International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 1 Issue 6, August - 2012

Improved Statistical Steganalysis Using Syndrome-Trellis Codes

SRAVANTHI.B[1] **RAMESH .J**[2] **NARESH.A**[3]

[1]**M.Tech 2nd year , QIS College of Engineering and Technology, Ongole**

[2]**Asst.Professor, QIS College of Engineering and Technology,Ongole**

[3]**Asst.Professor, Vignan's Nirula institute of Technology and science for women, Guntur**

ABSTRACT:*The word stego is a Greek word known as "Secret". The process of embedding a secret message into an image is called stenography. Basically, distortion occurs when length of pixels increased, if distortion increases hackers can easily attack the system and can easily view the information that is embedded in the image. Hence to minimize the distortion, in this paper we are using the non-binary embedding system through syndrome-trellis codes. A cover "X" with a message "M" and key "K" is embedded into a Stego element "Y" and is when passed through a stego channel image, security is increased. Here "X" and "Y" are random variables on image function. Here the probability distribution on X and Y are same hence no statistical test can detect the steganography. Stego element "Y" is produced by slightly modifying the element "X". In Syndrome trellis two rules are applied, the rule1is Embed entropy bits into MSB's with low costs. 2. Embed entropy bits into LSB's with costs. In this paper every pixels i.e. {0, 1, 2, 3} is embedded into MSB and LSB, hence MSB contains {0, 1} and LSB contains {2, 3} pixel values.*

Keywords:*Watermarking, Steganography, Non-binary embedding system, Stego-image, Stego element, Syndrome trellis code.*

1. Introduction

Conventional cryptographic systems permit only valid key holders access to encrypted data, but once such data is decrypted there is no way to track its reproduction or retransmission. Therefore, conventional cryptography provides little protection against data piracy, in which a publisher is confronted with unauthorized reproduction of information. A digital watermark is intended to complement cryptographic processes. It is a visible, or preferably invisible, identification code that is permanently embedded in the data and remains present within the data after any decryption process.

Watermarking is a technology for embedding various types of information in digital content. In general, information for protecting copyrights and proving the validity of data is embedded as a watermark. Watermarked content can prove its origin, thereby protecting the data.Here a cover "X" with a message "M" and key "K" is embedded into a Stego element "Y" and is when passed through a stego channel, security. Here "X" and "Y" are random variables on image functionwhich is exactly given in fig 1.

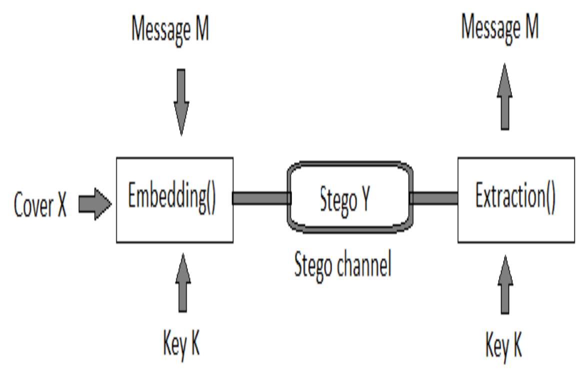

Fig1: Embedding and extraction of a message M into Cover X with key K

1.1 Steganography

ISSN: 2278-0181
Vol. 1 Issue 6, August - 2012

Steganography, coming from the Greek words "stegos", meaning roof or covered and"graphia" which means writing, is the art and science of hiding the fact that communicationis taking place. Using steganography, we can embed a secret message inside apiece of unsuspicious information and send it without anyone knowing of the existenceof the secret message.Steganography and cryptography are closely related. Cryptography scrambles messages so they cannot be understood. Steganography on the other hand, will hide the message so there is no knowledge of the existence of the message in the first place. In some situations, sending an encrypted message will arouse suspicion while an "invisible" message will not do so. Both sciences can be combined to produce better protection of the message. In this case, when the steganography fails and the message can be detected, it is still of no use as it is encrypted using cryptography techniques.With steganography you can send messages without anyone having knowledge of the existence of the communication. There are many countries where it is not possible to speak as freely as it is in some more democratic countries. Steganography can be absolution which makes it possible to send news and information without being censored and without the fear of the messages being intercepted and traced back to you.

