# Digital Steganalysis: Computational Intelligence Approach

Roshidi Din and Azman Samsudin

***Abstract*—**In this paper, we present a consolidated view of digital media steganalysis from the perspective of computational intelligence. In our analysis the digital media steganalysis is divided into three domains which are image steganalysis, audio steganalysis, and video steganalysis. Three major computational intelligence methods have also been identified in the steganalysis domains which are bayesian, neural network, and genetic algorithm. Each of these methods has its own pros and cons.

***Keywords*—**Steganalysis, Computational Intelligence, Image Steganalysis, Audio Steganalysis, Video Steganalysis

## I. Introduction

Over the last decade, one of the most significant current discussions in computer science is the field of information security. In general, information security is the techniques, policies and strategies used to protect and secure computer systems, in maintaining the operations of an organization. One of the concerns in information security is the concept of information hiding. It is the process of embedding information into digital content without causing perceptual degradation. In the new global economy, information hiding can be used to maintain the authenticity and confidentiality of valuable information, to protect the data from possible sabotage, theft, or unauthorized viewing. Before 1990s, the concept of information hiding has received less attention from the research community or the industry. However, this is changing rapidly and even so since the first academic conference on Information Hiding [1] in mid year 1996.

There are two main purposes in information hiding: (1) to protect against the detection of secret messages by a passive adversary, and (2) to hide data so that even an active adversary will not be able to isolate the secret message from the cover data. Information hiding system can be divided into four areas which are Covert Channels, Steganography, Anonymity, and Copyright Marking as shown in Fig. 1. A survey [2] of current information hiding has shown that steganography is one of the recent important subdisciplines. This is because most of the proposed information hiding system is designed based on steganography. Today, steganography is most often associated with the high-tech application where data are hidden with other information in an electronic file.

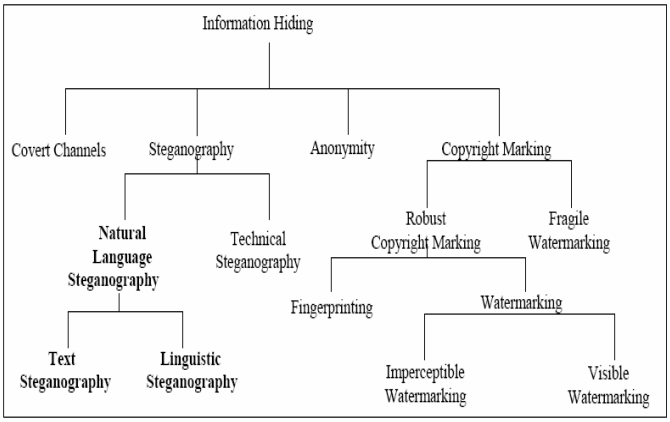


Fig. 1: Information hiding classification

Steganography is the art and science of secret communication, aiming to conceal the existence of a communication which has been used by revolutionaries, spies, the military, and perhaps terrorists.

As a 'hidden writing', steganography uses a covert communication between two parties whose existence is unknown to a possible attacker. If this is done properly, the exchanged messages should not arouse any suspicion since the information passed is an innocent looking message which does not require any secret key as part of its information hiding process. This can be done in many ways such as

Manuscript received December 17, 2008: Revised version received January 30, 2009. This work was supported in part by the School of Computer Sciences, Universiti Sains Malaysia (USM).

Roshidi Din is with the School of Computer Sciences, Universiti Sains Malaysia (USM), 11800, Pulau Pinang, Malaysia, on leave from the Universiti Utara Malaysia (UUM), Malaysia (e-mail: roshidi@uum.edu.my).

Azman Samsudin is with the School of Computer Sciences, Universiti Sains Malaysia (USM), 11800, Pulau Pinang, Malaysia, (phone: +604-653-3635; fax: +604-6532158; e-mail: azman@cs.usm.my, azman@mail.cs.usm.my).

inclusion of line break characters, and multiple spacing that represents a hidden message. Some of the steganography techniques, however, are actually not a new technique. They are some older practices in message hiding such as by using invisible ink, tiny pin punctures on selected characters and pencil mark on typewritten characters.

Specifically, steganography can be divided into two broad categories namely technical steganography and natural language steganography. Technical steganography is a technique of hiding information inside a medium such as image [3-7], audio [8-14], and video [15-20]. Several efforts on stegosystem based on technical steganography for hidden and unhidden messages such as StegaMage [21], Steganoflage [22], StegCure [23], Outguess [24], F5 [25], and PQ [26] have also been developed.

On the other hand, natural language steganography is the art of using natural language to conceal secret message [27]. It focuses on hiding information in text by using steganography [28-34] and linguistic steganography [35-39]. The crucial requirement for steganography is perceptual and algorithmic undetectability. In any case, once the presence of hidden information is revealed or even suspected, the purpose of steganography is defeated, even if the message content is not extracted or deciphered.

Many of the new attacks in steganography are derived by analyzing steganography techniques. This process of analyzing steganographic protocols is carried out in order to detect and extract secret messages. The process is called steganalysis which is generally starts with several suspected information streams but uncertain whether any of the information stream contains hidden messages. The goal of steganography is to avoid suspicion on the existence of hidden messages whereas steganalysis aims to discover the hidden message from useless covert messages in a given text or data.

Hence, steganalysis is the process of detecting steganography by analyzing variances among bit patterns on unusually large file sizes. There are at least five challenges of steganalysis:

- Suspected message may or may not have hidden data embedded into them
- Hidden data may or may not have been encrypted before inserted into the image, signal or text
- Analysis of the suspected message may or may not have noise or data encoded into them
- Suspected message may or may not possible to fully recover, decrypt or extract the hidden data
- Steganalysis is very time consuming process

Generally, steganalysis techniques could be classify into two broad categories namely [40]:

- Specific steganalysis - these types of steganalysis techniques are based on the specific targeted of the steganographic tool which is depend on the embedding technique being analyzed. It would give a very good result when tested only on the right embedding technique, whereas might fail on the other steganographic embedding techniques
- Universal steganalysis - these types of steganalysis are based on designing a classifier, which in turn, is based on statistical moments derived from a variety of embedding techniques. Then, the classifier is used to distinguish between cover-message and stego-message

Now, with the steganalysis becomes more mature, this technology has been rapidly being applied into practice [41]. Based on numerous practices, the pattern of steganalysis system can be divided into four approaches [42] which are:

- Supervised learning detection – use a classifier to identify any pattern
- Blind identification – pattern is detected based on the computed statistics of the digital medium
- Parametrical statistical detection – use a detector to identify pattern
- Hybrid techniques – integrate two or more approaches

One of the significant techniques used in steganalysis system is computational intelligence (CI). Thus, this study believed that CI can be implemented to solve steganalysis problems. Hence, this study suggests that to have a good steganalysis tool, the implementation of steganalysis system should involve some degree of CI.

