Analysis and Improving the Security of the Scalar Costa Scheme against Known Message Attack

R. Samadi* (C.A.) and S. A. R. Seyedin*

Abstract: The popular watermarking method with structured codebook is the Scalar Costa Scheme (SCS) which is highly robust but poorly secure. Known Message Attack (KMA) is a type of the security attacks which its countermeasure is more difficult than others. This paper proposes a novel scheme to increase security of SCS in KMA scenario. For this purpose, the SCS model is extended to a more general model. The security of Generalized SCS (GSCS) is analyzed using residual entropy as a security measure. Then, fundamental trade-off between security and achievable rate of GSCS is proved in KMA scenario. Based on this trade-off and the practical security attack on SCS, a new extension of SCS is proposed, called Surjective-SCS (SSCS). In comparison with SCS, SSCS clearly achieves more security and achievable rate in low Watermark to Noise Ratio (WNR) regime.

Keywords: Achievable Rate, Flat-host Assumption, Known Message Attack, Scalar Costa Scheme, Watermarking.

1 Introduction

The rapid development of digital information technology has made the multimedia distribution over public networks easy and popular. However, this public distribution is vulnerable to some serious threats such as unlimited duplication, illegally redistribution, etc. To prevent these illegal operations, the steganography [1], and watermarking [3] are used to hide a verification message into the multimedia. Two well-known watermarking schemes are Spread Spectrum (SS) and quantization-based codes. Scalar quantization-based data hiding [4] (also known as Scalar Costa Scheme (SCS) [5]) is the most popular watermarking scheme which is used in practice for its simplicity and good performance. Quantization-based data hiding schemes offer larger achievable rate versus SS schemes [6], however they are worse in term of security [7-8]. Hence, there is a demand for secure quantization-based data hiding schemes.

The security of a symmetric watermarking scheme (like SCS) may be enhanced by secure coding [9-11], host-dependent keys [12-13], host-dependent codebook (randomized codebook) [14-16] or secure embedding [17-21]. Although, secure coding is applicable to any watermarking scheme, it doesn't prevent the unauthorized removal attack [9]. Host-dependent keying suffers from key synchronization at the decoder side [13]. Randomizing codebook increases the entropy (security) of the codebook by means of non-structured codebooks, but it also increases computational complexity [15].

Secure embedding keeps the structure of the lattice codebook and increases the security by changing the law of embedding. One secure embedding method is to make a secret rotation of the embedding lattice [17]. However it is easy to model secret rotation as secure coding followed lattice embedding. Other method is to use a Look-Up Table (LUT) after SCS [18]. This method achieves less probability of error but more embedding distortion. Another method for security improvement is the use of Spread Transform Dither Modulation (STDM) in conjunction with LUT [19]. However, both spread transform and LUT techniques are not applicable in the low WNR. More recent works for increasing the security of SCS [20-21] are based on defining the decoding region with the aid of distribution matching. This method is perfectly secure only against Watermark Only Attack (WOA) which the adversary has a pool of hosts watermarked with the same secret key.

Sometimes the adversary has access to corresponding messages in addition to watermarked hosts. This scenario is called Known Message Attack (KMA). The problem of secure quantization-based embedding scheme in KMA scenario has been not

Iranian Journal of Electrical & Electronic Engineering, 2014.
Paper first received 12 June 2013 and in revised form 20 Oct. 2013.
* The Authors are with the Department of Electrical Engineering, Ferdowsi University of Mashhad, Mashhad, Iran.
E-mails: r.samadi@stu.um.ac.ir and seyedin@um.ac.ir.

addressed before in the literature which is the main scope of this paper. For this purpose, the SCS model is analyzed in general case. In this case to have an accurate result, the flat-host approximation is not applied. Moreover watermark is assumed to be an arbitrary function of quantization noise without transgressing orthogonality as in the Costa's construction. To analyze Generalized SCS (GSCS), the residual entropy is used as a security measure. The analysis shows a trade-off between security and achievable rate of the GSCS in KMA scenario. Based on this trade-off and the practical security attack on SCS, a new extension of SCS is proposed which is called Surjective-SCS (SSCS). The SSCS achieves more security and achievable rate than SCS in KMA scenario and in low Watermark to Noise Ratio (WNR) regime while it keeps the transparency and computational cost. The application of SSCS in multimedia is for the situations where attacker has large number of observations and wants to implement security attack; while simultaneously adding strong noise to decrease embedding bit rate.

Formal definitions of digital watermarking characteristics are reviewed in section 2. Section 3 provides the accurate security analysis of GSCS in KMA scenario, and proves the fundamental trade-off between security and achievable rate. As a special case, the security analysis of SCS in KMA scenario is examined. New scheme is proposed in section 4 with the proof of its superiority over SCS in term of security and achievable rate. Finally section 5 concludes the paper and gives some remarks and future research lines.

2 Problem Formulation

The theoretical model of digital watermarking followed in this paper is shown in Fig. 1. This model is the same as that of used in [4]. Coefficients x_i are i.i.d sequences of scalar features which are extracted from original digital content by Discrete Cosine Transform, Discrete Wavelet Transform, Fast Fourier Transform or other spatial/temporal transformations. The encoder hides an equi-probable watermark message $m_i \in \{0,...,p-1\}$ in x_i using secret key k_i to yield a watermark w_i. Then the watermark w_i is added to the host x_i yielding watermarked host y_i. The watermarked host y_i undergoes channel attack and is added by AWGN noise n_i. The detector that receives noisy watermarked host z_i should estimate embedded message m_i using secret key k_i.

The embedding distortion (or watermark power) is computed as $D_w = E\{|w|^2\}$. The transparency is measured by host variance to watermark power ratio $\lambda = \sigma_x^2 / D_w$. The attack channel is parameterized by watermark power to noise variance ratio $\zeta = D_w / \sigma_n^2$. DWR and WNR can be defined as $DWR = 10\log_{10}(\lambda)$ and $WNR = 10\log_{10}(\zeta)$, respectively.

