COMPLETE CHARACTERIZATION OF PERFECTLY SECURE STEGO-SYSTEMS WITH MUTUALLY INDEPENDENT EMBEDDING OPERATION

Tomáš Filler and Jessica Fridrich

Department of ECE, SUNY Binghamton, Binghamton, NY 13902-6000, USA
{tomas.filler, fridrich}@binghamton.edu

ABSTRACT

Without any assumption on the cover source, this paper presents a complete characterization of all perfectly secure stego-systems that employ mutually independent embedding operation. It is shown that for a fixed embedding operation, the only perfectly secure stego-systems are those whose cover distribution is an element of a linear vector space with basis vectors determined by the embedding operation. Moreover, we also prove that such stego-systems are perfectly secure if and only if the Fisher information with respect to the embedding change rate is zero and thus Fisher information can be seen as an equivalent descriptor of steganographic security. This result is important for deriving steganographic capacity of imperfect stego-systems with covers modeled as Markov chains [1]. It also suggests that Fisher information could be used for benchmarking.

Index Terms— steganography, perfect security, mutually independent embedding

1. INTRODUCTION

In steganography, the sender and receiver communicate by hiding their messages in generally trusted media, such as digital images, so that one cannot distinguish between the original (cover) objects and the objects carrying the message (stego objects). Formally, the security of a stego-system is evaluated using the Kullback-Leibler divergence between the distributions of cover and stego objects [2]. Systems with zero KL divergence are called perfectly secure.

Formally, a stego-system is a combination of an embedding algorithm and a cover source. The vast majority of practical stego-systems hide messages by modifying individual cover elements using mutually independent embedding operations, e.g., LSB and ± 1 embedding, F5 algorithm, perturbed quantization, MMx, stochastic modulation, and many others (see [3] and the references therein).

In this paper, we provide a complete characterization of perfectly secure stego-systems for the class of embedding algorithms that employ mutually independent (MI) embedding operations. The cover distributions of all perfectly secure systems form a linear vector space spanned by distributions determined by the embedding operation. Moreover, we show that perfect security (zero KL divergence) is equivalent to satisfying a simple condition related to Fisher information. This result suggests that Fisher information can be used as an equivalent descriptor of steganographic security.

In Section 2, we introduce the notation and definitions and review some preliminary facts. Section 3 and Section 4 contain the main results, as well as illustrative examples. Section 5 states the main results for the special case of Markov chain cover sources. Section 6 concludes the paper.

2. NOTATION, PRELIMINARIES, AND ASSUMPTIONS

We use $x_1^n \triangleq (x_1, \ldots, x_n) \in \mathcal{X}^n$, $\mathcal{X} = \{1, \ldots, N\}$ to represent an n-element cover object, obtained as a realization of random variable $X_1^n \sim P$ where P is the distribution of covers over $\mathcal{X}^n$. Similarly, the stego object $y_1^n \triangleq (y_1, \ldots, y_n) \in \mathcal{X}^n$ is a realization of random variable $Y_1^n \sim Q_\beta$, where β is a scalar parameter capturing the extent of embedding changes (It will be helpful to think of β as the change rate.).

The definition of steganographic security was given by Cachin [2].

Definition 1 *Steganography is* perfectly secure *iff*

$$d(\beta) \triangleq D_{KL}(P||Q_\beta) = \sum_{y_1^n \in \mathcal{X}^n} P(y_1^n) \log \frac{P(y_1^n)}{Q_\beta(y_1^n)} = 0,$$

or ϵ-secure *if* $d(\beta) \le \epsilon$.

