

On the Complexity and Hardness of the Steganography Embedding Problem

R. Chandramouli ^a Shalin Trivedi ^a and R. N. Uma ^b

^aMultimedia Systems, Communication and Networking (MSyNC) Laboratory

^aDepartment of Electrical and Computer Engineering

^aStevens Institute of Technology

^bDepartment of Computer Science

^bUniversity of Texas at Dallas

ABSTRACT

We analyze the complexity of the *steganography problem* and show that the decision version of the problem is NP-complete through transformation from the *Knapsack problem*. We also give a pseudo-polynomial time algorithm to optimally solve the *steganography problem*. This optimal algorithm can also be applied to image steganography.

1. INTRODUCTION

Steganography deals with hiding messages in a cover signal such that it can be extracted at the receiving side with the help of a secret key. Applications of steganography include watermarking and fingerprinting that seem to hold promise for copyright protection, tracing source of illegal copies, etc. There are several issues to be considered when studying steganographic systems. One among the key performance measures used to compare different message embedding algorithms is *steganography capacity*. In a general sense, it is the maximum message size that can be embedded subject to certain constraints. A number of ways to compute the steganography capacity using information theory, perceptually based methods, and detection theory have been proposed previously ([1–7] and references therein). Recent work discussed in [8] gives a capacity measure in the presence of a steganalysis detector. Whatever definition one adopts for the capacity measure, the common goal of steganography algorithms is to maximize this capacity measure subject to distortion bounds.

One important question in steganography is: what is the trade-off involved in attempting to embed larger and larger message sizes? This question can be answered in several different ways. In [8], it is shown that when the embedding message sizes are larger than a threshold then it becomes easier for a steganalysis algorithm to detect the presence/absence of a hidden message. This defeats the purpose of steganography where the idea is to hide messages in a cover signal such that its very presence can be concealed. For steganography applications such as watermarking, a higher watermark length could mean higher embedding induced distortion or less robustness to attacks [3]. Therefore, if constraints such as security, upper bound on embedding induced distortion, etc. are given, it is important for the embedding algorithm to be able to systematically embed the message such that its length is maximized subject to these bounds. This raises the key question: *what is the complexity of the general steganography embedding problem?* It appears that this fundamental question has not been addressed so far. Answers to this question could have serious implications on embedding algorithm design and analysis. In this paper, we attempt to address this issue and provide some initial thoughts. Without loss of generality we assume the host signal to be a digital image.

The paper is organized as follows. In Section 2 we formulate the *steganography problem*. We give the transformation from the *knapsack problem* and analyze the complexity in Section 3. In Section 4 we present the pseudo-polynomial time algorithm and conclude in Section 5.

Further author information:

R. Chandramouli: E-mail: rchandr1@stevens.edu, Telephone: 1 201 216 8642, Fax 1 201 216 8246

R. N. Uma: E-mail: rnuma@utdallas.edu

Shalin Trivedi: E-mail: shalin_trivedi@yahoo.com

2. STEGANOGRAPHY MODEL

Consider the following model for data embedding in digital images. Let $\{x_1, x_2, \dots, x_n\}$ denote n features (*symbols*) of the original cover image. These features could be in the spatial domain, frequency domain or a combination thereof. For example, for least significant bit encoding, x_i 's correspond to the image pixel intensity values and for spread spectrum steganography [3] these are the discrete cosine transform coefficients. By embedding a subset of the message (perhaps modulating a message carrier) in x_i let the resulting distortion induced to the host feature be d_i . Note that the total embedding induced distortion is usually bounded above, say, by D , due to perceptual and information theoretic (ϵ -security [9]) considerations. Let the number of message bits embedded in i^{th} original image symbol be m_i . We can formulate the *steganography problem* using 0-1 decision variables y_i where

$$y_i = \begin{cases} 1 & \text{if symbol } x_i \text{ is selected for embedding} \\ 0 & \text{otherwise} \end{cases}$$

Then, the steganography problem can be stated as follows:

$$\text{maximize } \sum_{i=1}^n m_i y_i \tag{1}$$

subject to

$$\sum_{i=1}^n d_i y_i \leq D; \tag{2}$$

$$y_i \in \{0, 1\}, i = 1, 2, \dots, n. \tag{3}$$

The goal is to maximize the length of the embedded message size in the image, given by the objective function (1). Constraint (2) ensures that the embedding induced distortion does not exceed the specified bound D . Here, we have implicitly assumed that each host symbol can carry a variable number of message size. Without loss of generality we also make the following assumptions about the coefficients m_i , d_i and D .

