ON THE SECURITY AND ROBUSTNESS OF ENCRYPTION VIA COMPRESSED SENSING

Adem Orsdemir, H. Oktay Altun, Gaurav Sharma, Mark F. Bocko

ECE Dept., University of Rochester, Rochester, NY, USA

ABSTRACT

The compressed sensing (CS) paradigm unifies sensing and compression of sparse signals in a simple linear measurement step. Reconstruction of the signal from the CS measurements relies on the knowledge of the measurement matrix used for sensing. Generation of the pseudo-random sensing matrix utilizing a cryptographic key, offers a natural method for encrypting the signal during CS. This CS based encryption has the inherent advantage that encryption occurs implicitly in the sensing process – without requiring additional computation. Additionally, the robustness of recovery from compressed sensing, allows a new form of "robust encryption" for multimedia data, wherein the signal is recoverable with high fidelity despite the introduction of additive noise in the encrypted data.

In this paper, we examine the security and robustness of this CS based encryption method. The security implications are investigated by considering brute force and structured attacks. Robustness is characterized empirically. Our analysis and results indicate that the computational complexity of these attacks renders them infeasible in practice. In addition, the CS based encryption is found to have fair robustness against additive noise, making it a promising "robust encryption" technique for multimedia.

1. INTRODUCTION

Digital multimedia signals often need to be transmitted through a channel or a network. Prior to transmission, it is desirable to compress the multimedia signal for efficient usage of storage resources and/or bandwidth of the communication channels. This compression step is performed either in a lossy or lossless way depending on the needs of the receiver. In addition, when the content of the media is private, security of the transmission must be considered. Typically, encryption of the compressed multimedia is performed following the compression. This step is performed either by conventional cryptographic algorithms or some custom design joint compression and encryption schemes [1].

The recently proposed compressed sensing (CS) framework [2] is known to unify sampling and compression in order to reduce the data acquisition and computational load at sensors, at the cost of increased computation at the intended receiver. Compressive sampling relies on the sparseness of the signal and gathers linear measurements $y = Ax$ of a sparse signal x, where size of y is a small fraction of the samples needed for Nyquist sampling. A is the linear transform which carries certain regularities. The receiver obtains the linear measurements y and reconstructs the image by solving an optimization problem.

Compressed sensing also provides nice encryption properties. The measurements y are a function of sensing matrix A. This matrix has pseudo-random entries that can be generated by using a cryptographic key shared between the sender and receiver. Since the receiver has to know this information in order to formulate the optimization problem and to reconstruct the signal, the CS measurements can be considered as an encrypted representation of the original signal. This idea has been briefly mentioned in the literature [3] but has not been addressed in detail.

In this paper we investigate the security of CS based encryption methods. The security of the encryption method relies on the fact that the sensing matrix A is not known to an attacker that does not have the pseudo-random key used to generate A. We therefore consider attacks aimed at estimating this matrix either based on brute force search or utilizing the symmetry and structure of the CS setup. In addition, we note that CS based encryption also represents a new type of "robust encryption" that is tolerant to additive noise in the CS based measurements that form the encrypted data and empirically characterize this robustness. Our results indicate that the CS based encryption is computationally secure against the investigated attacks, i.e., in order to be successful with high probability the computational requirements render the attacks infeasible.

This paper is organized as follows: Section 2 gives some brief information about compressed sensing. Section 3 explains the notion of encryption by linear measurement matrix. Section 4 mentions possible attack scenarios and their complexity. In Section 5 experimental results are given. Finally Section 6 concludes the paper.

2. COMPRESSIVE SAMPLING BASICS

Compressive sampling considers the problem of recovering an object $x \in \mathbb{R}^{n\times 1}$ from its linear measurements $y = Ax$ where $A \in \mathbb{R}^{m\times n}$ is a transform matrix. The transform matrix has fewer rows than columns, i.e., $m < n$. This system of equations has infinitely many solutions. When the signal x is known to be sparse, to recover x, an optimization problem can be cast as:

This work is supported by the Air Force Office of Scientific Research (AFOSR) under grant number FA9550-07-1-0017.