We assume that the impact of embedding with parameter $\beta \in [0, \beta_0]$ on the k-th element can be captured using the matrix $b_{i,j}(\beta) \triangleq Pr(Y_k = j | X_k = i) = \delta_{i,j} + \beta c_{i,j}$, for some constants $c_{i,j} \ge 0$ for $i \ne j$, $c_{i,i} = -\sum_j c_{i,j}$, where $\delta_{i,j}$ is the Kronecker delta. In a matrix form, $\mathbb{B}_\beta = \mathbb{I} + \beta\mathbb{C}$, where $\mathbb{B}_\beta \triangleq (b_{i,j}(\beta))$, $\mathbb{I}$ is the identity matrix, and $\mathbb{C} \triangleq (c_{i,j})$. We further assume that embedding operations are mutually independent, $Pr(Y_1^n|X_1^n) = \prod_{k=1}^n Pr(Y_k|X_k)$. By the definition of $b_{i,j}$, the matrix $\mathbb{B}_\beta$ is stochastic, $\sum_j b_{i,j} = 1$. Finally, we assume that $b_{i,i}(\beta) > 0$ for all $\beta \in [0, \beta_0]$. The matrix $\mathbb{B}_\beta$ represents an embedding algorithm with MI embedding operation (simply MI embedding). Many embedding methods can be formulated within this framework (see examples in Figure 1).

To simplify the language in this paper, we will speak of security of a cover source w.r.t. a given MI embedding meaning that the *cover source is perfectly secure w.r.t.* $\mathbb{B}$, if the resulting stego-system is perfectly secure. It does then make sense to inquire about all possible perfectly secure cover sources w.r.t. MI embedding with matrix $\mathbb{B}_\beta$.

We now review some results from the theory of ergodic classes borrowed from [4] that will be later applied to the stochastic matrix $\mathbb{B}_\beta$. For states $i, j \in \mathcal{X}$, we call j *a consequent* of i (of order k) $(i \to j)$ iff $\exists k, (\mathbb{B}_\beta^k)_{i,j} \ne 0$. State $i \in \mathcal{X}$ is *transient* if it has a consequent of which it is not itself a consequent, i.e., $\exists j \in \mathcal{X}$ such that $(i \to j) \Rightarrow (j \not\to i)$. We say $i \in \mathcal{X}$ is *non-transient* if it is a consequent of every one of its consequents, $\forall j \in \mathcal{X}$, $(i \to j) \Rightarrow$

This research was supported by Air Force Office of Scientific Research (AFOSR) under the research grant FA9550-08-1-0084. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation there on. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of AFOSR or the U.S. Government.

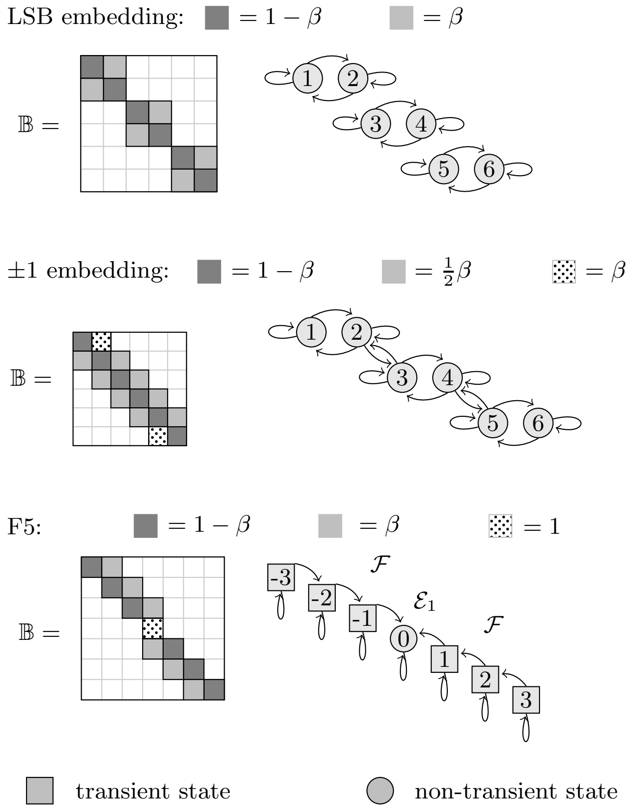
LSB embedding: ■ = 1 − β ▨ = β
𝔹 =
1
2
3
4
5
6
±1 embedding: ■ = 1 − β ▨ = ½β ▩ = β
𝔹 =
1
2
3
4
5
6
F5: ■ = 1 − β ▨ = β ▩ = 1
𝔹 =
-3
-2
-1
0
1
2
3
ℱ
ℰ1
ℱ
transient state
non-transient state

Fig. 1. Examples of several embedding methods and their ergodic classes.