The achievable rate can be computed by maximizing mutual information over embedding function $f(.)$, while λ and ζ are fixed.

$$R(\lambda,\zeta) = \max_{f(.)} I(Z;M \mid K) \tag{1}$$

In security evaluation, the purpose of attacker is disclosing the secret parameters and then implementation of the tampering attack. According to Kerckhoffs' principle, all details of the watermarking technique except the so-called secret key parameter of the embedding and decoding processes are publicly known. For evaluating security we use security level definition from [22]. The security level of a secrecy system is said to be the effort that attacker requires for estimating the secret key. In this paper, we only concentrate on KMA scenario which the attacker has access to the pool of independent messages and corresponding watermarked host when hosts are watermarked with the same secret key. By using residual entropy as information theoretic measure of security level [23], the γ-security level [7] is defined as the number of observation N_0 that attacker needs to holds inequality in Eq. (2). Using entropy as security measure is common in watermarking literature [24]. The security level of SCS in KMA scenario by using flat-host assumption, is derived theoretically in [7] only for $\alpha_{SCS} \geq 0.5$ where α_{SCS} is distortion compensation parameter. In [7] authors show that the trade-off between security and achievable rate of SCS is controlled by α_{SCS}.

$$h(K \mid Y^{N_0}, M^{N_0}) \leq \gamma \tag{2}$$

In this paper, we propose a security and achievable rate analysis of Generalized SCS after removing restrictive approximations like flat-host assumption, limited values of distortion compensation parameter and specific embedding law. Then, we propose a new scheme to achieve more security and achievable rate than SCS in the fixed transparency and computational complexity. An easy way to increase security is to decrease α_{SCS}. Although this choice decreases achievable rate as it is expected; however as we will prove in the next section, security doesn't increase so much for small α_{SCS}; besides, it is still large gap up to

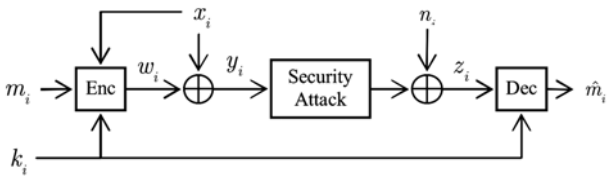


Fig. 1 Theoretical model for additive side-informed watermarking, including security attack and noisy channel.

security level of the SS scheme. In recent work [21], author chooses an extension of SCS by changing the embedding law which achieves perfect secrecy in WOA scenario, while obtains more achievable rate than SCS. This result motivates us to analyze the relation between security and achievable rate of Generalized SCS in KMA scenario, and then design embedding law with more security and achievable rate.

3 Security Analysis of General SCS

SCS can be extended into GSCS without transgressing orthogonality as in Costa's construction [25] and its implementation SCS [5]. In SCS, watermark is a linear function of quantization error. However in GSCS, watermark can be an arbitrary function of quantization error. It should be noted that in GSCS, watermark is still nearly orthogonal to x and host rejection is still possible. In GSCS, watermark can be written as

$$w = G(e_{m,k}(x)) = T(e_{m,k}(x)) - e_{m,k}(x) \tag{3}$$

where

$$\begin{aligned} e_{m,k}(x) &= x - \Delta\frac{m}{p} - k - Q_{\Delta}(x - \Delta\frac{m}{p} - k) \\ &\triangleq x - Q_{m,k}(x) \end{aligned} \tag{4}$$

and Q_{Δ} is uniform scalar quantizer over period $[-\Delta/2, \Delta/2]$, Δ is quantization step size and m is to-be-transmitted message symbol. $Q_{m,k}$ is shifted uniform scalar quantizer that its centroids is defined by a shifted lattice:

$$\Lambda_{m,k} \triangleq \Delta\mathbb{Z} + \Delta\frac{m}{p} + k \tag{5}$$

The variable k is called dither and it is used to reduce visible artifacts. However, in SCS k plays the role of secret key. The security of the embedder relies only on the randomization of the codebook via a dithering process. Previous works on the performance analysis of SCS assume that secret key k is statistically distributed uniformly over period $[-\Delta/2, \Delta/2]$, therefore the Schuchman condition is satisfied [26] and error signal $e_{m,k}(x)$ is nearly orthogonal to x, m. Hereafter, we use uniform pdf for k derive theoretical results.

Two variants of GSCS (other than SCS) have been used in the literature before. In [27], authors define $G(.)$ as a transform function from uniform distribution to Gaussian one, then they derive simple theoretical expression for probability of error and show that their scheme achieve lower error rate than SCS for a large WNR range. They call it Gaussian DC-DM (GDC-DM). In recent work [21], the author predefines pdf of host and watermarked host using flat host assumption; then he obtains function $T(.)$ using optimal distribution matching in such a way that SCS becomes secure in WOA scenario; the author called proposed scheme Soft-SCS. In both scheme, function $T(.)$ is nonlinear which it increases computational complexity.

Power of watermark in GSCS can be computed using crypto lemma [28] by Forney as bellow:

$$D_w = \frac{1}{\Delta}\int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} (T(e) - e)^2 de \tag{6}$$

For the sake of simplicity in deriving pdf of Watermarked host, we assume that $T(.)$ is odd function as in [5], [27] and [21]. Also without loss of generality, we assume that $|T(e)| \leq |e|$ holds, because if it doesn't hold, then we can translate the codebook (both host and watermarked host) by $\Delta/2$, so assumption will be hold for new embedder $T_1(e)$ as follows [21]:

$$T_1(e) = \Delta/2 - T(\Delta/2 - e) \tag{7}$$

For computing the pdf of watermarked host, we use the same approach as used in [6] by taking into account that SCS can be considered as a random variable transformation. Form Eq. (3) we have equality

$$Q_{m,k}(y) = Q_{m,k}(x) + Q_{m,k}(T(e_{m,k}(x))) = Q_{m,k}(x) \tag{8}$$

Therefore inverse GSCS would be as:

$$x = T^{-1}(y - Q_{m,k}(y)) + Q_{m,k}(y) \tag{9}$$

Now consider the following pdf transformation [29] resulting from Eq. (3).

$$p_Y(y \mid M = m, K = k) = \frac{p_X(x)}{T'(x - Q_{m,k}(x))} \tag{10}$$

Substituting Eq. (8) and Eq. (9) into Eq. (10) we have the following pdf of watermarked host which is used frequently during this paper.

$$\begin{aligned} &p_Y(y \mid M = m, K = k) \\ &\quad = \frac{p_X(T^{-1}(y - Q_{m,k}(y)) + Q_{m,k}(y))}{T'(T^{-1}(y - Q_{m,k}(y)))} \end{aligned} \tag{11}$$

Now we are ready to analyze the security of GSCS w.r.t definition in Eq. (2). First we analyze security for $N_0 = 1$ observation which gives a simple closed form. Theoretical security analysis for $N_0 = 1$ observation helps us to compare it with previous theoretical results.