978-1-4244-2677-5/08/$25.00 ©2008 IEEE

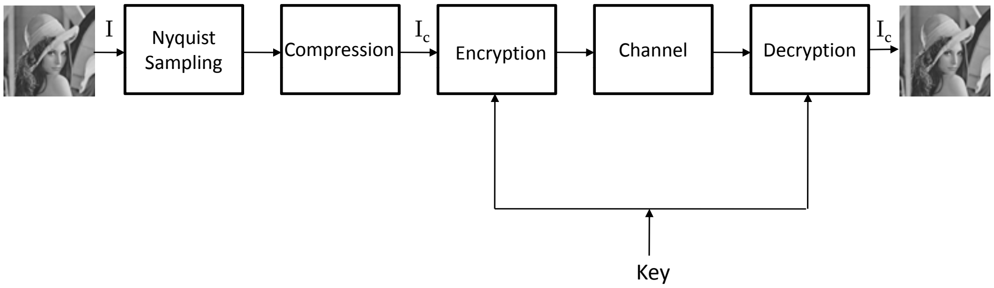

Figure 1: Flow diagram of a conventional sampling, compression and encryption scheme.

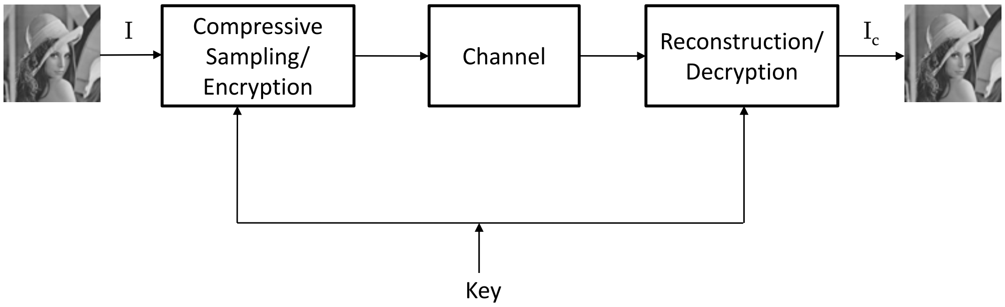

Figure 2: Flow diagram of compressive sampling for unified sampling, compression and encryption.

$$(P_0) \ \min_x \ \|x\|_0 \quad \text{subject to} \quad Ax = y \qquad (1)$$

where $\|x\|_0 = \sum \chi(x_i)$, where

$$\chi(t) = \begin{cases} 0 & t = 0 \\ 1 & \text{otherwise} \end{cases} \qquad (2)$$

This problem is a non-convex optimization problem and a solution requires an exhaustive search over the solution space, which is computationally infeasible for most problems of interest. Alternatively, the problem can be solved by greedy algorithms such as orthogonal matching pursuit [4], which however is not guaranteed to achieve the optimum.

If the measurement matrix A satisfies the *restricted isometry property* (RIP) an alternative convex formulation whose solution is identical to the solution for the problem (P_0) can be obtained as [5]:

$$(P_1) \ \min_x \ \|x\|_1 \quad \text{subject to} \quad Ax = y \qquad (3)$$

where $\|x\|_1 = \sum_i |x_i|$ denotes the ℓ_1 norm of the vector x. The ℓ_1 term penalizes the small components in x and pushes them close to zero. The convexity of the problem (P_1) allows for a number of effective, globally convergent solution techniques. Various computational methods have been proposed to solve the problems of this form, including but not limited to conjugate gradient (CG) methods [6], path-following methods [7], bound optimization methods, gradient projection algorithms, interior point methods [8] and preconditioned conjugate gradient methods [9].

In the presence of noise on measurements, a similar problem can be formulated as [10]:

$$(P_2) \ \min_x \ \|x\|_1 \ \text{subject to} \quad \|Ax - y\|_2 \leq \delta \qquad (4)$$

where δ represents the expected noise power in the observations. This formulation provides a *stable* recovery of the original signal value i.e. if the matrix A satisfies RIP, the error in solution to the (4) is bounded by a term proportional to the noise variance.

3. ROBUST ENCRYPTION VIA COMPRESSED SAMPLING

Conventionally, real life signals are sampled conforming the Nyquist sampling rule. The data is then compressed to reduce the data size and encrypted for security. The encrypted data usually goes through a channel and is decrypted by the intended user. Both the sender and recipient share a secret key. Figure 1 illustrates the conventional sampling, compression and encryption schemes in a flow diagram.