Thus, the purpose of this paper is to discuss the implementation of CI methods on steganalysis task. Several domain area of steganalysis environment has been formalized in order to justify each domain area against the right CI methods.

The rest of the paper is organized as follows: In Section 2 we introduce an artificial intelligence concept and tools. The discussion of conventional AI, computational intelligence and hybrid system are also included in this section. Section 3 discusses the implementation of computational intelligence on steganalysis environment such as image steganalysis, audio steganalysis, and video steganalysis. In Section 4 we discuss the counterattack tools based on computational intelligence which can be applied on steganalysis system. Finally, we conclude and suggest directions for further research in Section 5.

## II. Artificial Intelligence

Established in the mid 1950s, Artificial Intelligence (AI), in theory, is an artifact built of human's ability to construct and demonstrate an intelligent behavior such as reasoning, learning and perceptual processes by machine learning. However, AI does not have to confine itself to methods that are biologically observable.

From another perspective, artificial intelligence is both an art and a science [43]. Generally speaking, AI systems are

built into two types of automated inference reasoning engines which are forward reasoning and backwards reasoning. Meanwhile, AI applications can be also divided into two types, in terms of consequences:

- *Classifiers*

Classifier is a mechanism or algorithm that takes unclassified variables as an input and gives a prediction of the class as an output. The classifier provides rules, which identify potentially reliable future output. For example,

if $x$ = {"*shiny*" and "*valuable*" and "*in demand*"}
then $y$ = *diamond*

we might insert the variables $x$, such as {"*shiny*", "*valuable*", "*in demand*"} to the classifier and the prediction of output $y$, might be "*diamond*". There are two types of classifier either *trained classifier* or *fixed classifier*. First, it must be trained with classified variables before the classifier is utilized. That is all it takes to be a classifier. Second, if we might want our machine to be fixed classifier so that every time we train the classifier with the same training data and insert the same variables, we can get a same prediction output.

- *Controllers*

Controllers do however classify conditions before inferring actions, and therefore classification forms a central part of most AI systems. The design of a controller can be understood as the search of the best among all specification. The goal of the controller is done to fulfill a set of conflicting specifications. Thus, a control decisions can be thought of as a transformation from *state variables* to *action variables*. *State variables* describe the *current state* of the physical entity and the *desired state*. *Action variables* are those that can be directly altered by the controller. For example, a state condition $x$ of the electrical current and the action condition $y$ of the electrical controller;

if $x$ = {" *temperature is high" or "current is high"*}
then $y$ = *reduce current*

if $x$ = {" *temperature is low" or "current is low"*}
then $y$ = *increase current*

state variables ={ *temperature, current*}
action variable ={*change in current*}

the action $y$ would reduce the electrical current when a temperature and current state are high, whereas the action $y$ would increase the electrical current when a temperature and current state are low.

The ultimate achievement in the field of AI would be to develop a tool that can replicate or exceed human internal capabilities, including reasoning, recognition, understanding, imagination, creativity, and emotions. Due to these ultimate challenges in AI, the development of several useful computing tools has arisen in order to meet the expectation. The tools of particular interest can be roughly divided into conventional AI, computational intelligence, and hybrid intelligent systems as shown in Fig. 2.

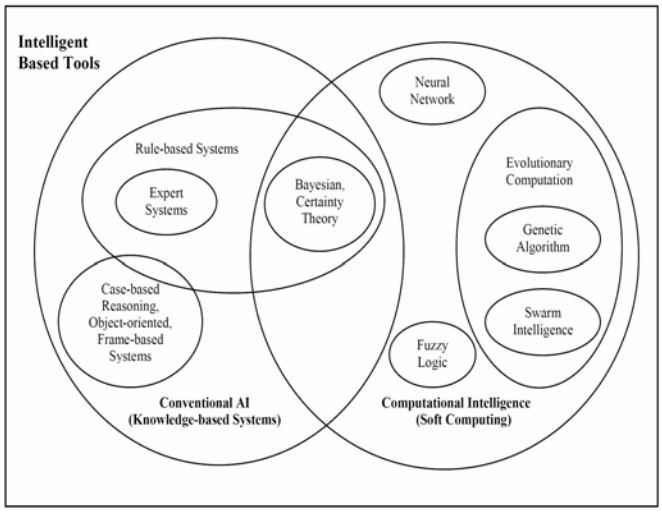


Fig. 2: Categories of computing tools (adapted from [44 - 45])

### A. Conventional AI

Conventional AI which is also known as knowledge-based Systems is being applied in many of the traditional Rule-based AI areas. Researchers are trying to develop AI systems that are capable of performing in a limited sense, "like a human being" [46]. Knowledge-based Systems includes Rule-based Systems (Expert Systems Bayesian, and Certainty Theory), Case-based Reasoning, Object-oriented, and Frame-based Systems.

Rule-based Systems such as on Expert Systems is usually work best in the organizations with high levels of management staff's know-how and difficult to transfer their knowledge to low level staff. The management staffs explain the ways of how to solve problems that are incorporated into the system. A reasoning capability of Rule-based Systems is based on large amounts of known information to provide a decision.

Case-based Reasoning (CBR) systems upon being presented with a problem finds a case in its knowledge base and presents its solutions as an output and attempts to solve new problems based on past solutions of similar problems. Meanwhile, Bayesian networks are heavily based upon probability theory and representing problem domain as a network.

Specifically, conventional AI is based on machine learning, which is the development of the techniques and algorithms that allow machines to "learn" or at least simulate learning. Machine learning attempts to use computer programs to

generate patterns or rules from large data sets. Machine learning makes heavy use of symbolic formalism and logic, as well as statistics. Thus, conventional AI can be also classified as symbolic AI, logical AI, neat AI and Good Old Fashioned Artificial Intelligence (GOFAI).

### *B. Computational Intelligence (CI)*

CI is also known as Soft Computing which refers to a collection of soft computing techniques in computer science, engineering, natural language processing and some business disciplines, contrasting it with classical artificial intelligence. It is a very young discipline compared with other disciplines such as philosophy, neurobiology, evolutionary biology, and psychology that have been studying intelligence much longer.

