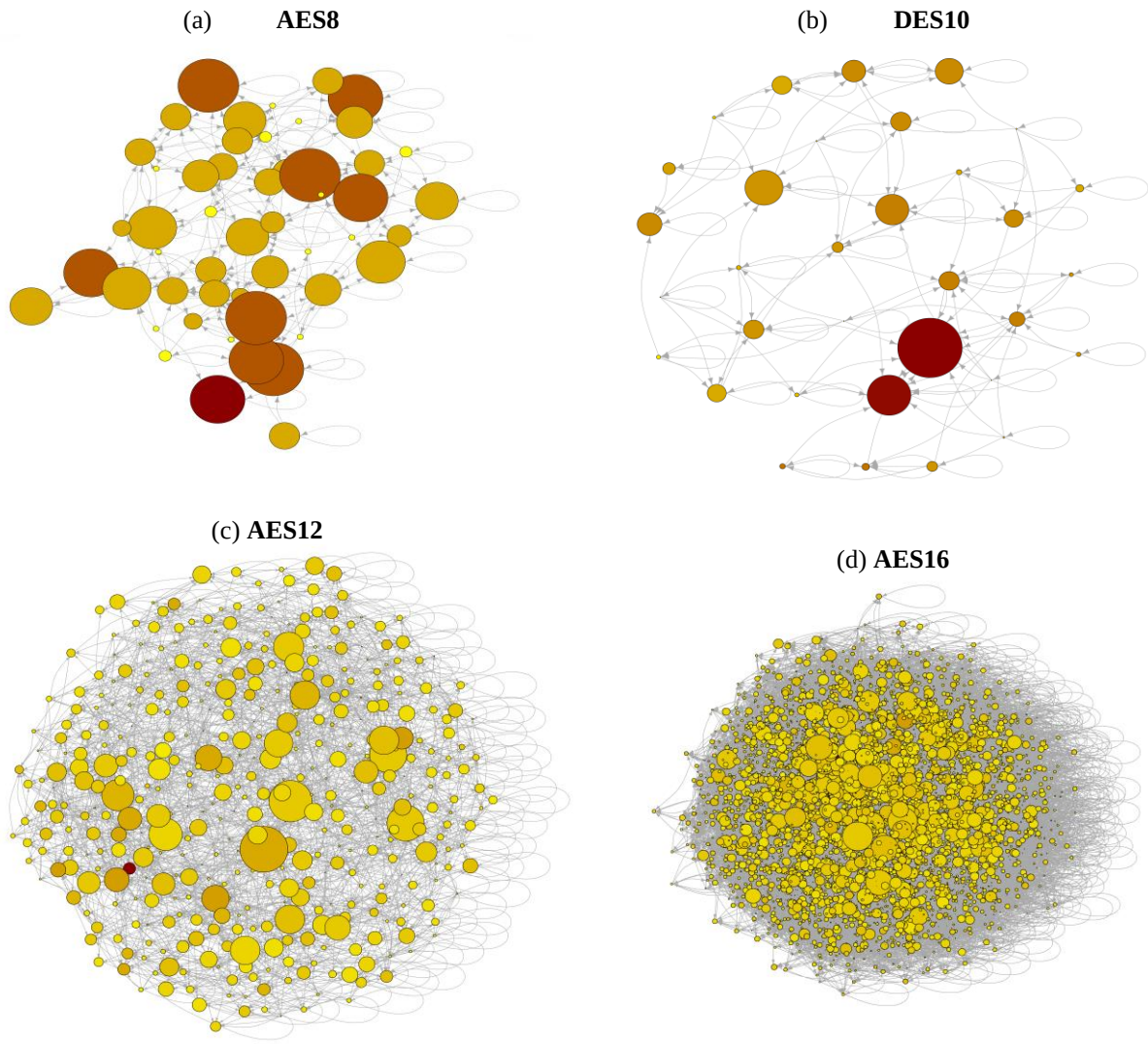


Fig. 9 Visualisation of LONs

7. Conclusion

In this paper, an unprecedented formal FLA has been conducted to study the fitness landscape of cryptographic keys of simple ciphers to gain insight into their landscape ruggedness. Consequently, this information is used to predict the hardness of search optimisation using metaheuristic algorithms. In the first part of the work, statistical and informational FLA have been conducted based on the work of Jones et al. [18] and Vassilev et al. [20], respectively. Results show that cryptographic key landscapes are rugged even for small key search spaces, which leads to the conclusion that cryptanalysis using metaheuristic algorithms is generally "difficult." Furthermore, the present study demonstrates that the ruggedness of cryptographic fitness landscape increases with the key size. The ruggedness of cryptosystems with larger keys is expected to be higher, hence harder to cryptanalyze using metaheuristic techniques.