3.1 Residual Entropy of $N_0 = 1$ Observation

Residual entropy of secret key conditioned on watermarked host and to-be-transmitted message can be written as follows:

$$\begin{aligned} h(K \mid Y, M) &= h(K) - I(K; Y, M) \\ &= h(K) - I(K; Y \mid M) \\ &= h(K) - h(Y \mid M) + h(Y \mid M, K) \end{aligned} \tag{12}$$

The first term is equal to $\log_2(\Delta)$. In order to compute the second term $h(Y \mid M)$, we need to know the pdf of watermarked host conditioned on message which can be computed via Eq. (11) as follows. The proof is the same as one used in [26] in addition to some simplifications.

$$p_Y(y \mid M = m) = \frac{1}{\Delta}\int_{-\frac{\Delta}{2}}^{+\frac{\Delta}{2}} p_X(y + r - T(r))dr \tag{13}$$

Obviously this conditioned pdf is independent of message $p_Y(y \mid M = m) = p_Y(y)$, i.e. in GSCS, watermarked host don't leak any information about to-be-transmitted message to attacker. Also, this form corresponds to previous exact result in [30], so it testifies Eq. (11).

The second term $h(Y \mid M, K)$ can be derived via Eq. (11) as follows. The proof is in Appendix A.

$$h(Y \mid M, K) = \frac{1}{\Delta}\int_{-\frac{\Delta}{2}}^{+\frac{\Delta}{2}} \log(T'(e))de + h(X) \tag{14}$$

So residual entropy of secret key conditioned on watermarked host and to-be-transmitted message can be written as

$$h(K \mid Y, M) = \log(\Delta) + h(X) - h(Y) + \frac{1}{\Delta}\int_{-\frac{\Delta}{2}}^{+\frac{\Delta}{2}} \log(T'(e))de \tag{15}$$

The term $h(X) - h(Y)$ for Gaussian host, can be bounded. Because x and w are independent (Schuchman condition) and $y = x + w$, by using Power Entropy Inequality (PEI) lemma [31] and some simplifications, we have:

$$h(X) - h(Y) \le -0.5\log_2(1 + \frac{2^{2h(W)}}{2\pi e\sigma_x^2}) \tag{16}$$

Also by using maximum entropy lemma [31] we have:

$$h(X) - h(Y) \ge -0.5\log_2(1 + \frac{D_w}{\sigma_x^2}) \tag{17}$$

Thus the term $h(X) - h(Y)$ is negative and dual bounded, also lower and upper bound can be converged if distribution of watermark tends to Gaussian shape (using maximum entropy lemma [31]). The third term in Eq. (15) is also negative, because $|T(e)| \le |e|$. From Eq. (15) and Eq. (17) we have:

$$\lim_{\substack{T(e)\to e,\\ DWR\to+\infty}} I(K;Y,M) = \lim_{DWR\to+\infty} h(Y) - h(X) = 0 \tag{18}$$

So, perfect secrecy for GSCS in KMA scenario is possible if $T(e) \to e$ and $DWR \to +\infty$. However this choice leads to zero embedding distortion (from Eq. (6)) and consequently zero achievable rate, since achievable rate is an increasing function of embedding distortion. This shows trade-off between security and achievable rate of GSCS in KMA scenario (at least for one observation).

3.2 Residual Entropy of $N_0 \ge 1$ Observations

Residual entropy of secret key conditioned on watermarked hosts and messages can be evaluated as:

$$\begin{aligned} h(K \mid Y^{N_0}, M^{N_0}) &= h(K) - I(K; Y^{N_0}, M^{N_0}) \\ &= h(K) - I(K; Y^{N_0} \mid M^{N_0}) \\ &= h(K) - h(Y^{N_0} \mid M^{N_0}) \\ &\quad + h(Y^{N_0} \mid M^{N_0}, K) \end{aligned} \tag{19}$$

It should be noted that Y^{N_0} conditioned on M^{N_0} and K, are composed of i.i.d sequence X_i. Using Eq. (14) we can write:

$$h(Y^{N_0} \mid M^{N_0}, K) = N_0 \frac{1}{\Delta}\int_{-\frac{\Delta}{2}}^{+\frac{\Delta}{2}} \log(T'(e))de + N_0 h(X) \tag{20}$$

For second term, we should first compute the pdf of watermarked hosts conditioned on messages. It can be written as follows:

$$\begin{aligned} p(y^{N_0} \mid m^{N_0}) &= \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} p(y^{N_0} \mid m^{N_0}, k)\, p_K(k)dk \\ &= \frac{1}{\Delta}\int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \prod_{i=1}^{N_0} \frac{p_X(T^{-1}(y_i - Q_{m,k}(y_i)) + Q_{m,k}(y_i))}{T'(T^{-1}(y_i - Q_{m,k}(y_i)))}dk \end{aligned} \tag{21}$$

Unfortunately, deriving theoretical closed form for this pdf is intricate, so we use numerical integration for deriving this pdf. The first solution is to compute Eq. (21) numerically, then replacing it in Eq. (19) and integrating over $\{y^{N_0}, m^{N_0}\}$ using Monte-Carlo integration. This solution is not accurate for large N_0 due to big error in large dimension in Monte-Carlo integration. The more exact solution is to average $h(K \mid y^{N_0}, m^{N_0})$ over large number N of outcomes $\{y^{N_0}, m^{N_0}\}_{i=1}^{N}$ instead of integration over them as

$$\begin{aligned} h(K \mid Y^{N_0}, M^{N_0}) &= E_{Y^{N_0}, M^{N_0}}\{h(K \mid y^{N_0}, m^{N_0})\} \\ &\cong \frac{1}{N}\sum_{i=1}^{N} h(K \mid \{y^{N_0}, m^{N_0}\}_i) \end{aligned} \tag{22}$$

which is application of *weak law of large numbers* in approximation used above. As N increases, variance of error tends to zero if outcomes $\{y^{N_0}, m^{N_0}\}_i$ are independent. This is true because $\{y^{N_0}\}_i$ in iteration i are function of mutually independent variables $\{x^{N_0}\}_i$, $\{m^{N_0}\}_i$ and k_i. In practice $N \approx 500$ is sufficient to get an accurate result.