Hiding information inside images is a popular technique nowadays. An image with a secret message inside can easily be spread over the World Wide Web or in newsgroups. To hide a message inside an image without changing its visible properties, the cover source can be altered in "noisy" areas with many color variations, so less attention will be drawn to the modifications. *Least-Significant Bit (LSB), Masking, Filtering and Transformations* on the cover image are the most commonly used methods to make these alterations. These techniques can be used with varying degrees of success on different types of image files.

2.Least-significant bit modifications:

The most widely used technique to hide data, is the usage of the LSB. Although there are several disadvantages to this approach, the relative easiness to implement it, makes it a popular method. To hide a secret message inside an image, a proper cover image is needed. Because this method uses bits of each pixel in the image, it is necessary to use a lossless compression format, otherwise the hidden information will get lost in the transformations of a loss compression algorithm. When using a 24 bit color image, a bit of each of the red, green and blue color componentscan be used, so a total of 3 bits can be stored in each pixel. Thus, an 800 × 600pixel image can contain a total amount of 1.440.000 bits (180.000 bytes) of secret data.For example, the following grid can be considered as 3 pixels of a 24 bit color image,using 9 bytes of memory:

(00100111 11101001 11001000)
(00100111 11001000 11101001)
(11001000 00100111 11101001)

When the character A, which binary value equals 10000001, is inserted, the followinggrid results:

(00100111 11101000 11001000)
(00100110 11001000 11101000)
(11001000 00100111 11101001)

In this case, only three bits needed to be changed to insert the character successfully.On average, only half of the bits in an image will need to be modified to hide a secretmessage using the maximal cover size. The resulting changes that are

ISSN: 2278-0181
Vol. 1 Issue 6, August - 2012

made to the least significantbits are too small to be recognized by the human eye, so the message iseffectively hidden.While using a 24 bit image gives a relatively large amount of space to hide messages, it is also possible to use an 8 bit image as a cover source. Because of the smaller space and different properties, 8 bit images require a more careful approach. Where 24 bit images use three bytes to represent a pixel, an 8 bit image uses only one. Changing theLSB of that byte will result in a visible change of color, as another color in the available palette will be displayed. Therefore, the cover image needs to be selected more carefullyand preferably be in gray scale, as the human eye will not detect the difference betweendifferent gray values as easy as with different colors.Disadvantages of using LSB alteration are mainly in the fact that it requires afairly large cover image to create a usable amount of hiding space. Even nowadays,uncompressed images of 800 x 600 pixels are not often used on the Internet, so usingthese might raise suspicion. Another disadvantage will arise when compressing an imageconcealing a secret using a lossy compression algorithm. The hidden message will not survive this operation and is lost after the transformation.

In special domain, the hiding process such as least significant bit (LSB) replacement is done in special domain, while transform domain methods .Hide data in another domain such as wavelet domain. Least significant bit (LSB) is the simplest form of Steganography. LSB is based on inserting data in the least significant bit of pixels, which lead to a slight change on the cover image that is not noticeable to human eye. Since this method can be easily cracked, it is more vulnerable to attacks. LSB method has intense effects on the statistical information of image like histogram. Attackers could be aware of a hidden communication by just checking the Histogram of an image. A good solution to eliminate this defect was LSB matching. LSB-Matching was a great step forward in Steganography methods and many others get ideas from it.

Now, it is planned to introduce a method that embed 2 bits information in a pixel and alter one bit from one bit plane but the message does not necessarily place in the least significant bit of pixel and second less significant bit plane and fourth less significant bit plane can also host the massage.Since in our method for embedding two bits message we alter just one bit plane, fewer pixels would be manipulated during embedding message in an image and it is expected for the steganalysis algorithm to have more difficulty detecting the covert communication. It is clear that in return complexity of the system would increase.
In our method there are only three ways that a pixel is allowed to be changed:

1) Its least significant Bit would alter (So the gray level of the pixel would Increase or decrease by one level)

2) The second less significant bit plane would alter (So the gray level of the pixel would increase or decrease by two levels)

3) The fourth less significant bit plane would alter (So the gray level of the pixel would increase or decrease by eight levels)

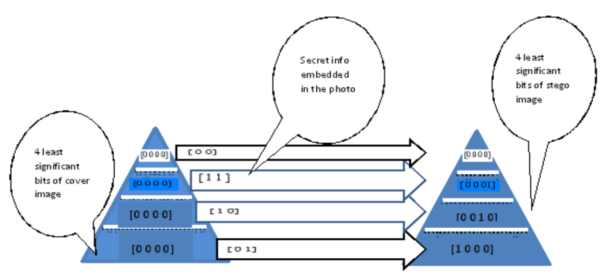

Fig 2: How cover pixel with four less significant bits of [0000] change according to different message.