CI is the study of the system design that acts intelligently in order to understand the principles that make intelligent behavior possible, in natural or artificial systems [47] which involve iterative development or learning. Whereas conventional AI is considered to be a top-down approach, CI is more bottoms-up, where solutions emerge from an unstructured initial state. The techniques are resemble biological processes and intended to complement each other.

A long-term goal for CI is to create cognitive systems that could compete with humans in large number of areas. The output of a CI system often includes predictions and/or decisions.

A good part of CI research is concerned with low-level cognitive functions such as perception, object recognition, signal analysis, discovery of structures in data, simple associations and control [48]. CI includes methods such as neural networks, evolutionary computation (genetic algorithms and swarm intelligence) and other optimization algorithms. Techniques for handling uncertainty, such as bayesian, fuzzy logic, certainty theory fit into both categories. All these techniques use a mixture of rules and associated numerical values. Currently, subjects in computational intelligence as defined by IEEE Computational Intelligence Society include fuzzy systems, neural networks and evolutionary computation (genetic algorithms and swarm intelligence).

### *C. Hybrid System*

Hybrid systems are becoming popular due to their predictions and/or decisions capabilities in analyzing and handling complex problems environment. With hybrid system, attempts are made to combine at least two AI or CI disciplines. There are several ways in which different computational techniques can be complementary as hybrid system which are including dealing with multifaceted problems, capability enhancement, parameter setting and clarification and verification. Mostly, hybrid system is very flexible and a decision is more accurate [49].

In recent years, hybridization of intelligent systems is a promising research field for the next generation of intelligent systems [50]. The integration of different learning techniques and adaptation techniques can overcome the limitations of each technique and to achieve better results. Hybrid intelligent system can be classify into four (4) categories namely stand-alone, transformational, hierarchical hybrid and integrated hybrid systems [51].

Research on hybrid systems is one of the key problems of developing hybrid intelligent systems and it is on integration of computing tools such as knowledge-based systems, certainty theory, bayesian, neural network, fuzzy logic, genetic algorithms, particle swarm optimization etc.

The hybrid intelligent systems has many important practical applications in science, technology, and business such as facial recognition [52], speech recognition [53], optical lens design [54], business intelligence systems [55] etc.

## III. COMPUTATIONAL INTELLIGENCE ON STEGANALYSIS

Until recently, there are lots of steganalysis methods that have been proposed by researchers [56-59]. All of these steganalysis methods can be classified into two types of the steganalysis paradigm which are identified as statistical steganalysis [60-65] and CI steganalysis (see Fig. 3).

Statistical steganalysis consists of Linear Regression such as Simple Pair Analysis and Regular and Singular Analysis, Support Vector Machine, Information Theory Approach and Machine Learning whereas Neural Network, Fuzzy Logic and Genetic Algorithm have been widely classified as CI steganalysis methods. Bayesian and Certainty Theory appear on both of these two steganalysis classification.

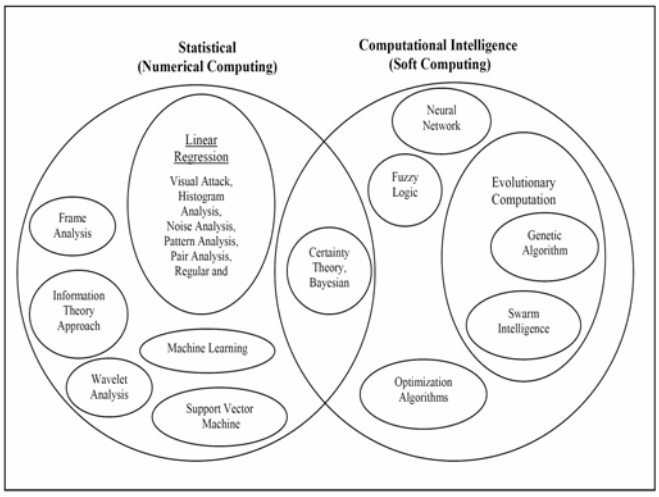


Fig. 3: The paradigm of digital steganalysis methods

Commonly, the implementation of computational intelligence, and their hybrid methods in steganalysis environment are collectively referred to as *intelligent steganalytic systems* (ISS) as shown in Fig. 4. ISS are becoming increasingly distributed in terms of both their applications and implementations when comes to utilizing CI

methods. Generally, ISS consolidate two, three or more CI methods that are either used in series or integrated in a way to produce advantageous results through synergistic interactions [66].

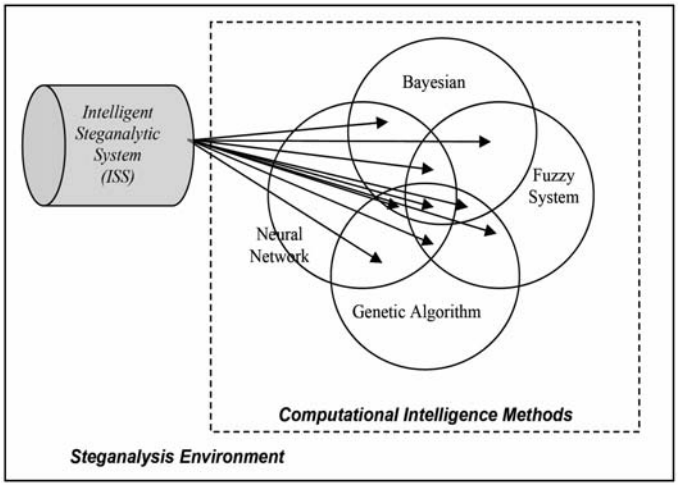


Fig. 4: Synthesis of CI methods on steganalysis environment (adopted from [67])

Many complex domains have various component problems. Each of which may require different types of processing. These limitations have been the central driving force behind the creation of intelligent systems where two or more techniques are being combined in a manner that overcomes the limitations of individual techniques. There are three (3) main reasons for creating ISS which are [68]:

- *Technique enhancement* – integration of different techniques to overcome the limitations of each individual technique
- *Multiplicity of application tasks* – this is created because there is no single technique applicable to solve the different subproblems that a given application may have
- *Realizing multifunctionality* – creating a solution that can exhibit multiple information processing capabilities within one architecture

Nowadays, many researchers have been applying CI on steganalysis environment. Most of their results have proven that the application of CI methods on ISS has given a great influence on steganalysis performance. Thus, we have identified that the digital steganalysis environment can be divided into three (3) domains which are image steganalysis, audio steganalysis, and video steganalysis as shown in Fig. 5. Despite different CI domains that have been proposed, the possibilities of using the techniques for steganalysis are still under-utilized.