Compressive sampling, on the other hand, unifies the sampling, compression, and encryption steps as illustrated in Fig. 2. Compressive samples are gathered by the sensor and delivered to the recipient throughout the channel. An estimate of the desired signal is then reconstructed by the receiver, utilizing these samples. The linear measurement matrix A should be available to the receiver side in order to recover the signal. Without the knowledge of A the gathered samples appear encrypted to anyone eavesdropping on the channel [3]. This encryption comes naturally and require no additional cost.

Compressive samples are proven to be robust against some level of noise. This noise can be either due to quantization or due to a noisy channel. The signal can still be reconstructed with some degradation that depends on the noise level. This provides the notion of *robust encryption* where the encrypted signal is robust against noise. This is an unusual property of encryption which is uncommon in conventional encryption schemes, where even minute changes can cause the loss of all the information.

The concept of robust encryption is in harmony with the nature of multimedia files. Multimedia files carry a considerable amount of perceptual redundancy and a human observer will tolerate some amount of degradation. From a security point of view, this brings two advantages. The intruder listening on the channel cannot decrypt the message. And even if he tries to block the communication with the addition of noise, the samples will be resilient against the attack up to a certain level.

4. ATTACKS ON COMPRESSED SENSING BASED ENCRYPTION

In this section, we consider two possible attacks on compressed sensing based encryption schemes for sparse signals: firstly a brute force attack through a search for the measurement matrix A and by a more informed signal processing attack that exploits the symmetry and sparsity structure inherent in compressed sensing.

4.1 Brute Force Attack

An immediate attack on the compressive sampling based encryption scheme would be guessing the linear measurement matrix A. The eavesdropper could directly try to do this by performing an exhaustive search over a “grid” of values for A. The step size of this grid is critical since too large a step size may cause the search to miss the correct value and too small a grid size will increase the computational burden unnecessarily.

The step size can be empirically determined by disrupting the measurement matrix A and determining the quality of the signal for that predetermined noise level. The corresponding noise level for acceptable quality level will provide us a confidence interval for the estimation of a optimum step size for a grid search.

The computational cost of reconstructing the signal is high. The best optimization algorithm as of now requires a computational cost in the order of $O(N^{1.2})$ [8]. This cost accompanied with a random search will already make the search too expensive. However, it is important to show that a large step size does not suffice for the attack, which would make the whole encryption scheme trivially insecure.

4.2 Attack Based on Symmetry and Sparsity Structure

Given the measurements y, an attacker's goal is to estimate an $m \times n$ matrix A and a $n \times 1$ sparse signal x with t nonzero coefficients such that $Ax = y$. We note that this problem is clearly degenerate[1] and allows for multiple solutions. In order to see this note that if we identically permute the columns of A and the coefficients of x we obtain the same measurements vector y. Mathematically the (original) measurements can be expressed as:

$$y_i = \sum_{j=1}^{n} (a_{ij} x_j) \quad i = 1, \dots m \tag{5}$$

if π denotes a permutation of integers $[1, \dots n]$ we have

$$y_i = \sum_{j=1}^{n} a_{i\pi(j)} x_{\pi(j)} \quad i = 1, \dots m \tag{6}$$

Thus the measurements y are also consistent with the alternate matrix A' and sparse signal x'(again with t nonzero coefficients), where the $(ij)^{th}$ element of A' is $a_{i\pi(j)}$ and j^{th} element of x' is $x_{\pi(j)}$. In particular, we note that without loss of generality, we can assume that the t leading coefficients of x' are nonzero. Consequently, an eavesdropper wishing to attack the compressed sensing based encryption, may decompose his attack in two phases: First estimate the t leading columns of A', viz. $A_t = [a'_1 a'_2 \dots a'_t]$, and the corresponding coefficients $x_t = [x'_1 x'_2 \dots x'_t]$ such that[2]

$$A_t x_t = y \tag{7}$$

where a'_i's denotes the i^{th} column of A'. Once such a combination is found, assuming the problem has no additional degeneracy, the attacker is faced with challenge of determining the permutation $\pi()$. Since $(n-t)$ of the coefficients in x are zero, a brute force search for this permutation requires $C(n,t) \times t!$ possible arrangements.