3.3 Investigation on SCS as a Special Case

GSCS can be easily simplified to SCS by substituting $T_{SCS}(e) = (1-\alpha_{SCS})e$ in GSCS. A good approximation for residual entropy conditioned on watermarked host and message for SCS Gaussian host in $N_0 = 1$ observation is as follows:

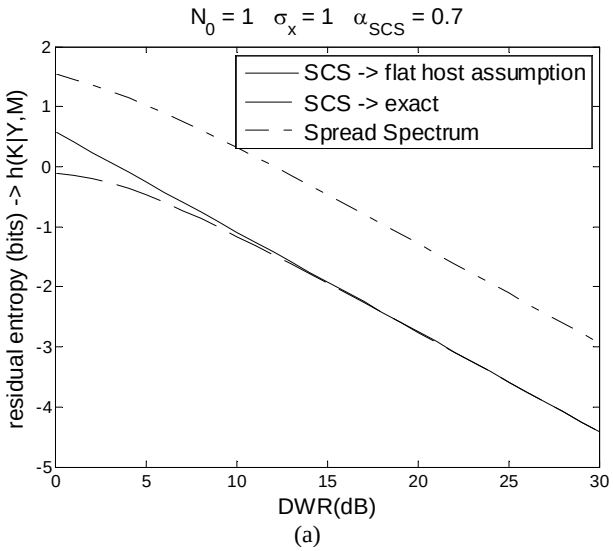
$$h(K \mid Y, M) \cong \log_2(\Delta_{SCS}) + \log_2(1-\alpha_{SCS}) - \frac{0.684}{\lambda} \qquad (23)$$

The proof comes after using Eq. (6), Eq. (13) and Eq. (15) and some simplifications. It is worthy to note that the previous security analysis of SCS [7] ignores host statistics and do not drive the third term in Eq. (23). It is because of using flat-host assumption which implies infinite host variance, so substituting $\lambda = +\infty$ in Eq. (23) gets previous result in [7]. Comparison between exact security analysis of SCS proposed here and the security analysis based on flat-host assumption [7] is plotted in Fig. 2(a). For $N_0 \geq 1$ we use numerical method as in Eq. (22). Comparison between previous theoretic result based on flat-host assumption [7] and our result is sketched in Fig. 2(b) and Fig. 2(c). Also theoretic result for SS scheme [32] is sketched for comparison. Accurate result for $N_0 = 1$ and numerical result for $N_0 \geq 1$ shows large difference between the security level of SCS derived here and the security level of SCS by flat-host assumption [7]. For example as depicted in Fig. 2(b), the previous result in [7] shows that we need N_0 ; 500 observations to make residual entropy lower than -10.1, but we show here that attacker only needs N_0 ; 70 observations to do this. Other key point is that even for low α_{SCS} , there is still large gap between security level of SCS and SS scheme.

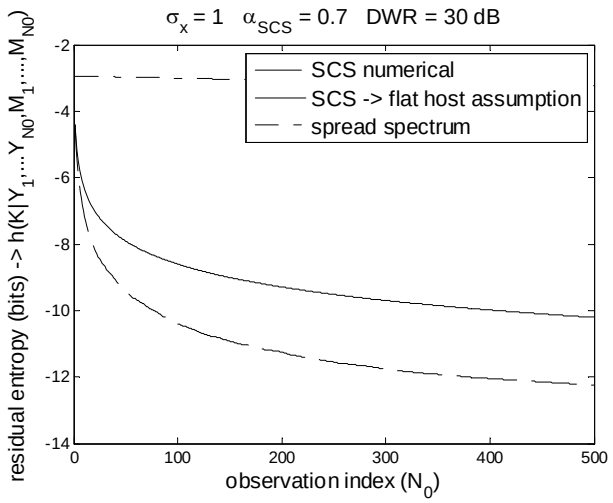
4 Surjective SCS

Security analysis made in section 3, clearly shows that there is a fundamental trade-off between security and achievable rate of GSCS in KMA scenario. So the perfect secrecy as defined in [22] and [33] or similar to one developed in [21] for improving security of SCS in WOA scenario, is not possible. As a result, we can only improve *security level* as defined in Eq. (2) in KMA scenario.

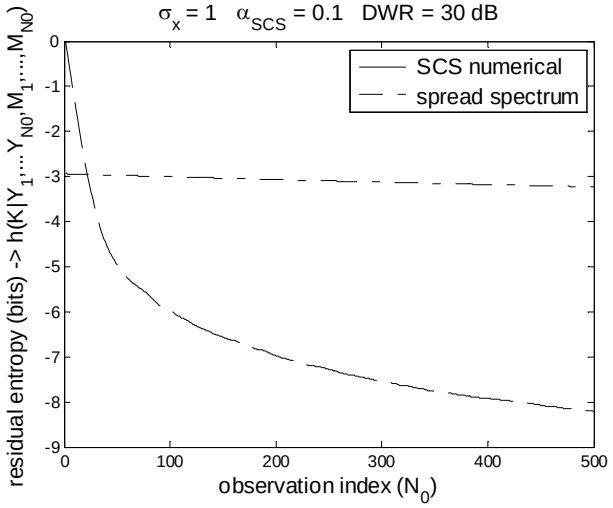
Security level of GSCS is completely dependent on the shape of function $T(.)$. In previous estimation attack on SCS [7], author benefits from the weakness that distribution of watermarked host doesn't cover the quantization cell for fixed message and secret key. It is clear by substituting $T_{SCS}(e) = (1-\alpha_{SCS})e$ in Eq. (11). To make connection with GSCS, the weakness in SCS comes arise from this fact that function $T(.)$ in SCS is not surjective over its domain, i.e. for some η there is no e such that $T(e) = \eta$. To overcome this weakness, we propose new scheme which called Surjective-SCS (SSCS) as follows.