$(j \to i)$. The set $\mathcal{X}$ can be decomposed as $\mathcal{X} = \mathcal{F} \cup \mathcal{E}_1 \cup \dots \cup \mathcal{E}_k$, where $\mathcal{F}$ is the set of all transient states and $\mathcal{E}_a$, $a \in \{1, \dots, k\}$, are so called ergodic classes. We put two non-transient states into one ergodic class if they are consequents of each other.

Let matrix $\mathbb{B}_\beta$ have k ergodic classes. Then, there exist k linearly independent left eigenvectors, denoted as $\pi^{(1)}, \dots, \pi^{(k)}$, of matrix $\mathbb{B}_\beta$ corresponding to eigenvalue 1, called *invariant distributions*. If $\pi^{(a)}\mathbb{B}_\beta = \pi^{(a)}$, for some $a \in \{1, \dots, k\}$, then $\pi_i^{(a)} > 0$ for all $i \in \mathcal{E}_a$, and $\pi_i^{(a)} = 0$ otherwise. Every other π satisfying $\pi\mathbb{B}_\beta = \pi$ is obtained by a convex linear combination of $\{\pi^{(a)} | a \in \{1, \dots, k\}\}$. For a complete reference, see [4, Chapter V, §2]. The set of ergodic classes for matrix $\mathbb{B}_\beta$ depends only on the set $\{(i,j) | b_{i,j}(\beta) \neq 0\}$. Since $b_{i,j}(\beta) = 0$ iff $c_{i,j} = 0$ for $i \neq j$ and $b_{i,i}(\beta) > 0$ for $\beta \in (0, \beta_0]$, the structure of ergodic classes does not depend on β. Moreover, if $\pi\mathbb{B}_\beta = \pi$ for some $\beta > 0$, then $\pi\mathbb{C} = 0$ and thus all invariant distributions are independent of β, because $\pi\mathbb{B}_{\beta'} = \pi\mathbb{I} + \beta'\pi\mathbb{C} = \pi\mathbb{I} = \pi$. By this reason, we frequently omit the index β.

3. PERFECTLY SECURE COVER SOURCES UNDER MUTUALLY INDEPENDENT EMBEDDING OPERATION

In this section, we let matrix $\mathbb{B}$ represent an arbitrary MI embedding with k ergodic classes $\mathcal{E}_a$ and invariant distributions $\pi^{(a)}$, $a \in \{1, \dots, k\}$. The following example describes a construction of perfectly secure cover sources w.r.t. $\mathbb{B}$.

Example 2 [Perfectly secure cover sources] *Let $P^{(2)}$ be a probability distribution on 2-element cover objects defined as $P^{(2)}(X_1^2 = (i,j)) = \pi_i^{(a)}\pi_j^{(b)}$ for some $a, b \in \{1, \dots, k\}$. Then $P^{(2)}$ is a perfectly secure cover source w.r.t. $\mathbb{B}$ because*

$$Q_\beta^{(2)}\big(Y_1^2 = (i,j)\big) = \Big(\sum_{\hat{i}} b_{\hat{i},i} P(X_1 = \hat{i})\Big)\Big(\sum_{\hat{j}} b_{\hat{j},j} P(X_2 = \hat{j})\Big)$$

$$= \big(\pi^{(a)}\mathbb{B}\big)_i \big(\pi^{(b)}\mathbb{B}\big)_j = \pi_i^{(a)}\pi_j^{(b)} = P^{(2)}\big(X_1^2 = (i,j)\big),$$

and thus both distributions $P^{(2)}$, and $Q_\beta^{(2)}$ are identical, which implies perfect security. Since this construction does not depend on the particular choice of $a, b \in \{1, \dots, k\}$, we can create k^2 perfectly secure cover sources w.r.t. $\mathbb{B}$. The probability distributions $P^{(2)}$ obtained from this construction are linearly independent and form a k^2-dimensional linear vector space. By a similar construction, we can construct k^n n-element linearly independent perfectly secure cover sources w.r.t. $\mathbb{B}$.

We next show that there are no other linearly independent perfectly secure cover sources w.r.t. $\mathbb{B}$.