3. Problem Formulation

The distortion function D is additive over individual cover pixels $D(x,y) = \sum_{i=1}^{n} \rho i(X, yi)$

Where ρi:X × Ii! [−K, K], 0 < K <1,are bounded functions expressing the cost of replacingthe cover pixel *xi* with *yi*. Note that Pi may arbitrarily depend on the entire cover image **x**, allowingthus the sender to place the embedding changes adaptively w.r.t. the image content. The fact thatthe value of ρi(**x**, *yi*) is independent of changes made at other pixels implies that the embeddingchanges do not interact. The boundedness of *D*(**x**, **y**) is not limiting the sender in practice since thecase when a particular value *yi is* forbidden can be resolved by excluding *yi* from I*i*. In practice, the sets I*i*, *i* ϵ {1, *n*}, maydepend on cover pixels and thus may not be available to the receiver. To handle this case, we expandthe domain of ρi to X × I and define ρi(**x**, *yi*) = 1 whenever *yi* ϵ I*i*.We intentionally keep the definition of the distortion function rather general. We assume the sender obtains her payload in the form of a pseudo-random bit stream, such as by compressing or encrypting the original message. We further assume that the embedding algorithm associates every cover image **x** with a pair {Y,∏}, where Y is the set of all stego images into which **x** can be modified and∏is their probability distribution characterizing the sender's actions, ∏(**y**) , *P*(**Y** = **y**|**x**). We think of **x** as a constant parameter that is *fixed in the very beginning* and thus we do not further denote the dependency on it explicitly. For this reason, we simply write *D*(**y**),*D*(**x**, **y**).If the receiver knew **x**, the sender could send up to *H*(∏) bits on average while introducing the average distortion *E*∏[*D*] by choosing the stego image according to ∏. Here**x** does not give any fundamental advantage to the receiver and the same performance can be achieved as long as **x** is known to the sender.

4. Syndrome Coding

Let us first assume a binary version of both embedding problems. Let P:I*i!* {0, 1} be parity function shared between the sender and the receiver satisfying P(*xi*) 6= P(*yi*) such as P(*x*) = *x* mod 2.The sender and the receiver need to implement the embedding and extraction mappings defined asEmb : X × {0, 1}*m* ! Y and Ext : Y! {0, 1}*m* satisfyingExt(Emb(**x**,**m**)) = **m** 8**x** 2 X, 8**m** 2 {0, 1}*m*,respectively. In particular, we do not assume the knowledge of the distortion function *D* at thereceiver and thus the embedding scheme can be seen as being universal in this sense. A commoninformation-theoretic strategy for solving the PLS problem is known as binning, which weimplement using cossets of a linear code. Such a construction, better known as syndrome coding, iscapacity achieving for the PLS problem if random linear codes are used.In syndrome coding, the embedding and extraction mappings are realized using a binary linearcode C of length *n* and dimension *n* − *m*:

$$\mathrm{Emb}(\mathbf{x},\mathbf{m}) = \arg\min_{\mathcal{P}(\mathbf{y})\in\mathcal{C}(\mathbf{m})} D(\mathbf{x},\mathbf{y}),$$

$$\mathrm{Ext}(\mathbf{y}) = \mathbb{H}\mathcal{P}(\mathbf{y})^T,$$

Where P(**y**) = (P (*y*1). . .P (*yn*)), H 2 {0, 1}*m*×*n* is a parity-check matrix of the code C, C(**m**) ={**z** 2 {0, 1}*n*|H**z***T* = **m**} is the cosset corresponding to syndrome **m**, and all operations are in binaryarithmetic.

4. Syndrome-Trellis Codes

We focus on solving the binary PLS problem with previous distortion function andpropose a large class of linear codes which we call the syndrome-trellis codes. The construction behind STCs is not new from an information-theoretic perspective, since theSTCs are trellis codes represented in a dual domain. However, STCs are very interesting forpractical steganography since they allow solving both embedding problems with a very small codingloss over a wide range of distortion profiles even with wet pixels. The same code can be used withall profiles making the embedding algorithm practically universal. STCs offer general and state-of the-art solution for both embedding problems in steganography. Here, we give the description of the codes along with their graphical representation, the syndrome trellis. Such construction is preparedfor the Viterbi algorithm, which is optimal for solving prior art. Important practical guidelines foroptimizing the codes and using them for the wet paper channel are also covered. Finally, we studythe performance of these codes by extensive numerical simulations using different distortion profilesincluding the wet paper channel. Syndrome-trellis codes targeted to applications in steganography were described in, whichwas written for practitioners. In this chapter, we expect the reader to have a working knowledge ofconvolutional codes which are often used in data-hiding applications such as digital watermarking.

