SECURITY OF CASS DATA HIDING SCHEME UNDER THE SCENARIOS OF KMA AND WOA

Dong Zhang[⋆] *Dah-Jye Lee*[†]

[⋆] School of Information Science and Technology, Sun Yat-sen University, China
[†]Department of Electrical and Computer Engineering, Brigham Young University, United States

ABSTRACT

This paper presents a theoretical analysis on the security of CASS (Correlation-and-bit-Aware Spread-Spectrum) data hiding scheme for the first time. By evaluated with the residual entropy of the secret key and the mutual information between the observations and the secret key, the security of CASS is investigated under the scenarios of Known Message Attack and Watermarked Only Attack. In addition, this paper compares data hiding security between CASS and the conventional Additive Spread-Spectrum (Add-SS) data hiding scheme. Theoretical analysis and simulation results show CASS scheme outperforms Add-SS scheme in terms of data hiding security when Document to Watermark Ratio (DWR) is low and performs comparably when DWR is high.

Index Terms— Security, CASS data hiding scheme, Additive Spread-Spectrum, Information leakage

1. INTRODUCTION

In recent years, security has attracted considerable attention from the data hiding community. Security of a data hiding scheme measures its performance against attacks aiming at gaining knowledge about the secret key [1] with the assumption of Kerchhoffs' principle which states that all functions should be declared as public except for the secret key [2][3]. Thus the security of a data hiding algorithm is equivalent to that of the secret key and can be evaluated by the difficulty of estimating the secret key [4].

As one of the most popular approaches, spread-spectrum (SS) based method is widely used in many data hiding schemes. Recently, a novel SS-based data hiding approach named Correlation-and-bit-aware Spread-Spectrum (CASS) was proposed [5]. This new approach explores the correlation between the host signal and the secret key as well as the information bit to be embedded into the host signal as the side information in the encoder. It has been shown that CASS is superior to its counterpart, the conventional Additive Spread-Spectrum (Add-SS) scheme, in terms of watermarking decoding performance [5].

This paper explores the security of CASS that has not been studied in the past. From the perspective of Shannon information theory, this paper evaluates the security of CASS with the residual entropy of the secret key given N-time observations and the mutual information between the secret key and the observations. The security of CASS under both Known Message Attack (KMA) and Watermarked Only Attack (WOA) is investigated. By comparing with Add-SS scheme, the result of this work shows that CASS outperforms Add-SS on security when Document to Watermark Ratio (DWR) is low and performs comparably when DWR is high.

2. METHOD

When analyzing security of data hiding schemes, it is assumed that the attacker is able to access the watermarked observations and possibly other kinds of information. For each observation, it is assumed that the same key is used [1][2]. Security analysis also assumes there is no noise imposed onto the observations [3].

Depending on the information the attacker possesses, several security attacking scenarios have been defined by researchers [1][2]. Among these scenarios, KMA and WOA are the two most critical cases. KMA is the type of attack that the attacker has access to the observations and corresponding messages [4]. WOA is the type of attack that the attacker has access only to the watermarked observations [4]. KMA posts the hardest challenge to the watermarker whose responsibility is to keep the data hiding system safe. Whereas, WOA is the most difficult scenario to the attacker. This paper focuses discussions on KMA and WOA scenarios.

Denote $\mathbf{Z}$ as the secret key and $\mathbf{Y}^N = \mathbf{Y}_1, \ldots, \mathbf{Y}_N$ the N-time observations. The security of CASS scheme is measured by the residual entropy of the secret key given N-time observations $h(\mathbf{Z}|\mathbf{Y}^N)$, which represents the uncertainty of the secret key given the observations, and the mutual information between the secret key and the N-time observations $I(\mathbf{Z};\mathbf{Y}^N)$, which represents the information leakage of the secret key from observations. From Shannon information theory [6], we have $h(\mathbf{Z}|\mathbf{Y}^N) = h(\mathbf{Z}) - I(\mathbf{Z};\mathbf{Y}^N)$ and $I(\mathbf{Z};\mathbf{Y}^N) = h(\mathbf{Y}) - h(\mathbf{Y}^N|\mathbf{Z})$ where $h(\cdot)$ is the differential entropy of a random vector and $h(\mathbf{A}|\mathbf{B})$ is the conditional entropy of $\mathbf{A}$ given $\mathbf{B}$. It can be seen that the residual entropy of the secret key depends on the entropy of the secret key