(a)



(b)



(c)

Fig. 2 Security analysis of SCS for Gaussian host and $N_0 \geq 1$ observation.

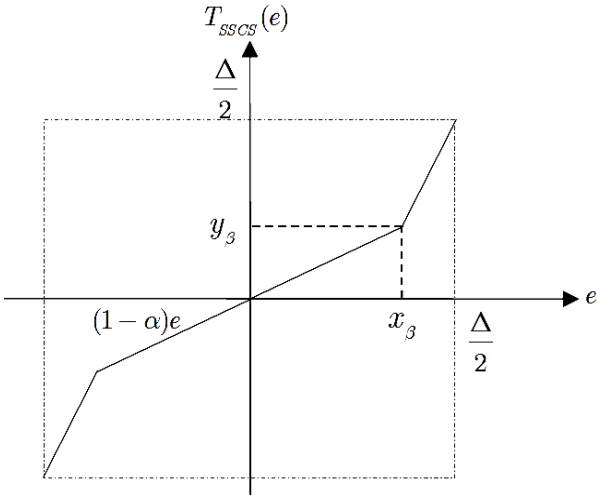


Fig. 3 Surjective-SCS.

The value for α and β is computed through numerical optimization.

$$T_{SSCS}(e)=\begin{cases}\dfrac{e+\beta\frac{\Delta}{2}}{1-\beta}, & -\dfrac{\Delta}{2}\le e\le -x_\beta\\ (1-\alpha)e, & |e|\le x_\beta=\dfrac{\Delta}{2}\dfrac{\beta}{\alpha+\beta-\alpha\beta}\\ \dfrac{e-\beta\frac{\Delta}{2}}{1-\beta}, & x_\beta\le e\le\dfrac{\Delta}{2}\end{cases} \quad (24)$$

Function $T_{SSCS}(.)$ is illustrated in Fig. 3. Notice that SSCS with $\beta=1$ is equal to SCS, also SSCS with $\beta\neq 1$ by definition is completely secure against estimation attack developed in [7]. But as stated in [7], attacker may use other non-convex estimation attack. To analyze security of SSCS against every estimation attack, we use theoretic security analysis of GSCS made in previous section and show that, SSCS is always more secure than SCS in the same DWR and WNR.

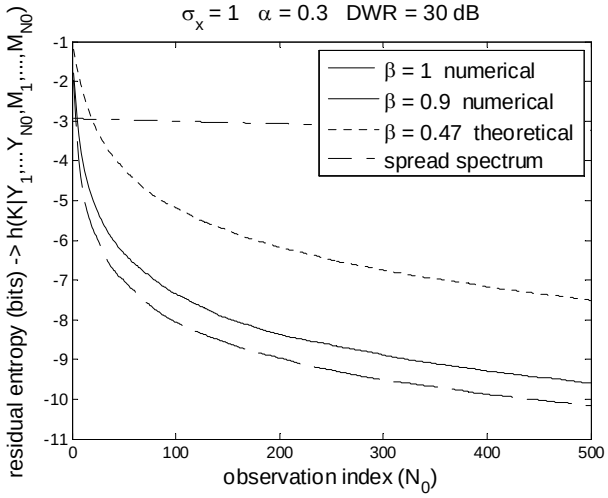
For $N_0=1$ observation, Eq. (15) simplifies to:

$$h(K\mid Y,M)=\log(\Delta)-\frac{0.684}{\lambda}+\frac{\beta\log(1-\alpha)-\alpha(1-\beta)\log(1-\beta)}{\alpha+\beta-\alpha\beta} \quad (25)$$

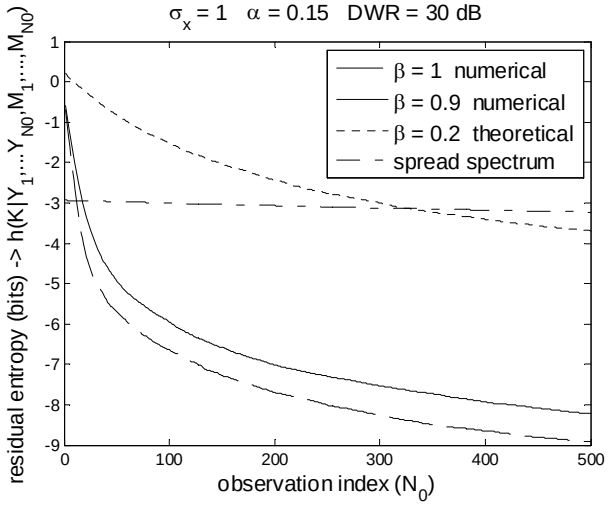
The proof is same as Eq. (23). Now we compare residual entropy of SCS and SSCS in Eq. (23) and Eq. (25). It is simple to find the region $(\alpha_{SCS},\alpha,\beta)$, which residual entropy of SCS is smaller than SSCS in same DWR. This region is derived from below which come after using Eq. (6) for SCS and SSCS and substituting it in Eq. (23) and Eq. (25).

$$\log\left(\frac{1-\alpha_{SCS}}{\alpha_{SCS}}\right)\le -\log\left(\frac{\alpha\beta}{\alpha+\beta-\alpha\beta}\right)+\frac{\beta\log(1-\alpha)-\alpha(1-\beta)\log(1-\beta)}{\alpha+\beta-\alpha\beta} \quad (26)$$

For $N_0\ge 1$, we use Eq. (22) to compare security level of SSCS and SCS. Comparison for some α,β,Δ is shown in Fig. 4. Simulation results state that, security level of SSCS is greater than SCS. As discussed in previous section, by reducing α, we can't achieve more security than SS, but here we can see that, for low α by reducing β, we can achieve more security even than SS. However reducing β, may lead to a lower achievable rate in the same DWR, so in following, we add AWGN attack channel and compare achievable rate of SSCS with SCS in the same DWR and WNR.



(a)



(b)

Fig. 4 Residual entropy of SSCS ($\beta=1$ is equal to SCS).