Theorem 3 [Mutually independent embedding] *There are exactly k^n linearly independent perfectly secure probability distributions P on n-element covers. Every perfectly secure probability distribution P w.r.t. $\mathbb{B}$ can be obtained by a convex linear combination of k^n linearly independent perfectly secure distributions described in Example 2.*

Proof It is sufficient to prove that there cannot be more than k^n linearly independent perfectly secure probability distributions P on n-element covers. We show the proof for $n = 2$ and later present its generalization.

We define the following matrices $\mathbb{P} \triangleq (p_{i,j})$, $p_{i,j} = P(X_1^2 = (i,j))$, and $\mathbb{Q} \triangleq (q_{i,j})$, $q_{i,j} = Q_\beta(Y_1^2 = (i,j))$. By defininition of MI embedding, we have

$$q_{ij} = \sum_{(v,w) \in \mathcal{X}^2} Q_\beta(Y_1^2 = (i,j) | X_1^2 = (v,w)) P(X_1^2 = (v,w))$$

$$= \sum_{v,w \in \mathcal{X}} b_{vi} b_{wj} p_{vw}.$$

Define matrix $\mathbb{D} \triangleq (d_{u_1^2, v_1^2})$ of size $N^2 \times N^2$, where $d_{u_1^2, v_1^2} = b_{u_1,v_1} b_{u_2,v_2}$. If $\vec{p}$ is defined as one big row vector of elements $p_{i,j}$ and similarly $\vec{q}$, then assuming perfect security of cover source w.r.t. $\mathbb{B}$ ($\mathbb{P} = \mathbb{Q}$), we have $\vec{q} = \vec{p}\mathbb{D} = \vec{p}$ and thus $\vec{p}$ is left eigenvector of $\mathbb{D}$ corresponding to 1. Matrix $\mathbb{D}$ is stochastic and thus it is sufficient to show that it has k^2 ergodic classes.

We first show that

$$u_1^2 \overset{(m)}{\rightarrow} v_1^2 \Leftrightarrow (u_1 \overset{(m)}{\rightarrow} v_1) \text{ and } (u_2 \overset{(m)}{\rightarrow} v_2),\ \ u_1^2, v_1^2 \in \mathcal{X}^2. \qquad (1)$$

By $u_1^2 \overset{(m)}{\rightarrow} v_1^2$ we mean that v_1^2 is a consequent of u_1^2 of order m in terms of matrix $\mathbb{D}$. If $u_1^2 \overset{(m)}{\rightarrow} v_1^2$, then there exist $m-1$ intermediate states ${}_1w_1^2, \dots, {}_{m-1}w_1^2$, such that $d_{u,{}_1w} d_{{}_1w,{}_2w} \cdots d_{{}_{m-1}w,v} > 0$. Since $d_{u_1^2,v_1^2} = b_{u_1,v_1} b_{u_2,v_2}$, this implies the existence of both paths $u_i \overset{(m)}{\rightarrow} v_i$ of order m, $i = 1, 2$. The converse is true by the same reason.

We show that $\mathcal{E}_a \times \mathcal{E}_b$, $a, b \in \{1, \dots, k\}$ are the only ergodic classes. If $u_1 \overset{(m_1)}{\rightarrow} v_1$ and $u_2 \overset{(m_2)}{\rightarrow} v_2$, then $u_1^2 \overset{(m_1+m_2)}{\rightarrow} v_1^2$ for all $u_1, v_1 \in \mathcal{E}_a$ and $u_2, v_2 \in \mathcal{E}_b$, because the path from u_i to v_i can be arbitrarily extended by adding self loops of type $j \to j$ since all diagonal terms $b_{j,j}$ are positive and thus by (1) we have $u_1^2 \overset{(m_1+m_2)}{\longrightarrow} v_1^2$. Finally by $u_1, v_1 \in \mathcal{E}_a$ and $u_2, v_2 \in \mathcal{E}_b$, $v_i \to u_i$ and by the same argument $v_1^2 \to u_1^2$, and therefore $\mathcal{E}_a \times \mathcal{E}_b$ are ergodic classes. Any other state $u_1^2 \in \mathcal{E}_a \times \mathcal{F} \cup \mathcal{F} \times \mathcal{E}_a \cup \mathcal{F} \times \mathcal{F}$ must be transient w.r.t. $\mathbb{D}$, otherwise by (1) we obtain contradiction with $u_i \in \mathcal{F}$ for some i.