$h(\mathbf{Z})$ and the mutual information between the secret key and observations $I(\mathbf{Z};\mathbf{Y}^N)$.

Assuming the secret key with length N_v follows Gaussian independent and identical distribution (i.i.d.) with zero mean and variance σ_z^2, i.e. $\mathbf{Z} \sim N(\mathbf{0}_{N_v}, \sigma_z^2\mathbf{I}_{N_v})$. The uncertainty of the secret key is evaluated by its differential entropy which can be obtained as

$$h(\mathbf{Z}) = \frac{N_v}{2}\log(2\pi e\sigma_z^2) \tag{1}$$

where the logarithm is taken to base e [7]. Thus, the residual entropy of the secret key is determined only by $I(\mathbf{Z};\mathbf{Y}^N)$. The larger the $I(\mathbf{Z};\mathbf{Y}^N)$ is, the lower security level the data hiding scheme will have.

In the analysis presented in the following sections, we assume all observations are independent to one another and every observation provides the same amount of information about the secret key [2]. The discussion on security is then simplified to concerning only one observation is available to the attacker [2].

3. SECURITY OF CASS

3.1. The CASS scheme

Assume that the host signal $\mathbf{x} = [x_1, x_2, \ldots, x_{N_v}]^T$ follows Gaussian i.i.d. with zero mean and variance σ_x^2, i.e. $\mathbf{X} \sim N(\mathbf{0}_{N_v}, \sigma_x^2\mathbf{I}_{N_v})$, where $\mathbf{I}_{N_v}$ and $\mathbf{0}_{N_v}$ are the $N_v \times N_v$ identity matrix and zero vector with dimension N_v, respectively. N_v is the number of the host coefficients used for conveying one information bit. The embedded message M is assumed to be selected from a binary set $m \in \{-1, +1\}$ with equal probability. The secret key $\mathbf{Z}$ is independent of the host $\mathbf{X}$ and the message M. The watermarked signal is denoted as $\mathbf{y} = [y_1, y_2, \ldots, y_{N_v}]^T$. Thus the scheme of CASS can be expressed as

$$\mathbf{y} = \begin{cases} \mathbf{x} + \mathbf{z}A_1, & \text{if } \mathbf{z}^T\mathbf{x} \geq 0, m = +1 \\ \mathbf{x} - \mathbf{z}A_2, & \text{if } \mathbf{z}^T\mathbf{x} \geq 0, m = -1 \\ \mathbf{x} - \mathbf{z}A_1, & \text{if } \mathbf{z}^T\mathbf{x} < 0, m = -1 \\ \mathbf{x} + \mathbf{z}A_2, & \text{if } \mathbf{z}^T\mathbf{x} < 0, m = +1 \end{cases} \tag{2}$$

where A_1 and A_2 $(0 < A_1 < A_2)$ are two amplitude levels [5]. The distortion and DWR of CASS embedding in (2) can be expressed as (3) and (4), respectively.