### *A. Image Steganalysis*

Currently, several methods for detecting image steganography with CI such as LSB embedding [69], spread spectrum steganography [70], and LSB matching [71-72], have been successfully being implemented in steganalysis [73].

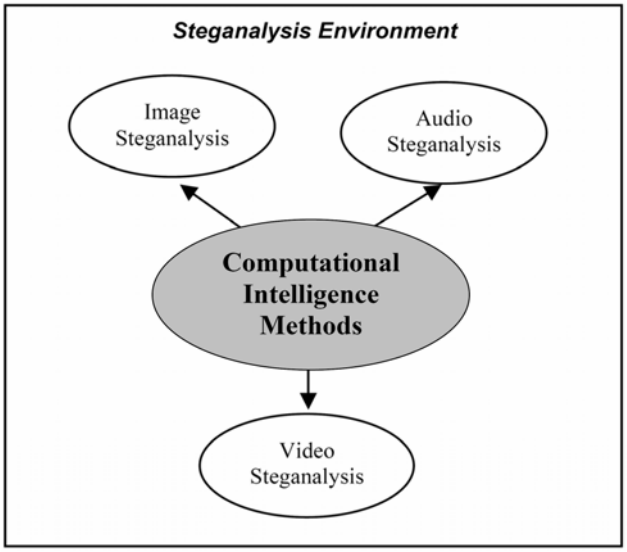


Fig. 5: Computational intelligence on steganalysis domain on digital medium

#### *1) Bayesian*

On analyzing an image, one steganalysis approach [74] had proposed to estimate the hidden message based on a Bayesian framework. Message embedding in bit planes of an image is modeled as a binary symmetric channel. However, this method does not work for LSB embedding due to the lack of statistical structure in the bit plane.

#### *2) Neural Network*

A neural network [75] has been applied to analyze the possible occurrences of certain image pattern through histogram to detect the presence of data. They have used neural network approach to check for those discrepancy patterns and train itself for better accuracy by automating the whole process from decomposition, signature searching, detection and elimination of the detection framework.

In another study, method based on neural network [76] had proposed to gather statistics features of images to identify the underlying hidden data. This study used neural network to analyze object digital image based on three different types of transformation which are Domain Frequency Transform (DFT), Domain Coefficient Transform (DCT) and Domain Wavelet Transform (DWT).

Meanwhile, the work on detection of wavelet domain information hiding techniques [77] has suggested statistical analysis on the texture of an image. Wavelet coefficients in each subband of wavelet transform are modeled as a

Generalized Gaussian Distribution (GGD) with two parameters. It appears that those parameters are a good measure of image features and can be used to discriminate stego-images from innocent images. Neural network is adopted to train these parameters to get the inherent characteristic of innocent and stego-images.

Other study also claimed [78] that an artificial neural network capable of supervised learning results in the creation of a surprisingly reliable predictor of steganographic content, even with relatively small amounts of embedded data. The interesting result is that clean color images can be reliably distinguished from steganographically altered images based on texture alone, regardless of the embedding algorithm.

Another study [79] utilized an artificial neural network as the classifier in a blind steganalysis system. They found that an artificial neural network performs better in steganalysis than Bayes classifier due to its powerful learning capability. Thus, IEEE Computer Society [80] has suggested Artificial Neural Network Technology System (ANNTS) specifically for this purpose. This technology is designed to recognize the digital files containing messages hidden by scanning an image or other file. ANNTS can accurately identify steganographic images between 85% and 100% of the time.

*3) Genetic Algorithm*

Through a Computational Immune System (CIS) [81], a genetic algorithm approach has been used in blind steganography detection. They have developed CIS classifiers, which evolved through a genetic algorithm (GA), that is able to distinguish between clean and stego-images by using statistics gathered from wavelet decomposition. A further study [82] has investigated an Artificial Immune System (AIS) approach to novel steganography detection for digital images. AIS typically mimic portions of the Biological Immune System (BIS) to provide a solution to a computational problem.

Meanwhile, an application of genetic algorithm to optimal feature set selection in supervised learning using Support Vector Machine (SVM) for image steganalysis [83] has also been presented. A genetic algorithm approach is used to optimize the feature set used by the classifier. Experimental results showed that the correct identification rates were as high as 98%, and as low as less than 2%.

*4) Hybrid System*

There are two studies that have been done on hybrid technique of image steganalysis [84-85]. This study has proven that the effectiveness of the AI hybrid in the dynamic environment is as good as Dynamic Evolving Neural Fuzzy Inference System (DENFIS) which was presented by [72] to steganalyze LSB in grayscale image.

*B. Audio Steganalysis*

Currently, interest in audio steganalysis is relatively low, despite obvious practical implications. It is most likely because of the limited implementation of audio application based on weak user demand.

*1) Bayesian*

Echo coding is one of the most effective coding methods in terms of the signal-to-perceived noise ratio in audio steganosystem. In Bayesian method, the process of distinguishing the audios with and without hidden data can be viewed as classification problem. Thus, a study [86] was carried out to detect hidden message by typical echo coding in audio steganalysis on statistical analysis of peak frequency with Bayes as a classifier. Experiments are conducted on a set of various types of audios and the correct rate of classification reaches to 80%. Compared with the method proposed by [87], this method is less time-consuming and produces high detecting accuracy for various embedding parameter combinations.

*2) Neural Network*

One of the audio steganalysis approaches is the use of the principle of Diminishing Marginal Distortions (DMD) [88]. This steganalysis technique is based on the effects of repeated data embedding on the morphological structure of the audio signals. Thus, the principle of DMD is used to detect the presence of hidden messages in uncompressed audio files by using a single layer Feed Forward Neural Network (FFNN) for classification.

Another study utilized a wavelet domain based on Principal Component Analysis (PCA) [89] by using Radial Basis Function (RBF) network as a classifier. This scheme is used to detect the stego-audio signals embedded by wavelet domain LSB, Quantization Index Method (QIM) and Addition Method (AM). Simulation results show that the performances of the detection rates are all greater than 92%. This scheme does not only reduce the dimension of the feature vector effectively and simplifies the design of the classifier, but also keeps the detection performance high.

*3) Genetic Algorithm*

In audio steganalysis, GA is chosen because of its robustness to noise and does not require gradient information to find a global optimal. Spread Spectrum Watermarking (SSW) is one of the most interesting and powerful methods for embedding hidden information into audio signal. It is expected to have high degree of robustness, security and perceptual transparency.

