



Feature selection for steganalysis using artificial bee colony algorithm

Neha Singh

Assistant Professor, Department of Computer Science, Indria Priyadarshini College for Women, Jhajjar, Maharshi Dayanand University, Haryana, India

Abstract

The process of steganalysis comprises of two major steps: extraction of features and classification based on the features that have been extracted. The classification step of steganalysis process is tortuous and time-consuming as the feature sets are of high dimensions. This paper proposes a swarm optimization technique- Artificial Bee Colony (ABC) algorithm for feature selection. The proposed ABC feature selection algorithm selects reduced feature set. The Artificial Bee Colony, is a swarm optimization technique based on swarm intelligence nature of honeybees, that has been adopted and modified successfully for steganalytic feature selection in this work. ELM is used as the classifier and Fisher score is used as measure of separability. The experiment is conducted on two different set of JPEG images of 256 x 256 size each. To show the adequacy and potency of the algorithm, the experiments is conducted on four different types of steganalytic feature sets. The results of the experiment shows that the dimensionality of the features has reduced efficaciously and the detection accuracy of the steganalysis process has enhanced when the proposed feature selection method is used. The code is implemented in Matlab.

Keywords: steganalysis, feature selection, swarm intelligence, artificial bee colony

1. Introduction

Steganography is the field of hiding information. The term steganography is derived from the two Greek words “stego” and “grafia”. Stego means cover and grafia means writing, thus defining it as “covered writing”. Steganography is an information concealing technique which hides the secret message as well as the presence of the message is hidden from the third party^[1]. Suppose a sender wants to send a secret message, say ‘m’, then the sender selects an innocuous message ‘c’, called as the cover-object. By using a key, say ‘k’, called as the stego- key, the sender embeds (embedding process) the secret message into c which can then be transmitted to the receiver without any speculation. The sender has hence changed the cover object c into stego object s. This whole process of changing the cover into stego is done with cautions, so that any third party cannot detect the presence of hidden message in the stego object s. On other hand, if the receiver has the stego key and knows the embedding algorithm he can easily reconstruct (extraction process) the secret message m^[2]. The embedding and extraction process are given as:

$$\begin{aligned} \text{Embedding process-} & \quad E: & C \times K \times M \rightarrow C \\ \text{Extraction process-} & \quad D: & C \times K \rightarrow M \end{aligned}$$

Where,

C: the set of possible covers

K: the set of possible keys

M: the set of possible messages

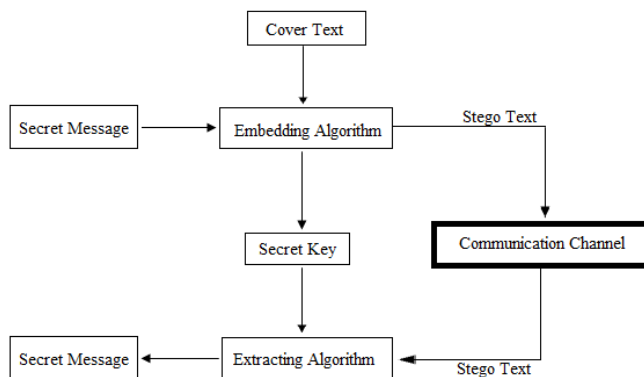


Fig 1: steganography process

Steganography can be categorized as text, image and audio steganography. *Text steganography*^[3] is applied in digital makeup format such as PDF. The clandestine embedding of data into digital images is known as *image steganography*^[3, 4]. Although information can be hidden in any digital media, digital images are the most popular cover objects because of the fact that they have large size, includes large amount of redundant bits (so, large amount of information can be stored) and are highly used over the internet. When secret information is embedded into digital sound, it is known as *audio steganography*^[3]. Steganalysis^[5] is the art of discovering the messages hidden using steganography. If a image contains a hidden message then the image is called a cover- image otherwise it is called as stego- image (images does not includes any hidden information).

Classification of steganalysis:

Based on whether the steganalysis algorithm is dependent the steganographic algorithm (SA), steganalysis is classified as follows ^[6]:

1. Specific / Target steganalysis.
2. Generic / Blind / Universal steganalysis.

In *Specific steganalysis*, the steganalysis algorithm is dependent on the steganographic algorithm (SA). Those statistical properties of an image that changes after the embedding process are analyzed in this steganalysis. The results of specific steganalysis are very precise but these can only be used with some specific embedding algorithms and image formats.