$$D_w = \frac{A_1^2 + A_2^2}{2}\sigma_z^2 \tag{3}$$

$$DWR = 10\log_{10}\frac{\sigma_x^2}{\frac{A_1^2+A_2^2}{2}\sigma_z^2} \tag{4}$$

3.2. Under KMA scenario

Under KMA scenario, the attacker is able to gather a collection of observations $\{\mathbf{Y}_i\}$, which is watermarked with the same secret key, and the attacker also knows the corresponding messages $\{m_i\}$. Thus the information leakage of the secret key will turn to $I(\mathbf{Z};\mathbf{Y}^N|M^N) = h(\mathbf{Y}^N|M^N) - h(\mathbf{Y}^N|\mathbf{Z}, M^N)$. When one-time observation is considered, we have

$$I(\mathbf{Z};\mathbf{Y}|M) = h(\mathbf{Y}|M) - h(\mathbf{Y}|\mathbf{Z}, M) \tag{5}$$

For the convenience of description, we assume the correlation between the host and secret key is positive. The same result can be obtained when the correlation is negative. Considering the embedding scheme of CASS (2) and the assumption on host and the secret key, the observation is Gaussian with zero mean and covariance $A_1^2\sigma_z^2\mathbf{I}_{N_v} + \sigma_x^2\mathbf{I}_{N_v}$ given $m = +1$ and $A_2^2\sigma_z^2\mathbf{I}_{N_v} + \sigma_x^2\mathbf{I}_{N_v}$ given $m = -1$. Since $h(\mathbf{Y}|M) = \sum_M p(M = m)h(\mathbf{Y}|M = m)$, and the embedded message takes the values '+1' and '-1' with the same probability, i.e. $p(M = +1) = p(M = -1) = \frac{1}{2}$, the value of $h(\mathbf{Y}|M)$ can be calculated analytically. The second term on the right-hand side of equation (5) can also be obtained. Given the secret key and the embedded message, the observation is also a Gaussian, i.e. $\mathbf{Y} \sim N(A_1\mathbf{z}, \sigma_x^2\mathbf{I}_{N_v})$ when $m = +1$ and $\mathbf{Y} \sim N(-A_2\mathbf{z}, \sigma_x^2\mathbf{I}_{N_v})$ when $m = -1$. Thus we can obtain the value of the information leakage of the secret key in CASS scheme by (6).

$$\begin{aligned} I(\mathbf{Z};\mathbf{Y}|M) &= \sum_M p(M = m)h(\mathbf{Y}|M = m) \\ &\quad - \sum_M p(M = m)h(\mathbf{Y}|\mathbf{Z}, M = m) \\ &= \frac{N_v}{4}\log(1 + \frac{A_1^2\sigma_z^2}{\sigma_x^2})(1 + \frac{A_2^2\sigma_z^2}{\sigma_x^2}) \end{aligned} \tag{6}$$

Formula (6) shows that the information leakage of the secret key in CASS depends on the length of the secret key, the amplitude levels A_1 and A_2, and the covariance ratio between the secret key and the host. When the distortion is constrained, different values assigned to A_1 and A_2 will lead to distinct security on CASS.

3.3. Under WOA scenario

The information leakage about the secret key under WOA scenario can be expressed as (7).

$$I(\mathbf{Y};\mathbf{Z}) = h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{Z}) \tag{7}$$

The first term on the right-hand side of (7) is the differential entropy of observation. The distribution of $\mathbf{Y}$ can be expressed as the summation of $p(\mathbf{Y}, M)$, which is the joint distribution of $\mathbf{Y}$ and M. The summation is performed with respect to M. Considering $p(\mathbf{Y}, M) = p(\mathbf{Y}|M)p(M)$, $p(\mathbf{Y}|M = +1) = N(\mathbf{0}_{N_v}, \sigma_x^2\mathbf{I}_{N_v} + A_1^2\sigma_z^2\mathbf{I}_{N_v})$, $p(\mathbf{Y}|M = -1) = N(\mathbf{0}_{N_v}, \sigma_x^2\mathbf{I}_{N_v} + A_2^2\sigma_z^2\mathbf{I}_{N_v})$, and the embedded message takes the values '+1' and '-1' with the same probability,

the distribution of $\mathbf{Y}$ is a mixture of Gaussians as shown in (8).