This proof can be generalized for $n \geq 3$ by proper definition of matrices $\mathbb{P}$, $\mathbb{Q}$, and $\mathbb{D}$. In general, matrix $\mathbb{D}$ has size $N^n \times N^n$. By similar construction we obtain k^n ergodic classes of generalized matrix $\mathbb{D}$, however we know k^n linearly independent distributions. ■

4. PERFECT SECURITY AND FISHER INFORMATION

In this section, we show that for stego-systems with MI embedding perfect security can be captured using Fisher information. From Taylor expansion of KL divergence, for small β, $d(\beta) = \frac{1}{2}\beta^2 I(0) + O(\beta^3)$ where $I(0) = \partial^2 d(\beta)/\partial\beta^2|_{\beta=0}$ is the Fisher information w.r.t. β. If for some stego-system $d(\beta) = 0$ for $\beta \in [0, \beta_0]$, then $I(0) = 0$ from the Taylor expansion. Even though the opposite does not hold in general, we will prove that for MI embedding zero Fisher information implies perfect security. In other words, a stego-system with MI embedding is perfectly secure for $\beta \in [0, \beta_0]$ if and only if $I(0) = 0$. This provides us with a simpler condition for verifying perfect security than the KL divergence. Fisher information also provides a connection to quantitative steganalysis because $1/I(\beta)$ is the lower bound on variance of unbiased estimators of β. Moreover, $I(0)$ could be used for comparing (benchmarking) stego-systems.

We start by reformulating the condition $I(0) = 0$.

Proposition 4 *Let P and Q_β be probability distributions of cover and stego objects with n elements embedded with parameter β. The Fisher information is zero if and only if the FI-condition is satisfied*

$$\forall y_1^n \in \mathcal{X}^n \quad \Big(P(X_1^n = y_1^n) > 0\Big) \Rightarrow \Big(\frac{d}{d\beta} Q_\beta(y_1^n)\big|_{\beta=0} = 0\Big). \tag{2}$$

Proof The second derivative of $d(\beta)$ at β, $d''(\beta)$, can be written as

$$I(\beta) = -\sum_{y_1^n \in \mathcal{X}^n} P(y_1^n)\Big(\frac{Q''_\beta(y_1^n)}{Q_\beta(y_1^n)} - \Big(\frac{Q'_\beta(y_1^n)}{Q_\beta(y_1^n)}\Big)^2\Big), \tag{3}$$

where $Q'_\beta(y_1^n) = \frac{\partial}{\partial\beta} Q_\beta(y_1^n)$. By $P(y_1^n) = Q_{\beta=0}(y_1^n)$, the first term in the bracket in (3) sums to zero at $\beta = 0$, and thus $I(0)$ is zero iff $Q'_\beta(y_1^n)\big|_{\beta=0} = 0$ is zero for all $y_1^n \in \mathcal{X}^n$ for which $P^{(n)}(y_1^n) > 0$ as was to be proved. Here, we assume the KL divergence $d(\beta)$ to be continuous w.r.t. β which is valid by the construction of the matrix $\mathbb{B}$. ■

The next theorem shows that the FI condition (2) is equivalent with perfect security for MI embedding.

Theorem 5 [Fisher information condition] *There are exactly k^n linearly independent probability distributions P on n-element covers satisfying the FI condition* (2). *These distributions are perfectly secure w.r.t. $\mathbb{B}$. Every other probability distribution P satisfying* (2) *can be obtained by a convex linear combination of k^n linearly independent perfectly secure distributions.*

Proof From Example 2, we know k^n linearly independent perfectly secure distributions. By Taylor expansion of $d(\beta)$, these distributions satisfy the FI condition, because $d(\beta) = 0 \Rightarrow I(0) = 0$. It is sufficient to show that there cannot be more linearly independent distributions satisfying the FI condition.

