Improving Data Integrity with Randomness – A Compressive Sensing Approach

Zainul Charbiwala, Sadaf Zahedi, Younghun Kim,
Supriyo Chakraborty, Mani B. Srivastava
University of California, Los Angeles

Ting He
Chatschik Bisdikian
IBM T. J. Watson Research Center

I. Introduction

Data loss in wireless sensor systems is inevitable, either due to exogenous (such as transmission medium impediments) or endogenous (such as faulty sensors) causes. While there have been many attempts at coping with this issue, recent developments in the area of Compressive Sensing (CS) enable a new perspective. Since many natural signals are compressible, it is possible to employ CS, not only to reduce the effective sampling rate, but to improve the robustness of the system at a given Quality of Information (QoI). This is possible because reconstruction algorithms for compressively sampled signals are not hampered by the stochastic nature of wireless link disturbances and sensor malfunctions, which has traditionally plagued attempts at proactively handling the effects of these errors. In this paper, we show how reconstruction error remains unchanged despite extreme independent data losses by marginally increasing the average sampling rate. We also show that a simple re-ordering of samples prior to communication could enable successful reconstruction when losses display bursty behavior.

The problem we seek to address is acquiring an n-length vector $f \in \mathbb{R}^n$ at a sensor node such that it can be recovered accurately at a base station one or more wireless hops away. Realistically, data losses creep into the system owing to two inevitable circumstances – wireless link quality variations because of noise and interference and temporary sensor faults[1]. To cope with these issues, reactive schemes like retransmissions (end-to-end or hop-by-hop) have been popularly employed. Proactive schemes such as error correcting codes have also been used, though in a limited sense. In this short paper, we introduce how recent developments in the area of Compressive Sensing (CS) [2] enable a low-encoding complexity, proactive sensing approach that can easily be made robust to even extreme data losses. Utilizing the fact that CS strategies make inherent use of randomness within the sensing process, we surmise that data lost through the stochastic nature of an erasure channel is indistinguishable from an *a priori* lower sensing rate at the fusion center. We verify this conjecture empirically and show that it is sufficient to proactively increase the sampling rate in order to maintain reconstruction accuracy.

[1]We assume an out-of-band technique such as [1] that exploits spatial correlations to classify and discard erroneous data at the fusion center.

II. Compressive Sensing Fundamentals

A key assumption for applying CS is that the vector f is representable in some domain Ψ as a sparse or compressible vector, $x \in \mathbb{R}^n$. We can represent f equivalently then as $f = \Psi x$, where $\Psi \in \mathbb{C}^{n \times n}$ represents an orthonormal basis such as the inverse Fourier basis ($\psi_{\omega,k} = \frac{1}{\sqrt{n}} \exp(i 2\pi \omega k/n)$). The signal f is then acquired in its natural domain by projecting it through a sensing matrix $\Phi \in \mathbb{R}^{m \times n}$ to generate m measurements $y = \Phi f = \Phi \Psi x = Ax$. The questions that CS theory answers are: how can x be recovered from y, how many measurements m ($\ll n$) are required for accurate recovery and what sensing matrices Φ facilitate recovery. We summarize some key results from [3] and references therein.

The proposed reconstruction technique involves solving a constrained ℓ_1 minimization problem [3] as follows:

$$\hat{x} = \underset{\tilde{x}}{\text{argmin}} \, \|\tilde{x}\|_{\ell 1} \quad \text{s.t. } \; y = A\tilde{x} \tag{1}$$

In order to prove that the solution to (1) is exact, a notion termed the restricted isometry property (RIP) was introduced.

Definition 1: [3] For each integer $s = 1, 2, ...$, define the isometry constant δ_s of a matrix A as the smallest number such that

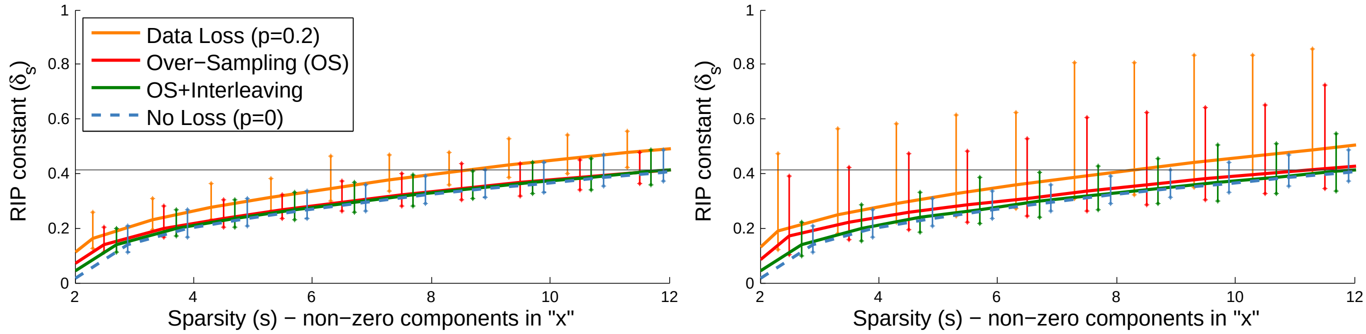
$$(1 - \delta_s) \|x\|^2 \leq \|Ax\|^2 \leq (1 + \delta_s) \|x\|^2 \tag{2}$$

holds for all s-sparse vectors x. A vector is said to be s-sparse if it has at most s non-zero entries.