Our main goal is to develop efficient syndrome-coding schemes for an *arbitrary* relative payload with the main focus on small relative payloads (think of$\alpha \leq 1/2$ for example). In steganography, the relative payload must decrease with increasing size of the cover object in order to maintain the samelevel of security, which is a consequence of the square root law. Moreover, recent results fromsteganalysis in both spatial and DCT domains suggest that the secure payload for digitalimage steganography is always far below 1/2. Another reason for targeting smaller payloads is thefact that as _ ! 1, all binary embedding algorithms tend to introduce changes with probability 1/2,no matter how optimal they are. Denoting with $R = (n - m)/n$ the rate of the linear code C, thenα->0 translates to R = 1 – α->1, which is characteristic for applications of syndrome coding insteganography.

5.1 Transforming Convolutional Codes to Syndrome-Trellis Codes

Convolutionalcodes were probably the first “practical” codes used for this problem. This is becausethe gap between the bound on the expected per-pixel distortion and the distortion obtained usingthe optimal encoding algorithm (the Viterbi algorithm) decreases exponentially with the constraintlength of the code. The complexity of the Viterbi algorithm is linear in the block length ofthe code, but exponential in its constraint length (the number of trellis states grows exponentiallyin the constraint length).This makes convolutionalcodes (of small constraint length) suitable for our application because the entire cover object

canbe used and the speed can be traded for performance by adjusting the constraint length. Note thatthe receiver does not need to know D since only the Viterbi algorithm requires this knowledge. Byincreasing the constraint length, we can achieve the average per-pixel distortion that is arbitrarilyclose to the bounds and thus make the coding loss approach zero. Convolutional codes areoften represented with shift-registers that generate the code word from a setof information bits. In channel coding, codes of rates $R = 1/k$ for $k = 2, 3 \ldots$ are usually consideredfor their simple implementation. The main drawback of convolutional codes, when implemented using shift-registers, comes fromour requirement of small relative payloads (code rates close to one). A convolutional code of rate$R = (k - 1)/k$ requires $k - 1$ shift registers in order to implement a scheme for _ = 1/*k*. Here,unfortunately, the complexity of the Viterbi algorithm in this construction grows exponentially with *k*. Instead of using puncturing which is often used to construct

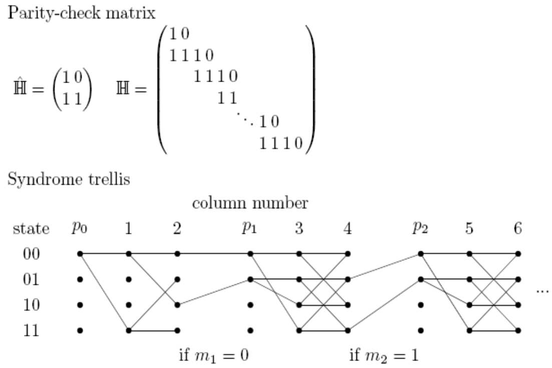

Figure 3: *Example of a parity-check matrix H formed from the sub matrix ˆH (h = 2,w= 2)and its corresponding syndrome trellis. The last h − 1 sub matrices in H are cropped to achieve thedesired relative payload α. The syndrome trellis consists of repeating blocks of w+ 1 column, where"p0" and "pi", i >0, denote the starting and pruning columns, respectively. The column labeledl 2 {1, 2, . . .} corresponds to the* l^{th} *column in the parity-check matrix H.*

high-rateconvolutional codes, we prefer to represent the convolutional code in the dual domain using its paritycheckmatrix. This approach is more efficient as α->0.In the dual domain, a code of length *n* is represented by a parity-check matrix instead of agenerator matrix as is more common for convolutional codes. Working directly in the dual domainallows the Viterbi algorithm to exactly implement the cosset quantizer required for the embedding function (6.2.1). The message can be extracted in a straightforward manner by the recipient using the shared parity-check matrix.