- These coefficients are non-negative integers; fractional values can be transformed to integers after multiplication by a suitable factor.
- $m_i > 0, \forall i$, if not, the corresponding feature need not be considered.
- $0 < d_i < D$. If $d_i = 0$ then it does not affect the solution and, if $d_i > D$ the corresponding image source symbol x_i will not be considered as a candidate for embedding.
- $0 < D < \sum_{i=1}^n d_i$. For, if $D = 0$ then we get the trivial solution by setting $y_i = 0, \forall i$ and, if $D \geq \sum_{i=1}^n d_i$ set $y_i = 1, \forall i$.

3. COMPLEXITY OF STEGANOGRAPHY PROBLEM

In this section, we give a transformation from the *knapsack problem* to the *steganography problem* and show that the *steganography problem* is NP-complete.

The decision version of the *knapsack problem* is stated as follows [10]:

Given a finite set S of items, a weight w_i and value v_i for each item, and positive integers W and K , does there exist a subset $S' \subseteq S$ such that $\sum_{i \in S'} w_i \leq W$ and such that $\sum_{i \in S'} v_i \geq K$?

We can recast the *steganography problem* (1)-(3) as the following decision problem:

Given a finite set S of features, an embedding induced distortion d_i and length of embedded message m_i for each image feature x_i , and positive integers D and L , does there exist a subset $S' \subseteq S$ such that $\sum_{i \in S'} d_i \leq D$ and such that $\sum_{i \in S'} m_i \geq L$?

Henceforth, *knapsack problem* and *steganography problem* will refer to their corresponding decision versions.

We can now state the following lemmas:

LEMMA 3.1. *The steganography problem is in NP.*

Proof. Given an embedding P , as a list of features in the image that are embedded, and an integer L , we can easily verify in polynomial time that the embedding induced distortion does not exceed D , i.e., $\sum_{i \in P} d_i \leq D$, and that the total length of the embedded message $\sum_{i \in P} m_i \geq L$. Therefore, the *steganography problem* is in NP. \square

LEMMA 3.2. *The steganography problem is NP-hard.*

Proof. To show that the *steganography problem* is NP-hard, we have to show that the *knapsack problem* can be transformed in polynomial time to the *steganography problem* and that the *knapsack problem* has a solution of value $\geq K$ iff the *steganography problem* has an embedding of length $\geq K$.

The transformation from the *knapsack problem* is as follows. Item i of the *knapsack problem* corresponds to feature x_i of the *steganography problem*. The distortion bound D is set to the knapsack capacity W . Weight w_i of item i corresponds to the embedding induced distortion d_i (i.e., set $d_i = w_i$). The value v_i of item i corresponds to the length of the message m_i used to embed in x_i . Clearly, the length of m_i and distortion d_i are independent quantities because the same m_i could give rise to different d_i values depending upon how the embedding is done. The transformation is trivially in polynomial time.

Let S' be a solution to the *knapsack problem*. Embedding the features corresponding to the items selected in S' gives a solution to the *steganography problem* with embedding length $\geq K$.

Let P be a solution to the *steganography problem* of total embedding length $\geq K$. Selecting items corresponding to features that are embedded gives a solution to the *knapsack problem* of value at least K .

Therefore the transformation is valid and hence the *steganography problem* is NP-hard. \square

Therefore, we have,

THEOREM 3.3. *The steganography problem is NP-complete.*

Proof. Follows from Lemma 3.1 and Lemma 3.2. \square

4. PSEUDO-POLYNOMIAL TIME EMBEDDING ALGORITHM

We give a dynamic programming solution (similar to [11,12]) to optimally solve the *steganography problem*.

Order the features arbitrarily as $x_1, x_2, x_3, \dots, x_n$. Let L be a three-dimensional matrix which can be viewed as slices of two-dimensional matrices. The rows of the two-dimensional matrices are labelled with subsets of features: row i corresponds to the subset $\{x_1, x_2, \dots, x_i\}$ and row $i + 1$ corresponds to the subset $\{x_1, x_2, \dots, x_i, x_{i+1}\}$. The columns are labelled 1 through D . The slices are labelled with the corresponding message index: slice k corresponds to the first k bits of the message being considered for embedding. Entry $L[i, d, k]$ corresponds to the optimal value of the message length for embedding a subset of features x_1, \dots, x_i with embedding induced distortion not to exceed d when the message under consideration is the first k bits. As boundary conditions, we have, $L[i, d, 0] = 0$, $L[0, d, k] = 0$ and $L[i, 0, k] \in \{0, \dots, k\}$ for values of $i = 1, \dots, n$, $d = 1, \dots, D$ and $k = 1, \dots, \ell$ where ℓ is the maximum message length (which is an upper bound on the total length that can be embedded within the allowed distortion).