$$\begin{aligned} p(\mathbf{Y}) &= \sum_{M} p(\mathbf{Y}|M=m_i)p(M=m_i) \\ &= \frac{1}{2}[N(\mathbf{0}_{N_v}, \sigma_x^2\mathbf{I}_{N_v} + A_1^2\sigma_z^2\mathbf{I}_{N_v}) \\ &\quad + N(\mathbf{0}_{N_v}, \sigma_x^2\mathbf{I}_{N_v} + A_2^2\sigma_z^2\mathbf{I}_{N_v})] \end{aligned} \tag{8}$$

It has been proved that there is no closed form for the differential entropy of a random vector which is distributed as a mixture of Gaussians [8]. Given the zero mean and the variance of the Gaussian components, $h(\mathbf{Y})$ can not be derived from its definition analytically. Numerical computation is required to sovle for $h(\mathbf{Y})$. It can be seen that the distribution of $\mathbf{Y}$ is composed of two Gaussians, both with zero mean. Since the two Gaussians appear with the same probability, the differential entropy of $\mathbf{Y}$ only depends on the covariance of them.

Solving the second term on the right-hand side of (8) also needs numerical computation. The reason is shown below in (9).

$$\begin{aligned} p(\mathbf{Y}|\mathbf{Z}) &= \sum_{M} p(\mathbf{Y}|\mathbf{Z}, M)p(M|\mathbf{Z}) \\ &= \sum_{M} p(\mathbf{Y}|\mathbf{Z}, M)p(M) \\ &= \frac{1}{4}N(A_1\mathbf{z}, \sigma_x^2\mathbf{I}_{N_v}) + \frac{1}{4}N(A_2\mathbf{z}, \sigma_x^2\mathbf{I}_{N_v}) \\ &\quad + \frac{1}{4}N(-A_1\mathbf{z}, \sigma_x^2\mathbf{I}_{N_v}) + \frac{1}{4}N(-A_2\mathbf{z}, \sigma_x^2\mathbf{I}_{N_v}) \end{aligned} \tag{9}$$

The first equation of (9) is from the definition of conditional probability. The second equation of (9) is due to the independence between the embedded message M and the secret key $\mathbf{Z}$, i.e. $p(M|\mathbf{Z}) = p(M)$. Equation (9) shows the distribution of $\mathbf{Y}$ given the secret key is also a mixture of Gaussians and numerical computation is required to obtain the value of $h(\mathbf{Y}|\mathbf{Z})$. Similar to the notes on the computation of $h(\mathbf{Y})$, it can be seen from (9) that the differential entropy of $\mathbf{Y}$ given the secret key will only be dependent on the means of each Gaussian components for a fixed covariance of host.

4. SIMULATION AND DISCUSSION

The simulation results were obtained with setting the length of secret key as 256 unless stated otherwise. The information leakage of the secret key was averaged into all dimensions. Since the logarithm used in analysis is taken to the base e, the information leakage is expressed in nat.

The information leakage of the secret key vs. DWR under KMA scenario is shown in Fig.1, where different amplitude values A_1 are investigated. We denote that $\eta = \frac{A_1^2}{A_1^2+A_2^2}$, and

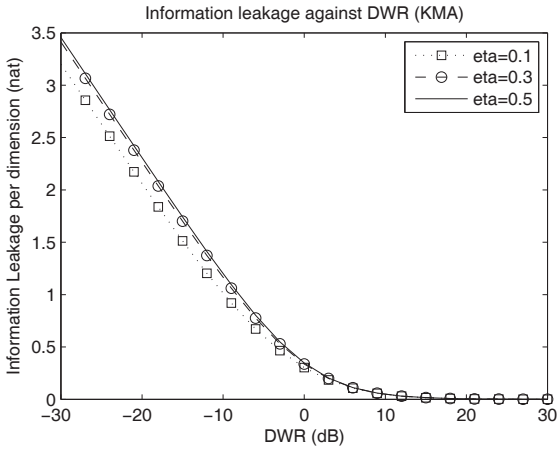

Fig. 1. Information Leakage vs. DWR under KMA scenario.