When the steganographic algorithm is not known to all, the steganalysis is called as *Blind / Universal steganalysis* ^[6]. Consequently, to find the existence of the concealed message, anyone can build a detector that is not dependent on the SA. Although, blind steganalysis are less coherent but they are popular because they are free of the SA. This paper emphasis on universal steganalysis. It includes the following 2 phases:

a. *Feature Extraction*: It is a process of creating a set of distinct statistical attributes of an image. These attributes are known as feature. Feature Extraction is nothing but a dimensionality reduction. The extracted features must be sensitive to the embedding artifacts. Image quality metrics, wavelet decomposition, moment of image statistic histograms, Markov empirical transition matrix, moment of image statistic from spatial and frequency domain ^[7], co-occurrence matrix ^[8, 9, 10] are some of the feature extraction methods.

b. *Classification*: Depending on the feature values of the image they are categorized into different classes; namely stego and non- stego images. Supervised learning is one of the primary classifications in steganalysis.

The process of steganalysis involves two phases: training phase and testing phase. In the training phase, the features which are sensitive to embedding are extracted of both cover and stego- image. The classifier is trained by these selected features. In the testing phase, the test data set is given to classifier to distinguish between stego and non- stego images.

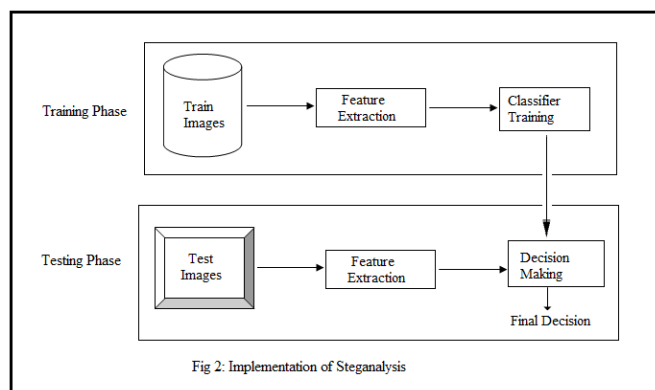


Fig 2: implementation of Steganalysis

Steganalysis Methods: Based on the way, the existence of concealed message is detected, steganalysis methods are divided as follows:

1. Statistical steganalysis.

- a) Spatial domain steganalysis
 - b) Transform domain steganalysis
2. Feature based steganalysis.

In Statistical steganalysis, for detecting the existence of the secret hidden message analysis of the pixels is done. It is further categorized as spatial domain steganalysis and transform domain steganalysis. In *spatial domain*, the difference between any two pair of pixels (within a block or over two blocks) is determined. Finally the histogram is plotted that shows the existence of the hidden message. In *transform domain*, to distinguish between the stego and non-stego (cover) images the frequency counts of coefficients are computed and thereafter, histogram analysis is conducted. The main snag of this method is that they don't provide any information about the embedding algorithms. In *Feature based steganalysis* ^[11], the features of the image which are affected after steganography are extracted. These extracted features tell whether the image contains any hidden message. These feature are also used to train the classifiers. This paper focuses on feature based steganalysis.

The layout of the rest of the paper is: Section 2 briefly accounts of swarm optimization and swam intelligence and introduces ABC. Section 3 presents ABC algorithm for feature selection. Section 4 comprises of experimental setup and results followed by conclusion in section 5.

2. Swarm Optimization and Swarm Intelligence

The word swarm is used in a general manner to refer to any controlled collection of cooperating agents. So *Swarm Intelligence* ^[12] is collective intelligence of a group of simple agents. The field of swarm intelligence has been used for redesigning algorithms or extensive problem solving devices taken by the collective behavior of social insect colonies and other animal societies concentrated on social insects lonely in their researches such as bees, termites, wasps and ants. Often a single agent is not competent enough to able to find its food source but when the intelligence, ability and skill of multiple such agent is combined, they are able to find the source not only quickly but the food source so found is also good in quality, quantity and have a shorter distance that need to be travelled by the agents. Swarm Intelligence has been exploited to solve various optimization problem. The art of using Swarm Intelligence to solve various optimization problem is known as *Swarm Optimization* ^[13] that is when intelligence of a group of agents is utilized and exploited to optimize the underlying objective.

In this paper, Swarm Intelligence is used for feature selection problem because the objective that need to be optimized is to reduce the search space that is the number of feature that can classify the class of image with as high accuracy as possible. The swarm optimization techniques discussed in this paper is Artificial Bee Colony.