5. Implementation Details

The construction of STCs is not constrained to having to repeat the same sub matrix ˆH along thediagonal. Any parity-check matrix H containing at most *h* nonzero entries along the main diagonal will have an efficient representation by its syndrome trellis and the Viterbi algorithm will have the same complexity O (2*hn*). In practice, the trellis is built on the fly because only the structure of thesub matrix ˆH is needed. As can be seen from the last two columns of the trellis in Figure 3, the connectivity between trellis columns is highly regular which can be used to speed up the implementation by "vectorizing" the calculations. In the forward part of the algorithm, we need to store one bit (the label of the incoming edge) to be able to reconstruct the path in the backward run. This space complexity is linear and should notcause any difficulty, since for *h* = 10, *n* = 106, the total of 210 · 106/8 bytes (122MB) of space isrequired. If less space is available, we can always run the algorithm on smaller blocks, say *n* = 104,without any noticeable performance drop. If we are only interested in the total distortion *D*(**y**) andnot the stego object itself, this information does not need to be stored at all and only the forwardrun of the Viterbi algorithm is required.

6. Design of Good Syndrome-Trellis Codes

A natural question regarding practical applications of syndrome-trellis codes is how to optimize the structure of ˆH for fixed parameters *h* and *w* and a given profile. If ˆH depended on the distortion profile, the profile would have to be somehow communicated to the receiver. Fortunately, this is notthe case and a sub matrix ˆH optimized for one profile seems to be good for other profiles as well. Inthis section, we study these issues experimentally and describe a practical algorithm for obtaininggood sub matrices.Let us suppose that we wish to design a sub matrix ˆH of size *h* × *w* for a given constraint height*h* and relative payload α= 1/*w*. Authors describe several methods for calculating theexpected distortion of a given convolutional code when used in the source-coding problem withHamming measure (uniform distortion profile). Unfortunately, the computational complexity of these algorithms does not permit us to use them for the code design. Instead, we rely on estimatesobtained from embedding a pseudo-random message into a random cover object. The author wasunable to find a better algorithm than an exhaustive search guided by some simple design rules.First, ˆH should not have identical columns because the syndrome trellis would contain two ormore different paths with exactly the same weight, which would lead to an overall decrease inperformance. By running an exhaustive search over small matrices, we have observed that the bestsub matrices ˆH had ones in the first and last rows. For example, when h = 7 and w = 4, morethan 97% of the best 1000 codes obtained from the exhaustive search satisfied this rule. Thus, wesearched for good matrices among those that did not contain identical columns and with all bitsin the first and last rows set to 1 (the remaining bits were assigned at random). In practice, werandomly generated 10 − 1000 sub matrices satisfying these rules and estimated their performance(embedding efficiency) experimentally by running the Viterbi algorithm with random covers andmessages. For a reliable estimate, cover objects of size at least n = 106 are required.To investigate the stability of the design w.r.t. to the profile, the following experiment wasconducted. We fixed h = 10 and w = 2, which correspond to a code with α= 1⁄2. The codedesign procedure was simulated by randomly generating 300 sub matrices ˆH1. . . ˆH300 satisfyingthe above design rules. The goodness of the code was evaluated using the embedding efficiency($e = m/D(\mathbf{x}, \mathbf{y})$) by running the Viterbi algorithm on a random cover object (of size n = 106) andwith a random message. The codes with a high embedding efficiency on the constant profile exhibithigh efficiency for the other profiles, we consider the code design to be stable w.r.t. the profile anduse these matrices with other profiles in practice. All further results are generated by using thesematrices.