However, a study [70] has shown that the SSW approach has low performance for detecting the exact location of the watermark signal through an attack based on genetic algorithm. Besides the work by [70], the use of genetic algorithm [90] has been explored to aid autonomous

intelligent software agents capable of detecting any hidden information in audio files, automatically. This agent would create the Detection Agent (DA) in architecture comprising of several different agents that collaborate together to detect the hidden information.

Another GA-based steganalysis approach called Stegobreaker [91] was proposed where the generated rules are used to classify audio documents in the real time environment. Experimental results showed that the Stegobreaker method worked effectively for the selected datasets and has the flexibility to be used to meet users' special requirements.

### C. Video Steganalysis

Based on our survey, only one work on video steganalysis called Inter-frame Correlation Steganalysis [92] has been explored. This study proposed a blind steganalysis method to compress video stream by using a three layer Feed Forward Neural Network (FFNN) as the blind classifier. The features of the blind classifier are selected from the global Discrete Cosine Transform (DCT) domain statistics in one single video scene on the collusion basis. Experimental results verify the availability of this scheme.

## IV. CI Based Counterattack Tools

Based on the above explanation, it can be concluded that computational intelligence has a strong chance to be applied successfully on steganalytic system or steganasystem. However, there are also several efforts on CI being proposed as a counterattack tool to the steganasystem, namely;

- *GA based methodology for breaking the steganalytic [93]*: This study applies genetic algorithm by adjusting gray values of the cover-image whereas creating the desired statistic features to generate the stego-images that can break the inspection of steganasystem. Experimental results showed that the algorithm of this study can pass the detection of current steganasystems. Besides, the proposed algorithm can also increase the capacity of the embedded message and enhance the peak signal-to-noise ratio of stego-images.

- *Secure steganographic encoding based on OutGuess [94]*: This study has proposed a genetic algorithm on JPEG images to make a secure steganographic encoding. This steganography encoding is based on the *OutGuess* steganalytic tool. A combination of OutGuess steganalysis approach and Maximum Absolute Difference (MAD) for the image quality is used as fitness function and is considered to give almost the optimum solution in JPEG steganography process. Experimental result showed that this combination method works properly to defeat *OutGuess* method.

In these works, however, the capabilities of CI methods are still under investigation in order to be utilized as strong counterattack mechanism in digital steganography environment.

## V. Conclusion

This paper has presented the paradigm of steganalysis application on digital environments. The key idea is to demonstrate the use of CI in steganalysis. We have also addressed three major methods of CI that had been successfully used on steganalysis; they are bayesian, neural network, and genetic algorithm. We found that neural network is a popular choice to be used in image steganalysis whereas genetic algorithm is the first choice for audio steganalysis. Each of these methods has its pros and cons. Therefore, it depends on the steganalyst to use and choose a suitable CI method based on their analysis purposes.

## Acknowledgment

We thank to the School of Computer Sciences, Universiti Sains Malaysia (USM) for financial support provided us for the realization of this research.

## References

[1] X. Ge, R. Jiao, H. Tian, and J. Wang, “Research on information hiding,” *US-China Education Review*, USA, vol. 5(3:18), May 2006, pp. 77-81.

[2] F.A.P. Petitcolas, R.J. Anderson, M.G., and Kuhn, “Information hiding: A survey,” *Proceedings of the IEEE, Special issue on Protection of Multimedia Content*, vol. 87(7), July 1999, pp. 1062 - 1078.

[3] P. C. Su, and C. C. J. Kuo, “Steganography in JPEG 2000 compressed images,” *IEEE Trans. Consumer Electronics*, vol. 49(4), pp. 824 – 832, 2003.

[4] J. Fridrich, and M. Goljan, “Digital image steganography using stochastic modulation,” *Proceeding of SPIE Electronic Imaging*, Santa Clara, CA, January 2003.

[5] C. C. Thien, and J. C. Lin, “A simple and high-hiding capacity method for hiding digit-by-digit data in images based on modulus function,” *Pattern Recognition*, vol. 36(12), 2003, pp. 2875 - 2881.

[6] P. Y. Chen, and H. J. Lin, “A DWT based approach for image steganography,” *International Journal of Applied Science and Engineering*, vol. 4(3), 2006, pp. 275 – 290.

[7] Roshidi Din, and Hanizan Shaker Hussain, and Sallehuddin Shuib, “Hiding secret messages in images: suitability of different image file types,” *WSEAS TRANSACTIONS on COMPUTERS*, vol. 6(1), January 2006, pp. 127 -132. http://www.worldses.org/journals/computers/computers-anuary2007.doc

[8] Ozer, H., and Sankur, B., “An SVD based audio watermarking technique,” *Proceedings of the IEEE 13th Signal Processing and Communications Applications Conference*, May 2005, pp. 452 - 455.

[9] K. Gopalan, “Audio steganography using bit modification,” *International Proceedings of Conference on Multimedia and Expo (ICME)*, vol.1, 6-9 July 2003, pp. 629 - 32.

[10] K. Gopalan, “Audio steganography by cepstrum modification,” *Proceedings of IEEE International Conference on Acoustics, Speech,*

*and Signal Processing (ICASSP)*, vol. 5, 18 - 23 March 2005, pp. v481-v484.

[11] K. Gopalan, S. Wenndt, A. Noga, D Haddad, and S. Adams, "Covert speech communication via cover speech by tone insertion," *Proc. of IEEE Aerospace Conference*, vol. 4, March 2003, pp. 4_1647 - 4_1653.

[12] X. Li, and H. H. Yu, "Transparent and robust audio data hiding in cepstrum domain," *Proc. IEEE International Conference on Multimedia and Expo, (ICME)*, New York, 2000.

[13] S.K. Lee, and Y.S. Ho, "Digital audio watermarking in the cepstrum domain," *IEEE Trans. Consumer Electronics*, vol. 46, pp. 744 - 750, August 2000.

[14] A. Delforouzi, and M. Pooyan, "Adaptive digital audio steganography based on integer wavelet transform," in *Proceeding of 3rd International Conference on Intelligent Information Hiding and Multimedia Signal Processing* (IIHMSP), vol. 2, November 2007, pp. 283 - 286.

[15] C. Xu, X. Ping, and T. Zhang, "Steganography in compressed video stream," *Proceeding of First International Conference on Innovative Computing, Information, and Control (ICICIC)*, vol. 1, 2006, pp. 269 – 272.

[16] J. J. Chae, and B. S. Manjunath, "Data hiding in video," *Proceedings of the 6th IEEE International Conference on Image Processing*, 1999, pp. 311 – 315.