With $\beta \neq 1$, we introduce more symbol interference (self-noise) which happens in SCS only when $\alpha \leq 0.5$, but we show that this more symbol interference, increases achievable rate as in [34]. We compute achievable rate to find the inherent performance limits of SSCS. To investigate this, we compute achievable rate from Eq. (1) and use the same approach as used in [6] by ignoring the flat-host assumption. The results in this section are derived for Gaussian hosts and binary signaling. Comparison of achievable rate between SSCS and SCS are sketched in Fig. 5. Optimum encoder parameters which maximize achievable rate are in Fig. 6. For WNR larger than -2 dB, $\beta = 1$ (SCS) is optimum. SSCS obtains more achievable rate than SCS only for WNR smaller than -2 dB, so we compare results only for negative WNRs. It is worthy to note that, we compare maximum achievable rate of SSCS and SCS by fixing $k = -\Delta/4$ in both SCS and SSCS. As discussed in [6], for lattices similar to those of in Eq. (5), using different dither may incur a loss of performance and comparison of maximum achievable rate is meaningless.

Finally, we compare security level of SSCS and SCS after adding Gaussian attack, to make a connection with achievable rate. Residual entropy of SSCS and SCS are compared in Fig. 7. As discussed in the first of this section, we can conclude that security level of SSCS is greater than SCS for WNR smaller than -2 dB.

We can conclude this section with four statements, 1) both security and achievable rate of SSCS is more than SCS for WNR smaller than -2 dB; 2) security level of SSCS in comparison with SCS increases as WNR decreases, e.g. in WNR=-8 dB as illustrated in Fig. 4(a) and Fig. 6, security level of SSCS is much higher than SCS while keeps other criteria, but in WNR=-15 dB as illustrated in Fig. 4(b) and Fig. 6, security level of SSCS is very much higher than SCS or SS while keeps other criteria; 3) by reducing β, we can fill the gap between the security level of SCS and SS while have more achievable rate than both of them; and 4) However $\beta = 1$ is optimum for WNR larger than -2 dB from achievable rate point of view, still we can use another β and increase the security level of SCS if we accept loss of achievable rate.

5 Conclusion and Future Work

The subject of this paper is to analyze and design a secure watermarking Scheme based on SCS in KMA scenario. We analyzed security of the GSCS with the aid of residual entropy as a security measure. As a special case, we proved that previous analysis of SCS overestimates its security level. Also, we showed the fundamental trade-off between security and achievable rate of the GSCS which is controlled by embedding law parameters. In SCS, this trade-off is controlled by DWR and distortion compensation α.

Based on security analysis of the GSCS, we proposed SSCS. Similar to SCS, the SSCS is piecewise linear, hence it keeps simplicity and computational complexity of SCS which is desirable from implementation point of view. We derived optimum embedding parameters for different values of DWR and WNR that maximize residual entropy. Using these optimum parameters, we evaluated achievable rate and residual entropy for specified numbers of observations. Surprisingly, comparison of achievable rates in low WNR shows that SSCS is more robust than both SCS and SS scheme in AWGN channel. However comparison of residual entropies for low WNR shows that SSCS is more secure than SCS. To use SSCS in a practical secure watermarking system, key management methods should be utilized.

Future work is to design a decoder for SSCS and replace it by SCS in applications which need high security. Another future work is to extend SSCS to multidimensional case by extending lattice embedding in Eq. (3). It should be noted that, the function $T(.)$ in multidimensional case may be more complex because of dependency between dimensions.

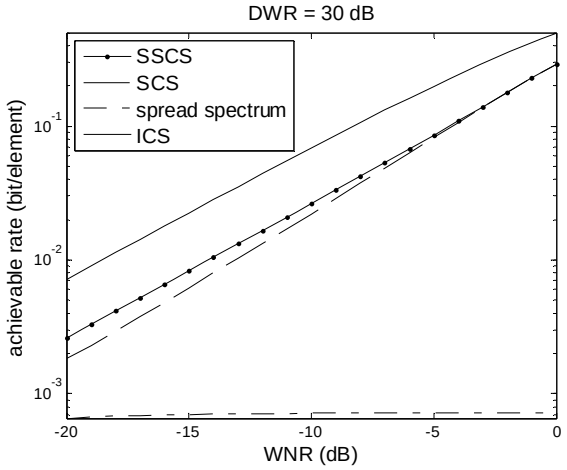


Fig. 5 Achievable rate of SSCS and SCS.

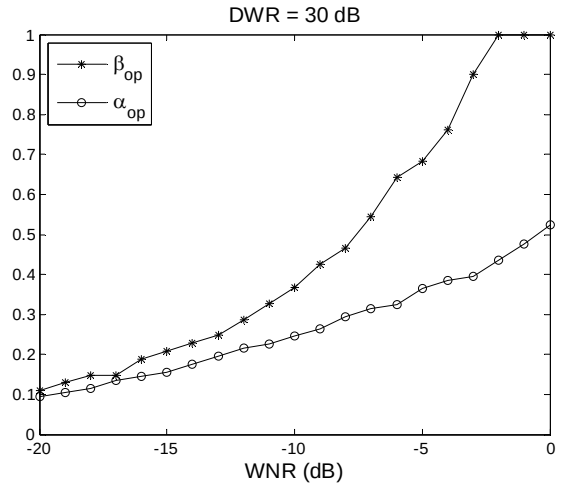
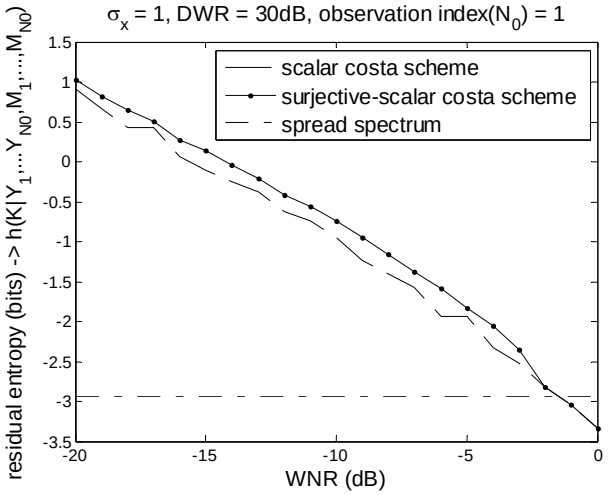
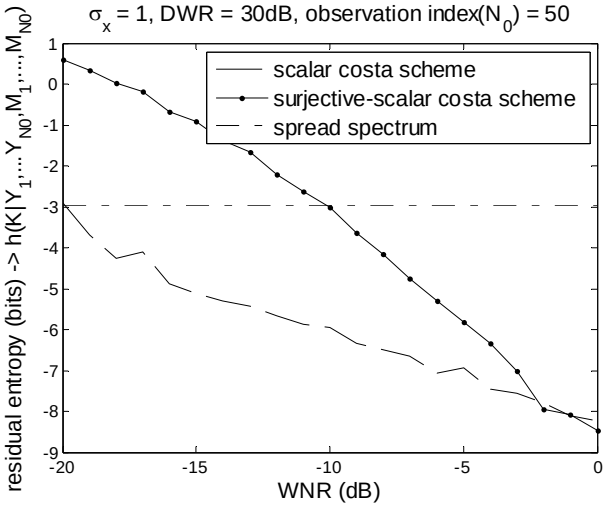