$1-\eta = \frac{A_2^2}{A_1^2+A_2^2}$. Considering $0 < A_1 < A_2$, we have the range of η, i.e. $0 < \eta \leq \frac{1}{2}$. It can be seen in Fig.1 that a small η (for example 0.1) has less information leakage than a large η (for example 0.5). As DWR increases, or the embedded power decreases, the secret key information leakage decreases for the same η. Note that the Add-SS watermarking scheme can be considered as a special case of CASS when $A_1 = A_2$, or $\eta = 0.5$. Fig.1 shows that the security of Add-SS is lower than CASS because Add-SS leaks more information of the secret key for the same DWR.

For the Add-SS scheme, it has been shown in [2] that for one observation, the information leakage of the secret key can be expressed as $I(\mathbf{Y};\mathbf{Z}|M) = \frac{N_v}{2}\log(1+\frac{\sigma_z^2}{\sigma_x^2})$. For a fare comparison, the same DWR is assumed for Add-SS and CASS schemes, e.g. DWR$= 10\log_{10}(\frac{1}{C})$, in which C is a fixed value. Then the secret key information leakage for Add-SS is obtained by (10).

$$I_{Add-SS}(\mathbf{Y};\mathbf{Z}|M) = \frac{N_v}{4}\log(1+2C+C^2) \tag{10}$$

For CASS, with the same DWR, which implies $\frac{A_1^2+A_2^2}{2}\sigma_z^2 = C\sigma_x^2$, the secret key information leakage can be computed by (11).

$$\begin{aligned} I_{CASS}(\mathbf{Y};\mathbf{Z}|M) &= \frac{N_v}{4}\log(1+\frac{2A_1^2C}{A_1^2+A_2^2})(1+\frac{2A_2^2C}{A_1^2+A_2^2}) \\ &= \frac{N_v}{4}\log[(1+2C\eta)(1+2C(1-\eta))] \end{aligned} \tag{11}$$

It can be seen that $(1+2C\eta)[1+2C(1-\eta)] \leq (1+2C+C^2)$ for $0 < \eta \leq \frac{1}{2}$. The equality is true when $\eta = \frac{1}{2}$, that means $A_1^2 = A_2^2$. This implies that, for the same DWR, the Add-SS scheme will not leak less information about the secret key than the CASS scheme, i.e. $I_{Add-SS}(\mathbf{Y};\mathbf{Z}|M) \geq$

$I_{CASS}(\mathbf{Y};\mathbf{Z}|M)$. For a high DWR, the difference in the secret key information leakage induced by different values of η is rather small because the embedded power of the watermark is rather weak compared with the power of the host signal.

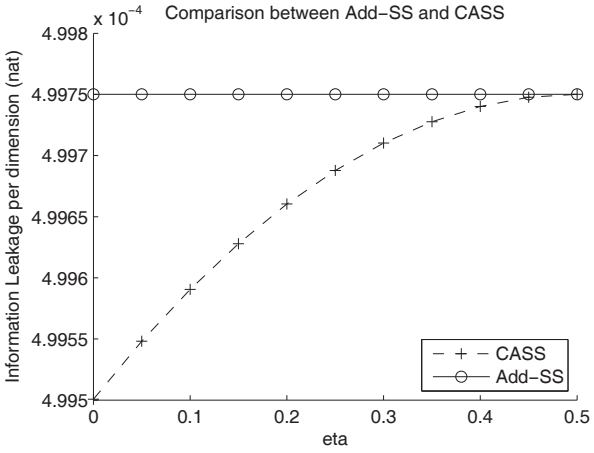

Fig. 2. Comparison between Add-SS and CASS under KMA scenario.