[17] M. Pazarci, and V. Dipcin, "Data embedding in scrambled digital video," in *Proceedings of the 8th IEEE International Symposium on Computers and Communication*, vol. 1, pp. 498 - 503, 2003.

[18] A. Giannoula, and D. Hatzinakos, "Compressive data hiding for video signals" *Proceedings of International Conference on Image Processing*, 2003, pp. 1529 - 1532.

[19] G. Caccia, and R. Lancini, "Data hiding in MPEG2 bit stream domain" *Proceedings of International Conference on Trends in Communications*, 2001, pp. 363 - 364.

[20] J. Zhang, J. Li, and L. Zhang, "Video watermark technique in motion vector," *Proceedings of XIV Brazilian Symposium on Computer Graphics and Image Processing*, 2001, pp.179 -182.

[21] Roshidi Din, and Hanizan Shaker Hussain, "The capability of image in hiding a secret message," in *Proceedings of the 6th WSEAS International Conference on Signal, Speech and Image Processing (SSIP)*, Portugal, September 2006.

[22] A. Cheddad, J. Condell, K. Curran, and P. McKevitt, "Enhancing steganography in digital images" *Canadian Conference on Computer and Robot Vision*, Windsor, Ont., May 2008, pp. 326 - 332.

[23] L.Y. Por, W. K. Lai, Z. Alireza, T. F. Ang, M.T. Su, and B. Delina, "StegCure: a comprehensive steganographic tool using enhanced LSB scheme", *WSEAS TRANSACTIONS on COMPUTERS*, vol. 7(8), August 2008, pp. 1309 – 1318.

[24] N. Provos, "Defending against statistical steganalysis," 1*0th USENIX Security Symposium*, 2001.

[25] A. Westfeld, "F5: a steganographic algorithm: high capacity despite better steganalysis," *4th International Workshop on Information Hiding*, 2001.

[26] J. Fridrich, M. Goljan, and D. Soukal, "Perturbed quantization steganography with wet paper codes", *ACM Multimedia Workshop*, Magdeburg, Germany, September 20-21, 2004.

[27] M. Chapman, G. I. Davida, and M. Rennhard, "A practical and effective approach to large-scale automated linguistic steganography" in *Proceedings of the Information Security Conference (ISC '01)*, Malaga, Spain, October 2001, pp.156 - 165.

[28] J. T. Brassil, S. Low, N. Maxemchuk, and L. O'Gorman, "Electronic marking and identification techniques to discourage document copying," *Proceedings of IEEE INFOCOM '94*, Toronto, Canada, June 1994, pp. 1278 - 1287.

[29] S. H. Low, N. F. Maxemchuk, J. T. Brassil, and L. O'Gorman, "Document marking and identification using both line and word shifting," *Proceedings of the 14th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '95)*, Washington DC, USA, , vol. 2, April 2 - 6, 1995, pp. 853 – 860.

[30] D. Huang, and H. Yan, "Inter-word distance changes represented by sine waves for watermarking text images," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 11(12), December 2001, pp. 1237 - 1245.

[31] Y. Kim, K. Moon, and I. Oh, "A text watermarking algorithm based on word classification and inter-word space statistics," *Proceedings of the 7th IEEE International Conference on Document Analysis and Recognition (ICDAR'03)*, Washington, USA, 2003, pp. 775 – 779.

[32] J. T. Brassil, S. Low, and N. F. Maxemchuk, "Copyright protection for the electronic distribution of text documents," in *Proceedings of the IEEE USA*, vol. 87(7), July 1999, pp. 1181 - 1196.

[33] N.F. Maxemchuk, and S. Low, "Marking text documents," *IEEE-ICIP'97*, Santa Barbara, California, vol. 3, 1997, pp. 13 - 16.

[34] Bender W, Gruhl D, Morimoto N, Lu A. Techniques for data hiding. *IBM Systems Journal*, 1996, 35(3 & 4): 313-336.

[35] H. Elkamchouchi, and M. Negm, "Hiding english information in extended Arabic characters (HEMERAC)," *Proceedings of the 20th National Radio Science Conference (NRSC 2003)*, March 2003, pp. C12-1-8.

[36] M. J. Atallah, C. J. McDonough, V. Raskin, S. Nirenburg, "Natural language processing for information assurance and security: an overview and implementations," *Proceeding of the Workshop on New Security Paradigms (NSPW'00)*, ACM Press, September 2000, pp. 51 – 65.

[37] M. Niimi, S. Minewaki, H. Noda, and E. Kawaguchi, "A framework of text-based steganography using SD-form semantics model," in *Proceedings of the Pacific Rim Workshop on Digital Steganography*, Kitakyushu, Japan, July 2003.

[38] M. J. Atallah, V. Raskin, M. Crogran, C. Hempelmann, F. Kerschbaum, D. Mohamed, and S. Naik, "Natural language watermarking: design, analysis and a proof-of-concept implementation," CERIAS Technical Report 2001-13, Purdue University, West Lafayette, 2001,

[39] H. Nakagawa, K. Sampei, T. Matsumoto, S. Kawaguchi, K. Makino, and I. Murase, "Text information hiding with preserved meaning – a case for Japanese documents," *Information Processing Society of Japan (IPSJ) Transaction*, vol. 42(9), 2001, pp. 2339 – 2350.

[40] M. Kharrazi, "Performance study of common image steganography and steganalysis techniques," Special Section on Security, Steganography, and Watermarking of Multimedia Contents, *Journal of Electronic Imaging*, vol. 15(4), October - December 2006, 041104.

[41] G. Luo, X. Sun, L. Xiang, and J. Huang, "An evaluation scheme for steganalysis-proof ability of steganalysis algorithms", *International Conference on Intelligent Information Hiding and Multimedia Signal Processing* (IIHMSP), vol. 2, 26-28 Nov 2007, pp. 126 - 129.

[42] R. Chandramouli, and S. K. Subbalakshmi, "Current trends in steganalysis: a critical survey," *Control, Automation, Robotics and Vision Conference (ICARCV)*, vol. 2, December 2004, pp 964–967.

[43] S. L. Tanimoto, "The elements of artificial intelligence: an introduction using LISP," Computer Science Press Inc., 1803 Research Boulevard, Rockville, Maryland 20850, 1987.

[44] A. A. Hopgood, "Intelligent systems for engineers and scientists," 2nd Edition, CRC Press, 2001.

[45] S. Kumar, "Neural networks - a class room approach," McGraw-Hill, 2004.