Similarly as in the previous proof, we reformulate the theorem as eigenvector problem and use ergodic class theory to give the exact number of left eigenvectors corresponding to 1. Again, we present the proof for the case $n = 2$ and then show how to generalize it.

If P satisfies (2), then the linear term in the Taylor expansion of $Q_\beta(y_1^2)$ w.r.t. β is zero. By the independence property, $(Q(y_1^n|x_1^n) = \prod_{i=1}^n Q(y_i|x_i))$, and the form of matrix $\mathbb{B}$ ($\mathbb{B}_\beta = \mathbb{I} + \beta\mathbb{C}$), condition (2) has the following form

$$\begin{aligned}\frac{dQ_\beta(y_1^2)}{d\beta}\Big|_{\beta=0} &= \lim_{\beta\to 0} \sum_{x_1^2 \in \mathcal{X}^2} P(x_1^2) \frac{d}{d\beta} \prod_{i=1}^{2} Q_\beta(y_i|x_i) \\ &= \sum_{x_1 \in \mathcal{X}} c_{x_1,y_1} P(x_1, y_2) + \sum_{x_2 \in \mathcal{X}} c_{x_2,y_2} P(y_1, x_2) = 0. \end{aligned} \tag{4}$$

We define matrix $\mathbb{P} \triangleq (p_{i,j})$ as $p_{i,j} = P(X_1^2 = (i,j))$ and represent it as a row vector $\vec{p}$. If we define matrix $\mathbb{D} \triangleq (q_{u_1^2,v_1^2})$ of size $N^2 \times N^2$ as

$$d_{u_1^2,v_1^2} = \begin{cases} c_{u_1,v_1} & \text{if } u_1 \neq v_1 \text{ and } u_2 = v_2 \\ c_{u_2,v_2} & \text{if } u_1 = v_1 \text{ and } u_2 \neq v_2 \\ 0 & \text{otherwise,} \end{cases} \tag{5}$$

and diagonal matrix $\mathbb{G} \triangleq (g_{u_1^2,v_1^2})$ of size $N^2 \times N^2$ as $g_{u_1^2,u_1^2} = -c_{u_1,u_1} - c_{u_2,u_2}$, then equation (4) can be written in a compact form as $\vec{p}\mathbb{D} = \vec{p}\mathbb{G}$. Both matrices $\mathbb{D}$ and $\mathbb{G}$ are non-negative by their definitions.

Let $\mathbb{H} = \mathbb{I} + \gamma(\mathbb{D} - \mathbb{G})$. If we put $\gamma = (\max_{u_1^2 \in \mathcal{X}^2} g_{u_1^2,u_1^2})^{-1}$, then matrix $\mathbb{H}$ is stochastic and $\vec{p}\mathbb{H} = \vec{p}$ iff $\vec{p}\mathbb{D} = \vec{p}\mathbb{G}$ and thus (2) is equivalent with an eigenvalue problem for matrix $\mathbb{H}$.

First, we observe that for $i \neq j$ $c_{ij} > 0$ iff $h_{(i,a),(j,a)} > 0$ for all $a \in \mathcal{X}$, because by (5) $h_{(i,a),(j,a)} = \gamma d_{(i,a),(j,a)} = \gamma c_{ij}$ (the first case when $u_2 = v_2$). Similarly, for $i \neq j$ $c_{ij} > 0$ iff $h_{(a,i),(a,j)} > 0$ for all $a \in \mathcal{X}$ (the second case when $u_1 = v_1$). This means that $i \to j$ iff $(i,a) \to (j,a)$ w.r.t. $\mathbb{H}$ for all $a \in \mathcal{X}$ and similarly $i \to j$ iff $(a,i) \to (a,j)$ w.r.t. $\mathbb{H}$ for all $a \in \mathcal{X}$. This can be proved by using the previous statement. By this rule used for a given $u_1^2 \in \mathcal{E}_a \times \mathcal{E}_b$, we obtain $u_1^2 \to v_1^2$ and $v_1^2 \to u_1^2$ for all $v_1^2 \in \mathcal{E}_a \times \mathcal{E}_b$ and thus $\mathcal{E}_a \times \mathcal{E}_b$ is an ergodic class w.r.t. $\mathbb{H}$. We show that there can not be more ergodic classes and thus we have all k^2 of them. If $u_1^2 \in \mathcal{F} \times \mathcal{E}$, then u_1^2 has to be transient w.r.t. $\mathbb{H}$, otherwise we will obtain contradiction with $u_1 \in \mathcal{F}$. This is because the only consequents of order 1 are of type $(i,a) \to (j,a)$ or $(a,i) \to (a,j)$, therefore if $u_1^2 \in \mathcal{F} \times \mathcal{E}$, we choose $v_1^2 \in \mathcal{X} \times \mathcal{E}$, such that $v_1 \not\to u_1$ (u_1 is transient and thus such v_1 must exist). State u_1^2 must be transient otherwise $u_1^2 \leftrightarrow v_1^2$ implies $u_1 \leftrightarrow v_1$ which results in contradiction with $v_1 \not\to u_1$. Similarly for $u_1^2 \in \mathcal{E} \times \mathcal{F} \cup \mathcal{F} \times \mathcal{F}$.