7. Conclusions

The concept of embedding in steganography that minimizes a distortion function is connected tomany basic principles used for constructing embedding schemes for complex cover sources today, including the principle of minimal-embedding-impact ,approximate model-preservation ,or the Gibbs construction . The current work describes a complete practical framework forconstructing steganographic schemes that embed by minimizing an additive distortion function. Once the steganographer specifies the form of the distortion function, the proposed frameworkprovides all essential tools for constructing practical embedding schemes working close to theirtheoretical bounds. The methods are not limited to binary embedding operations and allow theembedder to choose the amplitude of embedding changes dynamically based on the cover-imagecontent. The distortion function or the embedding operations do not need to be shared with therecipient. In fact, they can even change from image to image. The framework can be thought ofas an off-the-shelf method that allows practitioners to concentrate on the problem of designing the distortion measure instead of the problem of how to construct practical embedding schemes.The merit of the proposed algorithms is demonstrated experimentally by implementing them forthe JPEG and spatial domains and showing an improvement in statistical detectability as measured by state-of-the-art blind steganalyzers. We have demonstrated that larger embedding changes provide a significant gain in security when placed adaptively. Finally, the construction is not limitedto embedding with larger amplitudes but can be used, e.g., for embedding in color images, wherethe LSBs of all three colors can be seen as 3-bit symbols on which the cost functions are defined. Applications outside the scope of digital images are possible as long as we know how to define thecosts.The implicit premise of this chapter is the direct relationship between the distortion function *D*and statistical detectability. Designing (and possibly learning) the distortion measure for a givencover source is an interesting research problem by itself. Examples of distortion measures presentedin this work are unlikely to be optimal and we include them here mainly to illustrate the concepts.

9. References

1) Tomas Filler, Jan Judas, and Jessica Fridrich, "Minimizing Additive Distortion in Steganography Using Syndrome-Trellis Codes", IEEE Transactionson Information Forensics and Security, Vol.6, No.3, September 2011.

2) Tomas Filler "ImperfectStegosystems – Asymptotic Laws And Near-Optimal Practical Constructions"

3) Michael Arnold, Martin Schmucker, Stephen D. Wolthusen "Techniques and applicationsof Digital Watermarkingand Content Protection"

4) Hui-Yu Huang, Chi-Hung Fan, and Wen-Hsing Hsu "An effective watermark embedding algorithm for high JPEG compression" MVA2007 IAPR Conference on Machine Vision Applications, May 16-18, 2007, Tokyo, JAPAN

5) Diljith M. Thodi and Jeffrey J. Rodríguez, "Expansion Embedding Techniques for Reversible Watermarking" IEEE Transactions On Image Processing, Vol. 16, No. 3, March 2007

6) Toshio Modegi and Makoto Chiba "Nearly Lossless Audio Watermark Embedding echniques to be Extracted Contactlessly by Cell Phone" ICMU2006

7) R. Böhme, "Improved Statistical Steganalysis UsingModels of Heterogeneous Cover Signals," Ph.D. dissertation, Faculty of Comput. Science Technische Universität, Dresden, Germany, 2008.

8) Y. Kim, Z. Duric, and D. Richards, "Modified matrix encoding technique for minimal distortion steganography," in *Proc. 8th Int. Workshop Inf. Hiding*, J. L. Camenisch, C. S. Collberg, N. F. Johnson, and P. Sallee, Eds., Alexandria, VA, Jul. 10–12, 2006, vol. 437, Lecture Notes in Computer Science, pp. 314–327.

9) R. Zhang, V. Sachnev, and H. J. Kim, "Fast BCH syndrome coding for steganography," in *Proc. 11th Int. Workshop Inf. Hiding,*, S. Katzenbeisser and A.-R. Sadeghi, Eds., Darmstadt, Germany, Jun. 7–10, 2009, vol. 5806, Lecture Notes in Computer Science, pp. 31–47.

10) V. Sachnev, H. J. Kim, and R. Zhang, "Less detectable JPEG steganography method based on heuristic optimization and BCH syndrome coding," in *Proc. 11th ACM Multimedia Security Workshop*, J. Dittmann, S. Craver, and J. Fridrich, Eds., Princeton, NJ, Sep. 7–8, 2009, pp. 131–140.

11) T. Pevný, T. Filler, and P. Bas, "Using high-dimensional image models to perform highly undetectable steganography," in *Proc. 12th Int. Workshop Inf. Hiding*, P. W. L. Fong, R. Böhme, and R. Safavi-Naini, Eds., Calgary, Canada, Jun. 28–30, 2010, vol. 6387, Lecture Notes in Computer Science, pp. 161–177.

12) J. Kodovský and J. Fridrich, "On completeness of feature spaces in blind steganalysis," in *Proc. 10th ACMMultimedia SecurityWorkshop*, A. D. Ker, J. Dittmann, and J. Fridrich, Eds., Oxford, U.K., Sep. 22–23, 2008, pp. 123–132.

13) T. Filler and J. Fridrich, "Gibbs construction in steganography," *IEEE Trans.*

Inf. Forensics Security, vol. 5, pp. 705–720, Sep. 2010.