[46] N. M. Martin, and L. C. Jain, "Introduction to Neural Network, Fuzzy systems, Genetic Algorithms, and their Fusion in Fusion of Neural Networks, Fuzzy Sets, and Genetic Algorithms : Industrial Applications," *International Series on Computational Intelligence*, CRC Press, 1999, pp. 1–12.

[47] D. Poole, A. Mackworth, and R. Goebel, "Computational intelligence: a logical approach, Oxford University Press, New York, 1998.

[48] W. Duch, and J. Mandziuk, "Challenges for computational intelligence", *Studies in Computational Intelligence*, vol. 63, Springer, 2007, pp. 1-13.

[49] K.J. Cios, L. S. Goodenday, and L. M. Sztandera, "Hybrid intelligence system for diagnosing coronary stenosis: combining fuzzy generalized operators with decision rules generated by machine learning algorithms," *Engineering in Medicine and Biology Magazine*, vol. 13(5), Nov/Dec 1994, pp. 723 - 729.

[50] A. Abraham, "Hybrid intelligent systems: evolving intelligence in hierarchical layers," in *Studies in Fuzziness and Soft Computing*, vol. 173, Springer-Verlag, Berlin, Germany, 2005, , pp. 159-169.

[51] A. Abraham, "Intelligent systems: architectures and perspectives, recent advances in intelligent paradigms and applications, A. Abraham, L. Jain, and J. Kacprzyk (Eds.) in *Studies in Fuzziness and Soft Computing*, Springer Verlag Germany, 2002, pp. 1 - 35.

[52] A. Raouzaiou, N. Tsapatsoulis, V. Tzouvaras, G. Stamou, and S. Kollias, "A hybrid intelligence system for facial expression Recognition," *in Proceedings of EUNITE 2002, European Symposium on Intelligent Technologies: Hybrid Systems and their implementation on Smart Adaptive Systems*, Algarve, Portugal, 2002.

[53] N. Kasabov, and R. Kozma, "Introduction: hybrid intelligent adaptive systems, *INTERNATIONAL JOURNAL OF INTELLIGENT SYSTEMS*, vol. 13(6), 1998, John Wiley & Sons, Inc., pp. 453 – 586.

[54] S.M. Tam, C.K. Kwong, and W.H. Ip, "A hybrid artificial intelligence system for optical lens design," *International Journal of Computer Applications in Technology*, vol. 13(3-5):229 - 236, 2000.

[55] L. Cao, D. Luo, C. Luo, and C. Zhang, "Systematic engineering in designing architecture of telecommunications business intelligence system," Design and Application of Hybrid Intelligent Systems (HIS03), *Third International Conference on Hybrid Intelligent Systems*, Melbourne, Australia, December 14-17, 2003, pp. 1084 – 1093.

[56] I. Avcibas, M. Kharrazi, N. Memon, and B. Sankur, "Image steganalysis with binary similarity measures," *EURASIP J. Appl. Signal Process.* (17), 2005, pp. 2749 – 2757.

[57] S. Lyu and H. Farid, "Detecting hidden messages using higher-order statistics and support vector machines," in *Proc. 5th Int. Workshop on Information Hiding*, 2002.

[58] S. Lyu and H. Farid, "Steganalysis using color wavelet statistics and one-class support vector machines", *Proc. SPIE* 53062004, pp. 35 – 45,.

[59] J. Fridrich, "Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes," *Proceeding of 6th Information Hiding Workshop*, Toronto, 2004.

[60] M. Jiang, E. Wong, N. Memon, and X. Wu, "A simple technique for estimating message lengths for additive noise steganography," *8th International Conference on Control, Automation, Robotics and Vision Conference (ICARCV 2004)*, Kunming, China, vol. 26-9 Dec. 2004, pp. 983 - 986.

[61] S. Liu, and H. Yao, and W. Gao, "Steganalysis of data hiding techniques in wavelet domain," *IEEE Proceedings of International Conference on Information Technology: Coding and Computing, (ITCC 2004)*, vol.1, 2004, pp. 751 - 754.

[62] K. B. Raja, N. Shankara, K. R. Venugopal, and L. M. Patnaik, "Steganalysis of LSB embedded images using variable threshold color pair analysis," *Fourth International Conference on Intelligent Sensing and Information Processing (ICISIP 2006)*, 2006, pp. 11 - 16.

[63] A. Savoldi, and P. Gubian, "Blind multi-class steganalysis system Using wavelet statistics," *Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIHMSP 2007)*, vol. 2, 26-28 Nov. 2007, pp.93 – 96.

[64] L. Bin, L. Fenlin, and W. Ping, "Inter-frame correlation based compressed video steganalysis," *Congress on Image and Signal Processing (CISP08)*, vol. 3, 27-30 May 2008, pp. 42 – 46.

[65] W. Zeng, H. Ai, and R. Hu, "An algorithm of echo steganalysis based on power cepstrum and pattern classification," *International Conference on Audio, Language and Image Processing (ICALIP'08)*, Shanghai, July 2008, pp. 1344 - 348

[66] L.H. Tsoukalas, *Fuzzy and neural approaches in engineering*, John Wiley and Sons Ltd, Canada, 1997.

[67] T. Fukuda, and K. Shimojima, "Hierarchical intelligent robotic system - adaptation, learning and evolution," *Proceeding of International Conference on Computational Intelligence and Multimedia Applications (ICCIMA'97)*, Gold Coast, Australia, 1997, pp.1-5.

[68] S. Goonatilake, and S. Khebbal, *Intelligent hybrid systems*, John Wiley & Sons Ltd, England, UK, 1995.

[69] R. Benton, and H. Chu, "LSB embedding steganalysis using neural network," *3rd International Conference on Information Technology: Research and Education (ITRE 2005)*, Hsinchu, Taiwan, June 2005, pp. 105 - 109.

[70] S. Sedghi, H. R. Mashhadi, and M. Khademi, "Detecting hidden information from a spread spectrum watermarked signal by genetic algorithm," in *Congress on Evolutionary Computation (CEC '06)*, Vancouver, Canada, July 2006, pp. 173 – 178.

[71] R. Ji, H. Yao, S. Liu, and L. Wang, "Genetic algorithm based optimal block mapping method for LSB substitution," *International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP '06)*, December 2006, , pp. 215 - 218.

[72] Q. Liu, and A. H. Sung, "Feature mining and neuro-fuzzy inference system for steganalysis of LSB matching steganography in grayscale images," *Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI 2007)*, Hyderabad, India, January 2007, pp. 2808 - 2813.