This proof can be generalized for $n \geq 3$ by assuming larger matrices $\mathbb{P}$, $\mathbb{D}$, $\mathbb{G}$, and $\mathbb{H}$, obtaining exactly k^n linearly independent perfectly secure distributions satisfying the FI condition. ■

Next, we discuss the structure of the set of invariant distributions for a given MI embedding and show how to find ergodic classes from matrix $\mathbb{B}$ in practice. By Theorem 2.1 from [4, Chapter V, page 175], this can be done by inspecting the matrix limit $\mathbb{M} = (m_{i,j}) = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \mathbb{B}^i$. According to this theorem, state i

is non-transient iff $m_{i,i} > 0$ and is transient otherwise. We put two non-transient states $i, j \in \mathcal{X}$ into one ergodic class if $m_{i,j} > 0$. All rows of the matrix $\mathbb{M}$ corresponding to states in one ergodic class $\mathcal{E}_a$ are the same and equal to the invariant distribution of this class, $\pi^{(a)}$.

This section is closed with a short discussion of two practical embedding algorithms. For the F5 embedding algorithm [5], the set of states $\mathcal{X} = \{-1024, \ldots, 1024\}$. By the nature of the embedding changes (flip towards 0), there is only one ergodic set $\mathcal{E}_1 = \{0\}$ and $\mathcal{F} = \mathcal{X} \setminus \{0\}$. Thus, there is only one invariant distribution, $\pi_0 = 1$ and zero otherwise. Obviously, no message can be embedded in covers with this singular distribution.

For the case of LSB embedding over $\mathcal{X} = \{0, \ldots, 255\}$, we have $\mathcal{E}_a = \{2a, 2a+1\}$ for $a \in \{0, \ldots, 127\}$, $\mathcal{F} = \emptyset$ and $\pi_{2a}^{(a)} = \pi_{2a+1}^{(a)} = \frac{1}{2}$ and zero otherwise (LSB embedding cannot be detected in images with evened out histogram bins). Thus, sources realized as a sequence of mutually independent random variables with such a distribution are the only perfectly secure sources w.r.t. LSB embedding. Figure 1 shows examples of matrices $\mathbb{B}$ and ergodic classes of several known algorithms with MI embedding operation.

5. APPLICATION TO MARKOV COVER SOURCES

In this section, we reformulate the results obtained so far for a special type of cover sources that can be modeled as first-order stationary Markov Chains (MC). The results play a key role in proving the square root law of steganographic capacity of imperfect stego-systems for Markov covers [1, 6].

First, for stationary cover sources Theorem 3 leads to this immediate corollary.

Corollary 6 *There are exactly k (instead of k^n) linearly indpendent perfectly secure stationary cover sources. These sources are i.i.d. with some invariant distribution $\pi_a, a \in 1, \ldots, k$.*

The next corollary states that in order to study perfect security of n-element stationary MC covers, it is enough to study only 2-element covers.