Note that (7) represents a overdetermined system of m equations in t variables. The least squared solution to this overdetermined system is given by [11]

$$x_t^* = (A_t)^{\#} y \tag{8}$$

where $B^{\#}$ denotes the Moore-Penrose pseudo inverse of the matrix B. The corresponding least squares residual is given by

$$\rho = ||A_t x_t^*|| = ||(I - A_t (A_t)^{\#}) y|| = ||P_{A_t}^{\perp} y|| \tag{9}$$

[1] Though for multimedia signals the attacker may have additional methods for choosing multiple solutions, e.g a semantically meaningful signal.

[2] We have the reasonable assumption the attacker can estimate t.

where $P_{A_t}^{\perp}$ is the orthogonal projection matrix that projects onto the orthogonal complement of the columns of A_t.

Now let $A_t = Q_1 R$ denote the QR decomposition of the matrix A_t, where Q_1 is $m \times t$ matrix with orthonormal columns,and R is a $t \times t$ nonsingular upper triangular matrix. Let Q_2 be any $m \times (m-t)$ matrix with orthonormal columns such that $[Q_1 Q_2]$ forms an orthonormal basis for $\mathbb{R}^m$. Then in terms of this decomposition $P_{A_t}^{\perp} = Q_2 Q_2^T$ and using the orthonormality of the columns of Q_2, the residual is obtained as:

$$\rho = ||Q_2^T y||^2 = \sum_{i=1}^{m-t} (q_i^{(2)T} y)^2 \tag{10}$$

where $q_i^{(2)}$ denotes the i^{th} column of the matrix Q_2. It follows that the residual ρ is zero, equivalently the system of equations (8) is consistent, if and only if, the orthonormal vectors $\{q_i^{(2)}\}_{i=1}^{t}$ are orthogonal to the vector measurements y.

Using the above result we can now fully characterize the complexity of the attacker's task. The $m \times 1$ measurement vector y defines an $(m-1)$ dimensional space orthogonal to y, in which the attacker must choose $(m-t)$ orthonormal vectors to define Q_2. For the orthogonal complement of the column space of Q_2 an orthonormal basis set $Q_1 = [q_1^{(1)} \ldots q_t^{(1)}]$ can be obtained (for any given) Q_2[3]. The attacker must then select a nonsingular $t \times t$ upper triangular matrix R. The combination defines a consistent system of equations, $A_t x_t = y$, where $A_t = QR$ for which a solution x_t may be obtained. Once the solution x_t is available, the attacker must investigate (by other means) all possible $C(n,t) \times t!$ rearrangements of the terms in x_t into an $n \times 1$ vector in which other $(n-t)$ entries are zero.

The combinatorial expressions above indicate that the complexity of this structured attack is too high to be practical. However assuming that the elements other than R are known to the attacker, we can then characterize the difficulty of the (remaining) task. In a manner similar to Section 4.1 by determining the acceptable tolerance up to which the matrix R must be known for obtaining an acceptable approximation for x_t. This can be analytically estimated from known results in matrix perturbation theory [12, pp. 125]. In particular, if $\kappa(R) = ||R||||R^{-1}||$ is the condition number[4] of the matrix R, and $\kappa(R)||E||_F/||R||_F < 1$ we have

$$||\tilde{x} - x_t||/||x_t|| \leq \frac{\kappa(R)||E||_F/||R||_F}{1 - \kappa(R)||E||_F/||R||_F} \tag{11}$$

where $||E||_F$ denotes[5] the perturbation on R and $\tilde{x}_t$ denotes the resulting perturbed value of x_t. We explore this bound against empirical estimates as shown in Section 5.2.

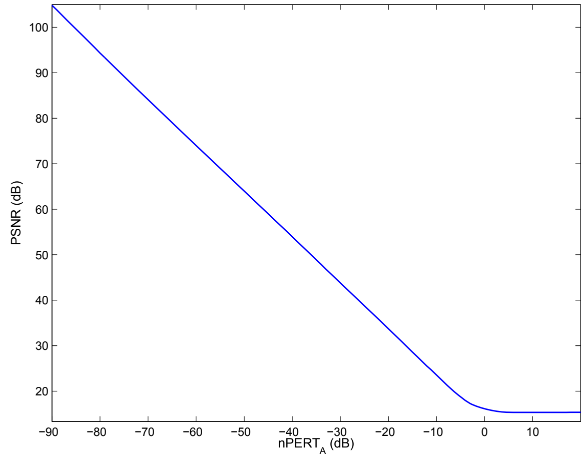

Figure 3: Normalized perturbation power on measurement matrix A vs PSNR of the reconstructed signal

5. EXPERIMENTAL RESULTS:

The experiments utilized randomly generated sparse signals of length $n = 1024$ with $t = 30$ non-zero coefficients each taking a value in $\{\pm 1\}$ randomly, where the locations of the t nonzero coefficients among the n samples was also chosen randomly.