Artificial bee colony

Artificial Bee Colony exploits the swarm intelligence nature of bees ^[14]. There are two kinds of bees: employed foragers and unemployed foragers. An *employed forager* is employed at a specific food source which she is currently exploiting. She carries information about this specific source and shares it

with other bees waiting in the hive. Unemployed foragers are the bees that look for a food source to exploit. It can be either a scout who searches the environment randomly or an onlooker who tries to find a food source based on the information provided by the employed bee.

Every cycle of the bees search encompasses three phases: employed bee phase, onlooker bees phases and scout bees phase. In the employed bee phase, the employed foragers find the best food source in their environment and share the information about this food source (its quality, quantity and distance from hive) with other bees (onlooker) waiting in the hive. In the onlooker bee phase, the onlooker bee goes to their respective food source and tries to find a better food source in the neighborhood of their food source. In the scout bee phase, the employee bees whose corresponding food source is abandoned becomes the scout and randomly surf the environment to find a food source. Hence the collective communicating nature of bees help them in getting best food source in their search environment.

3. Feature Selection using Artificial Bee Colony

Artificial Bee Colony (ABC) [15] is a stochastic, population-based evolutionary method proposed by Karaboga in the year 2005. In ABC algorithm, the colony of artificial bees consists of two kinds of bees [16]:

1. Employed foragers: An employed forager is employed at a specific food source which she is currently exploiting. She carries information about this specific source and shares it with other bees waiting in the hive.
2. Unemployed foragers: A forager bee that looks for a food source to exploit is called unemployed. It can be either a scout who searches the environment randomly or an onlooker who tries to find a food source based on the information provided by the employed bee.

The quantity of employed bees is equivalent to the quantity of food sources, which is equivalent to quantity of unemployed bees. In ABC algorithm, the actual position of a food source corresponds to a probable solution to the optimization problem. Every cycle of the bees search encompasses three phases: employed bee phase, onlooker bees phases and scout bees phase. Each employed bee is assigned to an individual food source. Each bee colony has scouts that are the colony's explorers. They have no guidance while searching food. The scout bee can be chosen and classified from one of the employed bees. The classification is handled by a control component called "limit". If an answer representing a food source is not enhanced by a pre-arranged number of trials, that food source is left through the employed bee, then that

food source is abandoned by its employed bee and the employed bee associated with that food source turns to a scout. The quantity of trials for releasing a food source is equivalent to the importance of "limit", that is a significant control component of ABC algorithm. Wrapper method [16] is used for feature selection and ELM [17, 18] is used as classifier to get the accuracy of bit strings.

The three phase of ABC:

Employee bee phase

A feature vector is a binary bit string of dimension equal to the amount of features which is given to per employed bee to act for its feature selection. In this, bit "1" declares that the feature is a selected feature and bit "0" declares the feature is not selected.

Each employed bee is assigned to each bit string. Each employed bee then checks in the neighborhood of the food source assigned to it whether a better food source exists or not. If a food source exist whose fitness is better than the one assigned to it, the employee bee memories that food source eliminating the previous food source. All employee repeat this process. After this the information gathered by all the employee bee is used to calculate probabilities of each food source thus sharing their information with onlooker bees.

Onlooker bee phase

An onlooker bee selects a topmost food source based on its probability value (these probabilities are calculated using information shared by employee bee). Each onlooker bee then try to explore the neighborhood of their corresponding food source in search of a better food source. If a food source with better fitness is found, all necessary updation are made by the onlooker bee. All the onlooker bees work in this manner.

Scout bee Phase

In this phase, first the abandoned food is discovered. For a food, when the limit exceeds a certain threshold it is considered as abandoned. Each cycle of ABC algorithm determines only one abandoned food. Then the employee bee corresponding to the food source that is abandoned becomes the scout bee. This bee randomly surf the environment and picks up one food source, which then tries to improve on this food source. If a better food source is found, it makes updation otherwise not.

These three phases are repeated until some stopping criteria is reached. The best food source (bit string with maximum classifier's accuracy) is the solution.