Corollary 7 *Let P, Q_β be first-order stationary MC cover distribution and its corresponding stego distribution after MI embedding with parameter β. For a given $n \geq 2$, an n-element stego-system is perfectly secure iff the corresponding stego-system narrowed to 2-element cover source is perfectly secure for some $\beta_0 > 0$:*

$$\exists \beta_0 > 0,\ \forall y_1^2 \in \mathcal{X}^2 \quad P^{(2)}(X_1^2 = y_1^2) = Q_{\beta_0}^{(2)}(X_1^2 = y_1^2). \tag{6}$$

Moreover, the FI condition for Markov sources simplifies to

$$\forall y_1^2 \in \mathcal{X}^2 \quad \Big(P^{(2)}(X_1^2 = y_1^2) > 0\Big) \Rightarrow \Big(\frac{d}{d\beta} Q_\beta^{(2)}(y_1^2)\big|_{\beta=0} = 0\Big). \tag{7}$$

Proof Because invariant distributions do not depend on β, Equation (6) must be valid for all $\beta > 0$ once it holds for some β_0 (see the arguments at the end of Sec. 2). By Corollary 6, if the stego-system is perfectly secure ($n \geq 2$), then the cover source is i.i.d. with some invariant distribution w.r.t. MI embedding and thus (6) and (7) hold. On the other hand, if (6) and (7) hold for $n = 2$ and stationary cover source, then this cover source is i.i.d. with one of k invariant distributions. This completes the proof since 2-element marginal is sufficient statistics for a first-order stationary MC. ∎

6. CONCLUSION

Most practical stego-systems for digital media embed messages by making independent changes to individual cover elements. In this paper, we fix the embedding operation and then inquire in which cover sources the embedding is statistically undetectable in Cachin's sense. The main contribution of this paper is a complete geometric characterization of such sources. Using the theory of ergodic classes, we show that all cover sources that are perfectly secure with respect to mutually independent embedding form a vector space spanned by invariant distributions determined by the embedding operation.

Additionally, we showed that perfect security of stegosystems with mutually independent embedding is completely captured using Fisher information formulated in Section 4 as the FI condition. This result not only provides a simpler and equivalent condition for perfect security, but it finds further applications in steganalysis. For example, Fisher information could be used for benchmarking such stego-systems, a direction we intend to pursue in our future research. Moreover, Fisher information provides fundamental lower bounds on the variance of unbiased estimators of the change rate, which connects our results to problems in quantitative steganalysis. Finally, the FI condition plays a key role in proving the square root law of steganographic capacity of imperfect stego-systems [1, 6].

7. REFERENCES

[1] T. Filler, J. Fridrich, and A. D. Ker, "The square root law of steganographic capacity for Markov covers," in *Proceedings SPIE, Electronic Imaging, Security and Forensics of Multimedia XI*, E. J. Delp, P. W. Wong, N. Memon, and J. Dittmann, Eds., San Jose, CA, January 18–21, 2009.

[2] C. Cachin, "An information-theoretic model for steganography," in *Information Hiding, 2nd International Workshop*, D. Aucsmith, Ed., Portland, OR, April 14–17, 1998, vol. 1525 of *Lecture Notes in Computer Science*, pp. 306–318, Springer-Verlag, New York.

[3] J. Kodovský, J. Fridrich, and T. Pevný, "Statistically undetectable JPEG steganography: Dead ends, challenges, and opportunities," in *Proceedings of the 9th ACM Multimedia & Security Workshop*, J. Dittmann and J. Fridrich, Eds., Dallas, TX, September 20–21, 2007, pp. 3–14.

[4] J. L. Doob, *Stochastic processes*, Wiley, New York, 1st edition, 1953.

[5] A. Westfeld, "High capacity despite better steganalysis (F5 – a steganographic algorithm)," in *Information Hiding, 4th International Workshop*, I. S. Moskowitz, Ed., Pittsburgh, PA, April 25–27, 2001, vol. 2137 of *Lecture Notes in Computer Science*, pp. 289–302, Springer-Verlag, New York.

[6] A. D. Ker, T. Pevný, J. Kodovský, and J. Fridrich, "The square root law of steganographic capacity," in *Proceedings of the 10th ACM Multimedia & Security Workshop*, A. D. Ker, J. Dittmann, and J. Fridrich, Eds., Oxford, UK, September 22–23, 2008, pp. 107–116.