We consider $m = 256$ measurements using a 256×1024 sensing matrix A that is formed by sampling *iid* entries from the normal distribution with mean zero and unity variance. Multiple experiments were conducted for different realizations of the signal and the results presented represent averages over these realizations.

5.1 Estimating the step size of a possible grid search

To estimate the grid size that needs to be utilized in order to estimate the encoder matrix A by a brute force grid search, we proceed as follows. We corrupt the entries in linear measurement matrix A by *iid* Gaussian noise and reconstruct the signal from the measurements by using this corrupted version of linear measurement matrix $\widetilde{A}$.

We examine the relation between the amount of perturbation on the linear measurement matrix versus reconstructed signal quality in Fig. 3. The horizontal axis in Fig. 3 shows the level of normalized perturbation ($nPERT_A$) power on A in decibels (dBs). The normalized perturbation level in dB is calculated as :

$$nPERT_A = 10log_{10}\frac{||\widetilde{A} - A||_F^2}{||A||_F^2} \tag{12}$$

where $\tilde{A}$ is $\tilde{A} = A + Z$, Z is the random perturbation having *iid* $\mathcal{N}(0, \sigma^2)$ entries. On the other hand, the vertical axis of Fig. 3 shows the *PSNR* of the reconstructed signal associated with $\tilde{A}$. The *PSNR* is calculated as follows:

$$PSNR = 10log_{10}\frac{1}{\frac{1}{n}||\tilde{x} - x||^2} \tag{13}$$

[3] Note that Q_1 is defined by Q_2, up to rotations and coordinate flips

[4] $||.||$ is the l_2 norm of a matrix, i.e maximum singular value of the matrix.

[5] $||.||_F$ denotes the Frobenious norm [11]

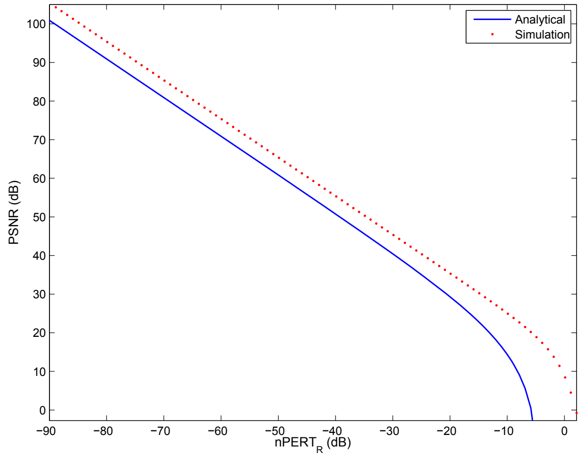

Figure 4: Normalized perturbation power on upper triangular matrix R vs PSNR of the reconstructed signal(Analytical and simulation results)

where $\tilde{x}$ is the signal recovered by CS recovery algorithm [10] using the measurements y and assuming a measurement matrix $\widetilde{A}$.

Figure 3 reveals the linear relation between the perturbation power and the reconstructed signal quality. The signal quality reaches a floor around 0 $dB\ nPERT_A$ value indicating random signal reconstruction. For many image and signal processing applications, 40 dB signal quality indicates an acceptable quality signal reconstruction. For the considered reconstruction problem, around $-26dB$ perturbation of reconstruction matrix provides a 40 dB PSNR signal. If the attacker can guess the reconstruction matrix with perturbation on A $-26dB$ or less, he can reach an acceptable quality signal. This is an important guideline for the choice of grid size for the attackers.

For 40 dB signal quality, the standard deviation(σ) of perturbation corresponding each element of the A matrix is 0.15×10^{-4}. The confidence interval for normal distribution with mean 0 and standard deviation 0.15×10^{-4} is (-0.004,0.004) for a confidence interval of 99 percent. The estimated number of trials in order to be assured of achieving a close enough reconstruction ($40dB$) with high probability is approximately $(\frac{1}{0.008})^{m \times n} = (125)^{mn}$ which is clearly prohibitive[6].