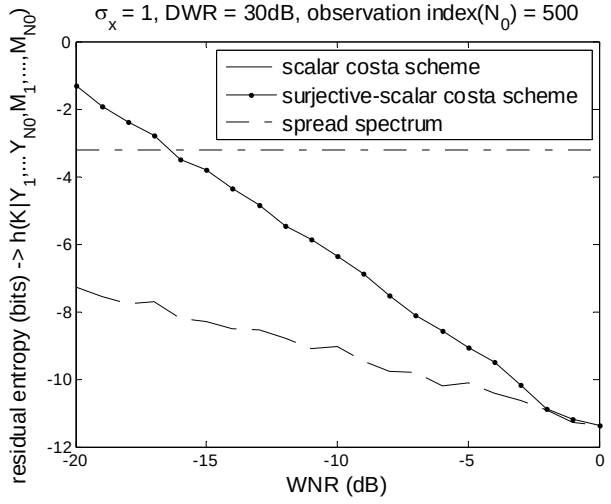
Fig. 6 Optimum encoder parameter in SSCS.



(a)



(b)



(c)

Fig. 7 Residual entropy comparison between SSCS and SCS.

Appendix

A. Entropy of Watermarked Host Conditioned on Message and Secret Key for GSCS

In this appendix to overcome the lack of space needed for long formula, we use notation $p_Y(y \mid m,k)$ instead of $p_Y(y \mid M=m, K=k)$ and $h(Y \mid m,K)$ instead of $h(Y \mid M=m,K)$. First we have:

$$h(Y \mid M,K) = \frac{1}{p}\sum_{m=0}^{p-1} h(Y \mid m,K) \tag{A1}$$

The term $h(Y \mid m,K)$ can be written as follows:

$$h(Y \mid m,K) = E_{Y,K|M=m}\{-\log_2(p_Y(y \mid m,k))\}$$
$$= -\int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} p_K(k)\int_{-\infty}^{+\infty} p_Y(y \mid m,k)\log_2(p_Y(y \mid m,k))dydk \tag{A2}$$

For uniform secret key k, we have:

$$h(Y \mid m,K) = \frac{1}{\Delta}\int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} h(Y \mid m,k)dk \tag{A3}$$

The term $h(Y \mid m,k)$ can be simplified as follows. The proof is the same as Eq. (13) or the one used in [26] plus some simplifications.

$$h(Y \mid m,k) = h(X) + \int_{-\infty}^{+\infty} p_X(e+\Delta\frac{m}{p}+k)\log_2(T'(e-Q_\Delta(e)))de \tag{A4}$$

After substituting Eq. (A4) in Eq. (A3) and then in Eq. (A1), we obtain the intended result in Eq. (14).

Acknowledgment

The authors are indebted to Prof. Gonzalez, Dr. Furon, Dr. Cayre, and Dr. Bas for their valuable helps.

References

[1] M. Mahdavi, Sh. Samavi, N. Zaker and M. Modarres-Hashemi, "Steganalysis Method for LSB Replacement Based on Local Gradient of Image Histogram", *Iranian Journal of Electrical & Electronic Engineering,* Vol. 4, No. 3, pp. 59-70, 2008.

[2] M. Soleimanpour-moghadam and S. Talebi, "A novel technique for steganography method based on improved genetic algorithm optimization in spatial domain", *Iranian Journal of Electrical & Electronic Engineering,* Vol. 9, No. 2, pp. 67-75, 2013.

[3] I. Cox, M. Miller and A. Mckellips, "Watermarking as communications with side information", *Proceedings of the IEEE,* Vol. 87, No. 7, pp. 1127-1141, 1999.

[4] B. Chen and G.W. Wornell, "Quantization index

modulation: a class of provably good methods for digital watermarking and information embedding", *Information Theory, IEEE Transactions on*, Vol. 47, No. 4, pp. 1423-1443, May 2001.

[5] J. J. Eggers, R. Bauml, R. Tzschoppe and B. Girod, "Scalar Costa scheme for information embedding", *Signal Processing, IEEE Transactions on* , Vol. 51, No. 4, pp. 1003-1019, Apr. 2003.

[6] L. Perez-Freire, F. Perez-Gonzalez and S. Voloshynovskiy, "An accurate analysis of scalar quantization-based data hiding", *Information Forensics and Security, IEEE Transactions on*, Vol. 1, No. 1, pp. 80-86, March 2006.

[7] L. Perez-Freire, F. Perez-Gonzalez, T. Furon and P. Comesana, "Security of Lattice-Based Data Hiding Against the Known Message Attack", *Information Forensics and Security, IEEE Transactions on*, Vol. 1, No. 4, pp. 421-439, Dec. 2006.

[8] L. Perez-Freire and F. Perez-Gonzalez, "Security of Lattice-Based Data Hiding Against the Watermarked-Only Attack", *Information Forensics and Security, IEEE Transactions on*, Vol. 3, No. 4, pp. 593-610, Dec. 2008.

[9] I. Cox, G. Doerr and T. Furon, "Watermarking is not cryptography", *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics),* Vol. 4283 LNCS, pp. 1-15, 2006.

[10] S. Mohammadi, and S. Talebi, and A. Hakimi, "Two Novel Chaos-Based Algorithms for Image and Video Watermarking", *Iranian Journal of Electrical & Electronic Engineering,* Vol. 8, No. 2, pp. 97-107, 2012.