In the second part of the work, a recent novel technique known as LONs for FLA is used to analyse relationships and connectivity among local optima within a network of cryptographic keys. These methods provide a unique prospect to perform an exhaustive local search and obtain a better grasp of the interrelated features and metrics of the fitness landscape. From a mathematical perspective, LONs "reduce" the search space from an exhaustive enumeration of all the potential solutions to a smaller set of local optima and their interrelated features.

To summarize, FLA has been successfully used to show that the cryptanalysis of miniaturized versions of DES and AES is categorised as “difficult” due to their highly rugged fitness landscape. FLA metrics based on resulting LONs provide additional insights into the complexity of the search problem and show trends, which can predict the metaheuristic search performance on block ciphers.

8. Future Work

The statistical information (entropy) and network analyses of fitness landscapes of cryptographic keys provide useful additional insights into the ruggedness of fitness landscapes. FLA provides an effective method to predict the efficiency of metaheuristic algorithms on the cryptanalysis problem. However, the methodology covered in this work requires an exhaustive enumeration of potential solutions for statistical analysis, which can be difficult for very large search spaces. Furthermore, the number of local optima for network analysis can nonetheless be too large for efficient processing despite being much less in number than when compared to an exhaustive search.

Due to memory and processing constraints, it is hard to perform exhaustive FLA on modern ciphers, like DES and AES, with key sizes larger than 32 bits. One potential solution to tackle this problem is to examine the use of statistical sampling of a large population of keys and look into the sampling of LONs for FLAs. The major challenges with this approach involve the efficiency of the sampling techniques and the sample representativeness of the exhaustive key space. The FLA of modern ciphers based on statistical and informational metrics using samples instead of the entire population is relatively straightforward. The statistical formulae for fitness-distance and fitness-autocorrelation coefficients can be modified for sample data. The informational metrics based on entropy do not require any special adaptation as they involve hill-climbing walks of a finite length. For the sampling of LONs, two promising techniques have been proposed, namely, the Snowball Sampling method [35], [36] and the Markov-Chain LON sampling approach [37]. Initial work has already been started in this direction, and the preliminary outcomes indicate to be promising. Based on the results of the current work coupled with the sampling techniques, it is expected that the FLA of modern block ciphers like DES and AES can be achieved to forecast the potential performance of metaheuristic algorithms on their cryptanalysis.

References

[1] R. Toemeh and S. Arumugam, Applying Genetic Algorithms for Searching Key-Space of Polyalphabetic Substitution Ciphers, Int. Arab J. Inf. Technol., 5(1) (2008) 87–91.

[2] A. Bhateja, S. Kumar, and A. K. Bhateja, Cryptanalysis of Vigenere Cipher Using Particle Swarm Optimization with Markov Chain Random Walk, Int. J. Comput. Sci. Eng., 5(5) (2013) 422–429.

[3] S. S. Omran, A. S. Al-Khalid, and D. M. Al-Saady, A Cryptanalytic Attack on Vigenere Cipher Using Genetic Algorithm, 2011 Ieee Conf. Open Syst., 1 (2011) 59–64.

[4] P. Saveetha, Cryptography and the Optimization Heuristics Techniques, 4(10) (2014) 408–413.

[5] J. Song, H. Zhang, Q. Meng, and Z. Wang, Cryptanalysis of Four-Round Des Based on Genetic Algorithm, Int. Conf. Wirel. Commun. Netw. Mob. Comput. Wicom , (2007) 2326–2329.

[6] T. Mekhaznia, Nature Inspired Heuristics for Attack of Simplified Des Algorithm, Sin 2013 - Proc. 6th Int. Conf. Secur. Inf. Networks, (2013) 311–315.

[7] W. Shahzad, A. B. Siddiqui, and F. A. Khan, Cryptanalysis of Four-Rounded Des using Binary Particle Swarm Optimization, Simulation, (2009) 1757–1758.

[8] D. H. Wolpert and W. G. Macready, No Free Lunch Theorems for Optimization, Ieee Trans. Evol. Comput., 1(1) (1997) 67–82.

[9] S. M. Almufti, Historical Survey on Metaheuristics Algorithms, Int. J. Sci. World, 7(1) (2019) 1.

[10] S. Wright, the Roles of Mutation, Inbreeding, Crossbreeding and Selection in Evolution, In the Sixth International Congress of Genetics, 1 (1932) 356–366.

[11] S. Kauffman and S. Levin, Towards a General Theory of Adaptive Walks on Rugged Landscapes Section of Ecology and Systematics , and Ecosystems Research Center , New York, 128(1) (1987) 11–45.

[12] F. Glover, Future Paths for Integer Programming and Links to Artificial Intelligence, 13(5) (1986) 533–549.