Another experiment was performed to compare the secret key information leakage vs. different η values with a fixed DWR, which was set to 30 dB. In Fig.2, the solid line marked with circles is the secret key information leaked from Add-SS and the dashed line marked with crosses is the secret key information leaked from CASS. Note that the security of Add-SS scheme is not affected by η because the amplitudes for embedding '+1' and '-1' are set to be equal. The secret key information leaked from CASS was much less than Add-SS. The performance difference in security becomes smaller as η increases, and turns to zero when η equals to 0.5 at which CASS is actually equivalent to Add-SS scheme.

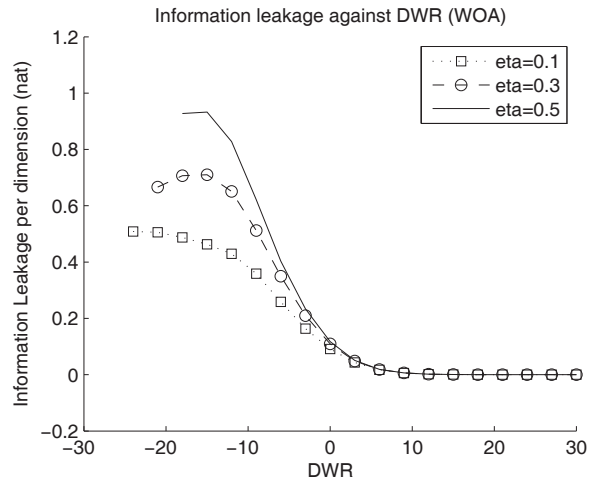

Fig. 3. Information leakage vs. DWR under WOA scenario.

Similar comparison was performed for WOA scenario. The curves showing the secret key information leakage vs. DWR for different η values are shown in Fig.3. Our result shows that the relation between the secret key information leakage and η is similar to the KMA scenario.

5. CONCLUSION

This paper investigates the security of CASS under both KMA and WOA scenarios from the perspective of Shannon information theory. Theoretical analysis and simulations show the information leakage of the secret key in CASS is not larger than Add-SS scheme. This result proves that CASS has a higher security level than Add-SS. Similar to the relation shown in the comparison of decoding performance [5], the security of CASS outperforms Add-SS distinctly when DWR is low, and has comparable performance when DWR is high.

6. ACKNOWLEDGEMENTS

This work was supported in part by the National Science Foundation of China under Grant (nos.61100170).

7. REFERENCES

[1] F. Cayre, C. Fontaine, T. Furon. "Watermarking security: theory and practice," in IEEE Trans. Signal Processing. 2005, vol.10, pp.3976-3987.

[2] P. Comesaña, L. Pérez-Freire, and F. Pérez-González. "Fundamentals of data hiding security and their application to spread-spectrum analysis," in Lecture Notes in Computer Science, IH05, Springer-Verlag. 3727, 2005.

[3] F. Cayre, P. Bas. "Kerckhoffs-based embedding security classes for WOA data hiding," in IEEE Transactions on Information Forensics and Security. 2008, vol.3, pp.11-15.

[4] L. Pérez-Freire, F. Pérez-González. "Spread Spectrum Watermarking Security," in IEEE Transactions on Inoformation Forensics and Security. 2009, vol.4, pp.2-24.

[5] A. Valizadeh, Z. J. Wang. "Correlation-and-bit-aware spread spectrum embedding for data hiding," in IEEE Transactions on Information Forensics and Security. 2011, vol.6, pp.257-282.

[6] T. M. Cover, J. A. Thomas. Elements of Information Theory. Wiley series in Telecommunications. 1991.

[7] L. Pérez-Freire, P. Comesaña, J. R. Troncoso-Pastoriza, and F. Pérez-González, "Watermarking security: a survey," in Transactions on Data Hiding and Multimedia Security I. 4300, 2006, pp.41-72.

[8] J. V. Michalowicz, J. M. Nichols, F. Bucholtz. "Calculation of differential entropy for a Mixed Gaussian Distribution," in Entropy. 2008, vol.10, pp.200-206.