[11] N. Liu, P. Amin and K. P. Subbalakshmi, "Security and Robustness Enhancement for Image Data Hiding", *Multimedia, IEEE Transactions on*, Vol. 9, No. 3, pp. 466-474, April 2007.

[12] M. Holliman, N. Memon and M. Yeung, "On the need for image dependent keys in watermarking", *Proceedings of IEEE Content Security and Data Hiding in Digital Media,* 1999.

[13] I. Nasir, F. Khelifi, J. Jiang and S. Ipson, "Robust image watermarking via geometrically invariant feature points and image normalisation", *IET Image Processing,* Vol. 6, No. 4, pp. 354-363, 2012.

[14] Fei. Chuhong, D. Kundur, and R.H. Kwong, "Analysis and design of secure watermark-based authentication systems", *Information Forensics and Security, IEEE Transactions on*, Vol. 1, No. 1, pp. 43-55, March 2006.

[15] P. Comesana, F. Balado and F. Perez-Gonzalez, "A novel interpretation of content authentication", In Edward J. Delp III and Ping W. Wong, editors, *Security, Steganography, and Watermarking of Multimedia Contents IX*, Vol. 6505, San Jose, California, USA, January 2007. http://dx.doi.org/10.1117/12.704271

[16] F. Perez-Gonzalez, C. Mosquera, M. Barni and A. Abrardo, "Rational Dither Modulation: A high-rate data-hiding method invariant to gain attacks", *IEEE Transactions on Signal Processing,* Vol. 53, No. 10 II, pp. 3960-3975, 2005.

[17] P. Moulin, and A. K. Goteti, "Block QIM watermarking games", *Information Forensics and Security, IEEE Transactions on*, Vol. 1, No. 3, pp. 293-310, Sep. 2006.

[18] M. Wu, "Joint security and robustness enhancement for quantization based data embedding", *Circuits and Systems for Video Technology, IEEE Transactions on*, Vol. 13, No. 8, pp. 831-841, Aug. 2003.

[19] A. Piva, T. Bianchi, and A. De Rosa, "Secure Client-Side ST-DM Watermark Embedding", *Information Forensics and Security, IEEE Transactions on*, Vol. 5, No. 1, pp. 13-26, March 2010.

[20] P. Bas, "Informed secure watermarking using optimal transport", *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, pp. 1848-1851, 22-27 May 2011.

[21] P. Bas, "Soft-SCS: improving the security and robustness of the Scalar-Costa-Scheme by optimal distribution matching", *In Information Hiding, IH'11, Lecture Notes in Computer Science*, Vol 6958, pp. 208-222, 18-20 May 2011.

[22] F. Cayre, C. Fontaine and T. Furon, "Watermarking security: theory and practice", *Signal Processing, IEEE Transactions on*, Vol. 53, No. 10, pp. 3976-3987, Oct. 2005.

[23] C. E. Shannon, "Communication theory of secrecy systems", *Bell Syst. Tech. J.,* Vol. 28, No. 4, pp. 656-715, 1949.

[24] A. Valizadeh and Z.J. Wang, "An Improved Multiplicative Spread Spectrum Embedding Scheme for Data Hiding", *Information Forensics and Security, IEEE Transactions on*, Vol. 7, No. 4, pp. 1127-1143, Aug. 2012.

[25] M. Costa, "Writing on dirty paper (Corresp.)", *Information Theory, IEEE Transactions on*, Vol. 29, No. 3, pp. 439-441, May 1983.

[26] L. Schuchman, "Dither Signals and Their Effect on Quantization Noise", *Communication Technology, IEEE Transactions on*, Vol. 12, No. 4, pp. 162-165, Dec 1964.

[27] F. Perez-Gonzalez, F. Balado and J. R. H. Martin, "Performance analysis of existing and

new methods for data hiding with known-host information in additive channels", *Signal Processing, IEEE Transactions on*, Vol. 51, No. 4, pp. 960-980, Apr 2003.

[28] R. Zamir and M. Feder, "On lattice quantization noise", *Information Theory, IEEE Transactions on*, Vol. 42, No. 4, pp. 1152-1159, Jul 1996.

[29] A. Papoulis and U. Pillai, *Probability, Random Variables and Stochastic Processes*. New York: McGraw-Hill, 2002.

[30] Y. Wang and P. Moulin, "Steganalysis of block-structured stegotext", in *Proc. Security, Stenography Watermarking of Multimedia Contents*, Vol. 5306, pp. 477-488, San Jose, CA, 2004.

[31] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley-Interscience, 2006.

[32] L. Perez-Freire and F. Perez-Gonzalez, "Spread-Spectrum Watermarking Security", *Information Forensics and Security, IEEE Transactions on*, Vol. 4, No. 1, pp. 2-24, March 2009.

[33] F. Cayre and P. Bas, "Kerckhoffs-Based Embedding Security Classes for WOA Data Hiding", *Information Forensics and Security, IEEE Transactions on*, Vol. 3, No. 1, pp. 1-15, March 2008.

[34] U. Erez, S. Shamai and R. Zamir, "Capacity and lattice strategies for canceling known interference", *Information Theory, IEEE Transactions on*, Vol. 51, No. 11, pp. 3820-3833, Nov. 2005.

Reza Samadi (M'2011) is graduated from Sharif University of technology in 2006. Now he is pursuing his Ph.D. in Ferdowsi University of Mashhad. His research interest covers information security, physical layer security, and applications of information theory.

Seyed Alireza Seyedin was born in Mashhad. He received the B.Sc. degree in Electronics Engineering from Isfehan University of Technology, Isfehan, Iran in 1986, and the M.E. degree in Control and Guidance Engineering from Roorkee University, Roorkee, India in 1992, and the Ph.D. degree from the University of New South Wales, Sydney, Australia in 1996. He has been an Associate Professor with the Department of Electrical Engineering, the Ferdowsi University of Mashhad, Mashhad, Iran. His research interest includes image processing, computer vision, signal processing, and pattern recognition. In these fields, specially, he is interested in image analysis, motion detection and estimation in image sequences, autonomous vehicles, and diverse applications of the radon transform.