The ABC algorithm is described in figure:

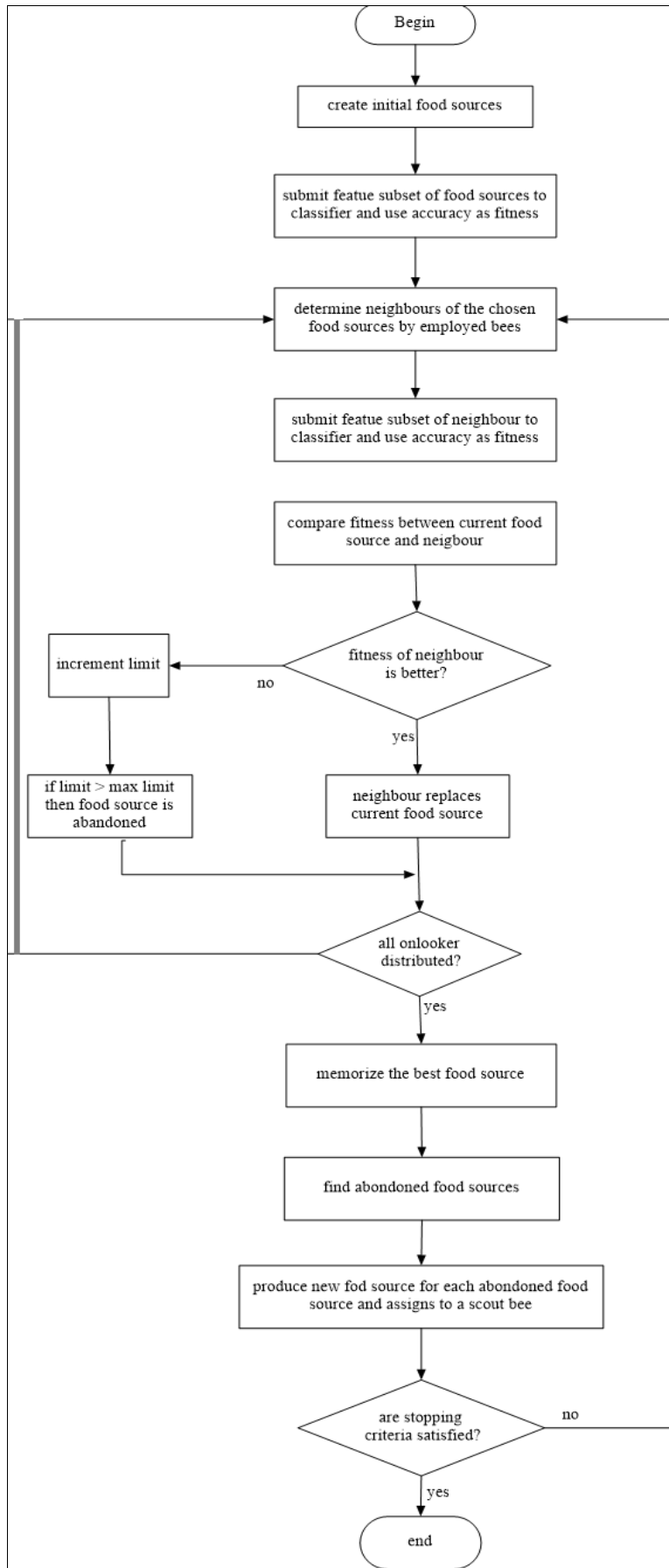


Fig 3: ABC Algorithm

4. Experimental Setup and Result

For the experiment, two different data set consisting of both stego and non- stego images are considered. One data set consists of 1098 images and other consists of 1892 images. 256 x 256 JPEG images are used which are freely available on internet. JPEG image Steganographic tools: OutGuess, F5 and Model-based steganography [8, 21]. are used to create the stego-images. The image dataset is divided randomly into non-overlapping training set and testing set.

Different feature extraction method [22], namely: co-

occurrence in DCT coefficient [10], neighborhood joint density based features (absolute), neighborhood joint density based features (normal) [11, 19], markov based [20] features are used for extracting the feature of the images. The extracted feature matrix is then supplied to ABC which in turn gives the best feature subset. The accuracy of the full feature set obtained using different feature extraction method are calculated using ELM. Similarly accuracy of the feature returned by ABC is calculated. The computations are shown in the Table 4.1.

Table 1: Performance of ABC

Feature Extraction Method	No. of features	Accuracy (%) using original features	No. of features selected by ABC	Accuracy (%) using selected features
Co-occurrence in DCT coefficient	120	81.42	12	94.50
NJD (absolute)	72	75.97	12	78.31
NJD (normal)	338	78.82	187	85.71
Markov Based	486	81.26	296	87.83

5. Conclusion

The paper presented an Artificial Bee Colony algorithm, a swarm optimization techniques for feature selection for steganalysis. Two different set of JPEG images, each of size 256 x 256 were used for the experiment. Extreme Machine Learning (ELM) was used as the classifier. To create the stego images different JPEG image Steganographic tools were used namely: OutGuess, F5 and Model-based steganography. To show the effectiveness and usefulness of the proposed algorithm, four different types of steganalytic feature sets are used for the experiment. The experimental result shows that the number of features that determines the class label of an image is reduced successfully and the accuracy which identifies the class label has improved significantly. Greater accuracy of the classifier with feature selection is achieved as compared to classifier's accuracy on the entire data set.