Say we have computed the optimal lengths when features x_1, \dots, x_i and the first k message bits are considered. Next we consider feature x_{i+1} . We have two choices — we can either embed feature x_{i+1} or not embed feature x_{i+1} . If we do not embed feature x_{i+1} , then the current optimal value $L[i + 1, j, k] = L[i, j, k]$. If we embed feature x_{i+1} , then the next question is how many bits can we embed to comply with the distortion limit. The distortion of embedding features x_1, \dots, x_i should not exceed $d - d_{i+1,c}$ since we need to guarantee that we do

not violate the distortion bound when feature x_{i+1} is embedded. $d_{i+1,c}$ denotes the distortion generated by embedding the last c message bits (i.e., message bits $k - c + 1$ through k) in the c least significant bits of pixel $i + 1$. Ideally, c can take any value from 1 to 8, but in reality, due to perceptual considerations, we will restrict c to take any value in the range 1 through 4, for example. Note that maximum distortion generated by embedding LSB is minimum among all the distortions generated by changing all other bits of the given pixel. So we will assume that LSB is embedded first. Writing it more formally,

$$L[i + 1, d, k] = \max \begin{cases} L[i, d, k] \\ \text{Do not embed feature } x_{i+1} \\ L[i, d - d_{i+1,c}, k - c] + c, c = 1, \dots, 4 \\ \text{Embed feature } x_{i+1} \text{ with } c \text{ bits} \end{cases} \quad (4)$$

The boundary conditions are given by:

$$\begin{aligned} L[i, d, 0] &= 0, i = 0, \dots, n, d = 0, \dots, D \\ L[0, d, k] &= 0, d = 0, \dots, D, k = 0, \dots, \ell \\ L[i, 0, k] &\in \{0, \dots, k\}, i = 0, \dots, n, k = 0, \dots, \ell \end{aligned} \quad (5)$$

Note that in Eq. 4 we will ignore those entries of L matrix for which $d - d_{i+1,c}$ is less than zero or $k - c$ is less than zero.

The optimal value is in entry $L[n, D, \ell]$. The running time of the algorithms is $O(nD\ell)$ which is a pseudo-polynomial time algorithm.

5. CONCLUSIONS

To the best of our knowledge, our paper is the first to discuss the complexity and hardness of the *steganography embedding problem*; in particular, we show that the decision version of the problem is NP-complete. Furthermore, we have given a pseudo-polynomial time algorithm to solve the embedding problem optimally. The fact that we have a pseudo-polynomial time algorithm essentially implies that the *steganography problem* is *weakly* NP-complete; this also follows from a similar result for the knapsack problem.

ACKNOWLEDGMENTS

This material is based on research sponsored by National Science Foundation under agreement number NSF DAS 0242417.

REFERENCES

1. R. Chandramouli, "Data hiding capacity in the presence of an imperfectly known channel," *SPIE Proceedings of Security and Watermarking of Multimedia Contents II* **4314**, 2001.
2. P. Mouli and M. Mihcak, "The data hiding capacity of image sources," *preprint available at <http://www.ifp.uiuc.edu/~moulin/paper.html>*.
3. M. Miller, I. Cox, and J. Bloom, *Digital Watermarking*, Morgan Kaufman, 2001.
4. L. Yung and S. Chang, "Watermarking capacity of digital images based on domain specific masking effects," *Proc. International Conference on Information Technology*, pp. 90–94, 2001.
5. R. Wolfgang, C. Podilchuk, and E.J.Delp, "Perceptual watermarks for digital images and video," *Proceedings of IEEE* **87**, July 1999.
6. M. Barni, Bartolini, and F. Piva, "Improved wavelet-based watermarking through pixel-wise masking," *IEEE Trans. on Image processing* **10**(5), pp. 783–791, 2001.
7. S. Somasundaram and R. Chandramouli, "Perceptually based waterfilling for watermarking," *Proc. ISCAS*, 2002.

8. R. Chandramouli and N. Memon, "Analysis of lsb image steganography techniques," *Proc. ICIP* **3**, pp. 1019–1022, 2001.
9. C. Cachin, "An information-theoretic model for steganography," *Proc. 2nd International Workshop Information Hiding LNCS 1525*, pp. 306–318, 1998.
10. M. Garey and D. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W.H. Freeman and Company, New York, 1979.
11. G. B. Dantzig, "Discrete-variable extremum problems," *Operations Research* **5**, pp. 266–277, 1957.
12. E. Lawler, *Combinatorial Optimization: Networks and Matroids*, Holt, Rinehart and Winston, 1976.