5.2 Estimating the step size of a possible grid search based on symmetry and sparsity structure

For the attack exploiting symmetry and sparsity described in Section 4.2 we can use a methodology similar to that of the preceding section in order to estimate the complexity of grid search step or estimating R. We compare this against the analytic estimates based on (11). We first apply a Q-R decomposition on the matrix $A_t = QR$. We reconstruct the x_t vector via the relation $x_t = R^{-1}Q^T y$. We perturb the matrix R with Gaussian noise of different variances and obtain an estimate of the sparse signal x through x_t. The attack power versus reconstructed signal quality is given in Fig.4 where it is compared against the analytic (lower bound) estimates from (11).

Figure 4 reveals the relation between the perturbation power on the upper triangular matrix R and the reconstructed signal quality. Similar to the grid size prediction analysis in the previous section, for 40 dB signal quality, the standard deviation (σ)of perturbation corresponding each element of the A matrix is 0.12×10^{-4}. The confidence interval for normal distribution with mean 0 and standard deviation 0.12×10^{-4} is (-0.002,0.002)for a confidence interval of 99 percent. Similar to the 5.1 to reconstruct the signal with a PSNR value close to $40dB$ with high probability number of trials needed is approximately $(1/0.004)^{t^2} = (250)^{t^2}$. This result suggests that, although the attacker enjoys the reduced search space and by-passes the computation cost corresponding to the optimization problem, he needs to perform the search on a finer grid, with half the step size compared to the brute force attack. Furthermore as described in Section 4.2 overall attack has much higher complexity.

5.3 Robustness of the encryption system

In this section, we consider the scenario where we do not have the exact encrypted message in the form of the measurements y. Instead we have the noisy $\tilde{y}$ version of the measurements y. Thus, we examine the robustness of the encryption system against this noise. This corresponds exactly to the analysis of "stable recovery" for CS [10]. For completeness, however, we include an empirical evaluation here. Figure 5 illustrates the robustness of the encryption. Horizontal axis is the noise power in the measurement vector y which is given by the formula:

$$nPERT_y = 10log_{10}\frac{||\widetilde{y}-y||^2}{||y||^2} \qquad (14)$$

where $\widetilde{y}$ is the noisy measurement vector. Vertical axis is the normalized PSNR which is described by (13). We again used the same parameters. Note that the PSNR of the recovered signal remains above $40dB$ for noise perturbation power under $-30dB$. Thus the system demonstrates fair robustness against the noise.

6. CONCLUSION AND DISCUSSION

This paper examines the security and robustness of a compressed sensing based encryption algorithm where a shared key-based generation of the measurement matrix provides the encryption. The tecnique is generic for signals which are sparse or have sparse representations. The immediate attack for such a scheme would be to estimate the linear measurement matrix by a grid search. We consider the complexity of brute force and structured approaches for the problem of

[6] Note that the entries of A matrix can at most be 1 due to normalization.

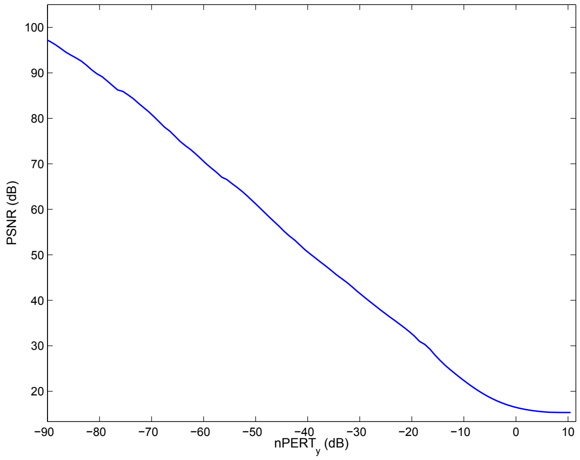

Figure 5: Normalized perturbation power on measurement vector *y* vs PSNR of the reconstructed signal

estimating the measurement matrix. The high computational complexity of these approaches makes them practically infeasible. Compressed sensing therefore provides a “robust encryption” method that is resilient to the attacks investigated.

In our investigation we noted in Section 4.2 that from the attacker perspective the problem of estimating a measurement matrix is a degenerate one with multiple solutions. The attacker must select among these by potentially examining the semantic meaningfulness of the recovered signal. This renders the attack even harder. On the other hand, for multimedia signals additional statistical information is often available regarding the sparsity structure, which if incorporated can partly reduce the attacker’s complexity. Considerations of both of these aspects would be useful in further investigations of this problem.