6. Reference

- Sumathi CP, Santanam T, mamaheswari GU. A Study of Various Steganographic Techniques Used for Information Hiding"; International Journal of Computer Science & Engineering Survey (IJCES). 2013; 4:6.
- Abbas Cheddad, Joan Condell, Kevin Curran, Paul Mc Kevitt: Digital Image Stegnography: Survey and Analysis of Current Methods"; Springer, 7-9, 2002, 310-323.
- Morkel T, Eloff JPH, Olivier MS, An overview of image Stegnography"; Information and Computer Security Architecture (ICSA) Research Group.
- Sumeet Kaur, Savina Bansal, Bansal RK. Steganography and Classification of image steganography techniques"; International Journal of Innovative Research in Computer and Communication Engineering (An ISO 3297: 2007 Certified Organization), 2016; 4:(6).
- Davidson JL, Jalan. Feature Selection for Steganalysis using the Mahalanobis distance"; SPIE Proceedings Media Forensics and Security, 7541, San Jose, CA, USA 2010, 7541.
- Arooj Nissar AH. Mir: Classification of Steganalysis techniques: A study; Digital Signal Processing, December 2010; 20; 1758-1770.
- Hossein Malekmohamadi, Shahrokh Ghaemmaghami. Steganalysis of LSB based Images Stenography using Spatial and Frequency Domain Features"; International Journal of Computer Applications by IJCA Journal, 2014 102.
- Abolghasemi M, Aghainia H, Faez K, Mehrabi MA, LSB data hiding detection based on gray level co-occurrence matrix; International symposium on Telecommunications,. Iran, Tehran. 2008, 656-659.
- Kekre HB, Athawale AA, Patki SA. Steganalysis of LSB Embedded Images using Gray Level Co-occurrence Matrix; International Journal of Image Processing (IJIP), 2011, 5(1).
- Guorang Xuan, Yun Q Shi, Cong Huang, Dongdong Fu, Xiuming Zhu, Peiqi Chai and Jianjiong Gao: Steganalysis using High- Dimensional Features derived from Co-occurrence Matrix and class-wise Non-Principal components Analysis (CNPCA).
- Fridrich J. Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes"; 6th Information Hiding Workshop, Toronto, ON, Canada, 2004.
- Muhammad Imrana, Rathiah Hashima, Noor Elaiza Abd Khalidb. An overview of Particle Swarm Optimization variants; Malaysian Technical Universities Conference on Engineering & Technology, MUCET 2012 Part 4 Information And Communication Technology, 2012.
- Kennedy J, Eberhart R. Particle swarm optimization; IEEE conference.
- Karaboga D, Akay B.A survey: Algorithms simulating bee swarm intelligence; Artificial Intelligence Review, 2009; 31 (1-4):61-85.
- D.Karaboga, B.Basturk: On the performance of artificial bee colony (ABC) algorithm; Applied Soft Computing, Volume, 2008; 8(1):687-697.
- Ghareh F, Mohammadi M. Saniee Abadeh: Image Steganalysis using a Bee Colony feature selection algorithm; Engineering Applications of Artificial Intelligence, 2003, 13,
- Huang, Guang-Bin, Zhu, Qin-Yu; Siew, Chee-Kheong.

- Extreme learning machine; theory and application, 2006.
18. Veenu Basin, Punam Bedi: Steganalysis of Coloured JPEG Images using Ensemble of Extreme Machine Learning; Int. J. on Recent Trends in Engineering and Technology, 2014; 11:(1).
 19. Katzenbeisser, S., Petitcolas, F.: "Information Hiding Techniques for Steganography and Digital Watermarking"; Artech House, Boston (2000).
 20. Fridrich J, Kodovsky J. Calibration revisite. 11th ACM Multimedia & Security Workshop. Princeton, NJ, 2009, 63-74.
 21. Neha Singh, Jyoti Kumari, Charu Aggarwal. Feature Selection for Steganalysis using Glow-worm Algorithm International Journal of Advanced Research and Development, 2017; 2(5):200-206.
 22. Veenu Basin, Punam Bedi, Neha Singh, Charu Aggarwal. FS-EHS: Harmony Search based Feature Selection Algorithm for Steganalysis using ELM; Innovations in Bio-inspired Computing and Applications, Springer International Publishing Switzerland, 2016.