[73] Q. Liu, and A. H. Sung, "Detect information-hiding type and length in JPEG images by using neuro-fuzzy inference systems", *Congress on Image and Signal Processing (CISP)*, Sanya, China, vol. 5, May 2008, , pp. 692 - 696.

[74] A. Ambalavanan, and R. Chandramouli, "A bayesian image steganalysis approach to estimate the embedded secret message," *International Multimedia Conference, Proceedings of the 7th Workshop on Multimedia and Security*, ACM Press, New York, USA, 2005, pp. 33 – 38.

[75] U.M. Sekarji, "Detection of hidden information in images using neural networks," SASTRA Tanjore, India, unpublished.

[76] S. Liu, H. Yao, and W. Gao, "Neural network based steganalysis in still images," *Proceedings of the 2003 International Conference on Multimedia and Expo (ICME 2003)*, vol.2, July 2003, pp. II – 509–512.

[77] S. Liu, H. Yao, and W. Gao, "Steganalysis based on wavelet texture analysis and neural network," *Fudan Journal (Natural Science)*, 43 vol. 5 (2004/10), pp. 910 - 913, 2004.

[78] P. Lafferty, and F. Ahmed, "Texture-based steganalysis: results for color images," *Proceedings of the SPIE Mathematics of Data/Image Coding, Compression, and Encryption VII, with Applications*, vol. 5561, 2004, pp. 145 - 151.

[79] Y.Q. Shi, G. Xuan, D. Zou, J. Gao, C. Yang, Z. Zhang, P. Chai, W. Chen, and C. Chen, "Steganalysis based on moments of characteristic functions using wavelet decomposition, prediction-error image, and neural network," *Conference on Multimedia and Expo (IEEE ICME 2005)*, Amsterdam, The Netherlands, July 2005.

[80] L. D. Paulson, "New system fights steganography purpose artificial neural network technology for steganalysis (ANNTS), *News Briefs*, Computer, IEEE Computer Society, vol. 39(8), August, 2006, pp. 25 - 27.

[81] J.T. Jackson, G.H. Gunsch, R.L. Claypoole, and G.B. Lamont, "Blind steganography detection using a computational immune system: a work in progress," *International Journal of Digital Evidence*, vol. 4(1), 2002.

[82] J.T. Jackson, G.H. Gunsch, R.L. Claypoole, and G.B. Lamont, "Novel steganography detection using an artificial immune system approach," *Congress on Evolutionary Computation (CEC '03)*, vol. 1, December 2003, pp. 139 - 45.

[83] T. Knapik, E. Lo, and J. A. Marsh, "Application of genetic algorithm to steganalysis," *Proceedings of the SPIE Modeling and Simulation for Military Applications*, vol. 6228, May 2006, pp. 62280X.

[84] T. Iba, and Y. Takefuji, "Adaptation of neural agent in dynamic environment: hybrid system of genetic algorithm and neural network," *Proceedings of the Second International Conference on Knowledge-Based Intelligent Electronic Systems (KES '98)*, Adelaide, SA, Australia, vol. 3, April 1998, pp. 575 - 584.

[85] I.S. Oh, J.S. Lee, and B.R. Moon, "Hybrid Genetic Algorithms for Feature Selection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26(11), November 2004, pp. 1424 - 1437.

[86] W. Zeng, H. Ai, R. Hu, and S. Gao, "An algorithm of echo steganalysis based on bayes classifier," *International Conference on Information and Automation (ICIA 2008)*, Changsha, China, June 2008, pp. 1667-1670.

[87] H. Ozer, I. Avcibas, B. Sankur, and N. Memon, "Steganalysis of audio based on audio quality metrics," *SPIE Electronic Imaging Conference on Security and Watermarking of Multimedia Contents*, Santa Clara, vol. V, January 2003, pp. 55 – 66.

[88] O. Altun, G. Sharma, M. Celik, M. Sterling, E. Titlebaum, and M. Bocko, "Morphological steganalysis of audio signals and the principle of diminishing marginal distortions," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '05)*, Philadelphia, PA, USA, vol. 2, March 2005, pp. 21 - 24.

[89] J. Fu, Y. Qi, and J. Yuan, "Wavelet domain audio steganalysis based on statistical moments and PCA", *Proceedings of the International*

*Conference on Wavelet Analysis and Pattern Recognition (ICWAPR '07)*, Beijing, China, vol. 4, November 2007, pp. 1619 - 1623.

[90] S. Geetha, S.S. Sivatha Sindhu, and A. Kannan, "An active rule based approach to audio steganalysis with a genetic algorithm," in *Proceedings of the IEEE 1st International Conference on Digital Information Management*, Bangalore, India, December 2006, pp. 131 - 136.

[91] S. Geetha, S.S. Sivatha Sindhu, and A. Kannan, "StegoBreaker: audio steganalysis using ensemble autonomous multi-agent and genetic algorithm," *Annual India Conference*, New Delhi, September 2006, pp. 1- 6.

[92] B. Liu, F. Liu, and P. Wang, "Inter-frame correlation Based compressed video steganalysis," *Congress on Image and Signal Processing (CISP '08)*, Sanya, China, vol. 3, May 2008, , pp. 42 - 46.

[93] Y. Wu, and F.Y. Shih, "Genetic algorithm based methodology for breaking the steganalytic systems," *IEEE Transactions on Systems, Man and Cybernetics*, Computer Vision Lab., New Jersey Institute of Technol., Newark, NJ, USA, vol. 36(1), February 2006, pp. 24 – 31.

[94] A. M. Fard, M. R. Akbarzadeh, and F. Varasteh, "A new genetic algorithm approach for secure JPEG steganography," in *Proceedings of the IEEE International Conference on Engineering of Intelligent Systems*, Islamabad, April 2006, pp. 216 – 219.

**Roshidi Din** is a researcher at the school of Computer Sciences, Universiti Sains Malaysia (USM). He is currently on leave from the Universiti Utara Malaysia (UUM), Malaysia. He received his M.Sc.IT and B.I.T degrees from Universiti Utara Malaysia in 1996 and 1999, respectively. His current research interests lie in information security, steganalysis and natural language steganology.

**Azman Samsudin** is an Associate Professor at the School of Computer Sciences, Universiti Sains Malaysia (USM). He received his Ph.D. and M.Sc. degrees from University of Denver in 1998 and 1992, respectively, and his B.Sc. from University of Rochester in 1989. Since 2005, he has been the Deputy Dean of Postgraduate Studies at the School of Computer Sciences, USM. His research interests lie in the areas of computer systems especially cryptography, network security, interconnection switching network, parallel computing and distributed computing. He has published over 80 academic papers.