[13] C. Blum and A. Roli, Metaheuristics In Combinatorial Optimization: Overview and Conceptual Comparison, Acm Comput. Surv., 35(3) (2003)189–213.

[14] P. F. Stadler and S. F. Institute, Towards a Theory of Landscapes, Complex Syst. Bin. Networks, (2008) 78–163.

[15] K. M. Malan, A Survey of Advances in Landscape Analysis for Optimisation, Algorithms, 14(2) (2021) 40.

[16] E. Pitzer and M. Affenzeller, A Comprehensive Survey on Fitness Landscape Analysis, Stud. Comput. Intell., 378 (2012) 161–191.

[17] K. M. Malan and A. P. Engelbrecht, A Survey of Techniques for Characterising Fitness Landscapes and Some Possible Ways Forward, Inf. Sci. (Ny)., 241 (2013) 148–163.

[18] T. Jones and S. Forrest, Fitness Distance Correlation As A Measure of Problem Difficulty for Genetic Algorithms, In Proc.\ of 6th Int.\ Conf.\ on Genetic Algorithms, (1995) 184–192.

[19] E. Weinberger, Correlated and Uncorrelated Fitness Landscapes and How to Tell the Difference, Biol. Cybern., 63(5) (1990) 325–336.

[20] V. K. Vassilev, T. C. Fogarty, and J. F. Miller, Information Characteristics and the Structure of Landscapes, Evol. Comput., 8(1) (2000) 31–60.

[21] Ramanathan.L and Ulaganathan.K, Nature-Inspired Metaheuristic Optimization Technique-Migrating Bird"S Optimization in Industrial Scheduling Problem Ssrg International Journal of Industrial Engineering,1(2) (2014) 12-17.Crossref, Https://Doi.Org/10.14445/23499362/Ijie-V1i3p101
[22] J. Horn and D. E. Goldberg, Genetic Algorithm Difficulty and the Modality of Fitness Landscapes, Second Edi., Morgan Kaufmann Publishers, Inc.,3 (1995).
[23] L. Altenberg, the Schema Theorem and Price's Theorem, (1995) 23–49.
[24] G. Ochoa, S. Verel, F. Daolio, and M. Tomassini, Local Optima Networks: A New Model of Combinatorial Fitness Landscapes, (2014) 233–262.
[25] J. P. K. Doye, Network Topology of a Potential Energy Landscape: A Static Scale-Free Network, Phys. Rev. Lett., 88 (23) (2002) 4.
[26] S. Vérel, F. Daolio, G. Ochoa, and M. Tomassini, Local Optima Networks with Escape Edges, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), 7401 (2012) 49–60.
[27] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani, the Architecture of Complex Weighted Networks, Proc. Natl. Acad. Sci. U. S. A., 101(11) (2004) 3747–3752.
[28] M. E. J. Newman, Assortative Mixing in Networks, Phys. Rev. Lett., 89(20) (2002) 1-5.
[29] P. T. Breuer, An 8- and 12-Bit Block Aes Cipher. Academia.Edu, (2013) 1–33.
[30] Edward F. Schaefer, A Simplified Data Encryption Standard Algorithm, Cryptologia, 20(1) (1996) 77–84.
[31] R. C. Phan, Mini Advanced Encryption Standard (Mini-Aes), Publ. Cryptologia, 26(4) (2002).
[32] S. Verel, F. Daolio, G. Ochoa, and M. Tomassini, Local Optima Networks with Escape Edges, In International Conference on Artificial Evolution (Ea-2011), (2011) 10–23.
[33] J. E. Fieldsend, Computationally Efficient Local Optima Network Construction, Gecco 2018 Companion - Proc. 2018 Genet. Evol. Comput. Conf. Companion, (2018) 1481–1488.
[34] M. Barthélemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani, Characterization and Modeling of Weighted Networks, Phys. A Stat. Mech. Its Appl., 346(1-2) Spec. Iss (2005) 34–43.
[35] S. Verel, F. Daolio, G. Ochoa, and M. Tomassini, Sampling Local Optima Networks of Large Combinatorial Search Spaces: the Qap Case, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), 11102 (2018) 257–268.
[36] S. L. Thomson, G. Ochoa, S. Verel, and N. Veerapen, Inferring Future Landscapes: Sampling the Local Optima Level, Evol. Comput., 28(4) (2019) 621–641.
[37] S. L. Thomson, G. Ochoa, and S. Verel, Clarifying the Difference in Local Optima Network Sampling Algorithms, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), 11452 (2019) 163–178.
[38] G. W. Greenwood and X. Hu, are Landscapes for Constrained Optimization Problems Statistically Isotropic?, Phys. Scr., 57(3) (1998) 321–323.