A theoretical characterization of the impact of a perturbation of the basis matrix A on the quality of the recovered signal is also of interest and a subject of our ongoing investigation. In this setting, we anticipate that the error in variables framework utilized for set theoretic estimation [13] is likely to provide relevant background.

In practical applications of CS based encryption, the matrix A would normally be employed after normalization. For instance, in our implementation the rows were normalized to have a unit norm. It would be useful to incorporate this knowledge in even more “informed attacks”, though it seems that such an approach would not cause a significant compromise in security.

We also note that we have not considered known plaintext attacks in our work where both the signal x and the CS measurements y are available to attacker (for a number of signals). These attacks can be avoided by refraining from reuse of a fixed A matrix, for instance by generating a sequence of matrices based on the key.

7. APPENDIX

7.1 Restricted Isometry Property:

We provide a brief introduction to the notion RIP [5, 10]. Let A_T be a submatrix of A formed by taking any T columns of A, where $T \leq S$; with S denoting an upper bound the sparsity of the signal of interest x_0.The S-restricted isometry constant of A is then defined as the smallest value of δ_S which satisfies:

$$(1-\delta_S)||c||_{\ell_2}^2 \leq ||A_T c||_{\ell_2}^2 \leq (1+\delta_S)||c||_{\ell_2}^2 \qquad (15)$$

for all $T \leq S$ and all coefficient sequences $(c_j)_{j \in T}$.

Observe that when A_T is a matrix with orthonormal columns, we get $\delta_S = 0$. Thus δ_S represents how closely the system of linear equations with coeffiecients given by A_T for $T \leq S$ behave as an orthonormal system.

The matrix A satisfies the RIP if:

$$\delta_S + \delta_{2S} + \delta_{3S} < 1 \qquad (16)$$

and in this case the solution of the ℓ_1 minimization in problem (P1) recovers x_0 [10].

REFERENCES

[1] M. Johnson, P. Ishwar, V. M. Prabhakaran, D. Schonberg, and K. Ramchandran, “On compressing encrypted data,” *IEEE Trans. on Signal Processing*, vol. 51, pp. 2992–3006, Oct 2004.

[2] E. Candes and M. Wakin, “An introduction to compressive sampling [a sensing/sampling paradigm that goes against the common knowledge in data acquisition],” *IEEE Sig. Proc. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[3] D. Takhar, J. N. Laska, M. B. Wakin, M. F. Duarte, D. B. S. Sarvotham, K. F. Kelly, and R. G. Baraniuk, “A new camera architecture based on optical-domain compression,” in *Proc. IST/SPIE Symposium on Electronic Imaging: Computational Imaging*, vol. 6065, 2006, pp. 129–132.

[4] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, “Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition,” in *Conf. Rec. 27th Asilomar Conf. Signals, Syst. Comput.*, vol. 1, 1993.

[5] E. Candes and T. Tao, “Decoding by linear programming,” *IEEE Trans. Inform. Theory*, vol. 51, no. 12, Dec 2005.

[6] E. Candes, “L1-magic: Recovery of sparse signals,” *http://www.acm.caltech.edu/l1magic/*. [Online]. Available: http://www.acm.caltech.edu/l1magic/

[7] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, “Least angle regression,” Jun 2004. [Online]. Available: http://arxiv.org/abs/math.ST/0406456

[8] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, “An interior-point method for large-scale ℓ_1-regularized least squares,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 606–617, 2007.

[9] M. L. S. Kim, K. Koh and S. Boyd, “An efficient method for compressed sensing,” in *ICIP (3)*, 2007, pp. 117–120.

[10] E. Candes, J. Romberg, and T. Tao, “Stable signal recovery from incomplete and inaccurate measurements,” *Comm. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, Aug 2006.

[11] G. Strang, *Introduction to Linear Algebra*. Wellesley-Cambridge Press, 1993.

[12] G. Stewart and J. Sun, *Matrix Perturbation Theory*. New York: Academic Press, 1990.

[13] G. Sharma and H. J. Trussell, “Set theoretic signal restoration using an error in variables criterion,” *IEEE Trans. Image Proc.*, vol. 6, no. 12, pp. 1692–1697, Dec. 1997.