Theorem 1: [3] Assume that $\delta_{2s} < \sqrt{2} - 1$ for some matrix A, then the solution $\hat{x}$ to (1) for an s-sparse vector x is exact.

It has been shown in [2] that the matrix A can be constructed randomly using an *i.i.d.* Gaussian r.v. such that $A_{ij} = \mathcal{N}(0, \frac{1}{n})$ or an equiprobable $\pm \frac{1}{\sqrt{n}}$ Bernoulli r.v. Using such matrices in low-power sensing devices, however, is difficult since implementing the sensing matrix $\Phi = A\Psi^{-1}$ involves sampling and buffering f and computing $y = \Phi f$ explicitly through complex floating point operations.

It was shown in [2] that A can also be constructed by randomly selecting the rows of a Fourier basis matrix, such as Ψ. This implies then, that Φ is essentially an $m \times n$ random sampling matrix constructed by selecting m rows independently and uniformly from an $n \times n$ identity matrix $\mathbf{I}_n$. This Φ is trivially implemented by pseudo-randomly sampling f, m times and communicating the stream of samples and their timestamps to the fusion center. Matrix Φ can then be recreated at the fusion center from the timestamps.



(a) Bernoulli Channel

(b) Gilbert-Elliott Channel ($p_{g\to b} = 0.0312$, $p_{b\to g} = 0.125$)

Fig. 1. Effect of data loss on RIP constant with average loss probability p. Also shown is the improvement in RIP constant by increasing rate to $m/(1-p)$ and shuffling samples prior to transmission. Error bars indicate the min-max across 1000 Monte-Carlo runs. The $\delta_{2s} < \sqrt{2}-1$ bound is included for reference.

III. Handling Data Losses

We model the erasure introduced by the transmission channel or sensor malfunction with an average loss probability, $p \in [0, 1]$. We initially assume an independent Bernoulli process so that the probability of any packet (or sample, assuming samples are transmitted individually) being dropped is equal to p. The question we would like to address is how the loss of $m \cdot p$ samples dropped in this way affects CS reconstruction. From Theorem 1, we understand that reconstruction accuracy depends on the RIP constant δ_{2s} of A. However, A has been modified by the channel so that $A' = \mathbf{C}A = \mathbf{C}\Phi\Psi$, where $\mathbf{C}$ is an $m \cdot (1-p) \times m$ matrix (constructed from $\mathbf{I}_m$) that enumerates which samples got delivered. To evaluate the extent of performance loss through the erasure channel, we thus rely upon quantifying $\delta_s(A')$.

Computing $\delta_s(A')$ exactly from Definition 1, however, is exhaustive because it is defined over all s-sparse vectors. We approximate it by evaluating the eigenvalues of the Grammian [4] over 10^3 random $s \times n$ submatrices. Increasing this number to 10^6 provides little improvement. The result from this calculation for 1000 randomly generated 256×1024 random sampling matrices (as described in §II) is shown in Figure 1. The dashed curve labeled "No Loss" indicates $\delta_s(A)$. The error bars illustrate the min-max values over all Φ about the mean and have been offset slightly to improve legibility. With loss probability $p = 0.2$, we see an increase in RIP constant, which implies that the sparsity for guaranteed ℓ_1 reconstruction drops (from about 6 to about 4 based on the bound $\delta_{2s} < \sqrt{2}-1$).

It can be shown that the probability distributions of time-stamps extracted from Φ and $\Phi' = \mathbf{C}\Phi$ are identical when $\mathbf{C}$ comes from an independent Bernoulli channel. This means that losses due to the channel are indistinguishable from an *a priori* reduced random sampling rate at the sensor node. This in turn means that, if the channel is not congested, increasing the sensing rate by a factor of $p/(1-p)$ will restore the delivery rate to m. The effect of this increase is substantiated in Figure 1(a) and establishes $\delta_s(A') \approx \delta_s(A)$ for the Bernoulli channel.

Realistic wireless channels, however, exhibit bursty characteristics [5], which may be modeled using a stationary discrete-time binary (good-bad) Markov process, popularly known as the Gilbert-Elliott (GE) channel [6]. To test the effect on $\delta_s(A')$ and hence CS reconstruction, we constructed a GE model with an expected loss burst of 8 samples and $p = 0.2$. Thus, the same number of samples are delivered to the fusion center, but with a modified time-stamp distribution. The effect of this change is immediately evident in the variance of δ_s in Figure 1(b) indicating that some Φ matrix choices will be particularly bad for a GE channel. While an increase in sensing rate improves the mean, the variance still remains high. This variance issue can be resolved by also applying randomized interleaving [7] prior to transmission, which results in a roughly uniform distribution of the sample losses. It can be shown that interleaving recovers the original time-stamp distribution (up to a bound) and Figure 1(b) illustrates this empirically. Note, however, that interleaving requires buffering y (but not f), which increases reconstruction latency.

In conclusion, we have explored the application of Compressive Sensing to handling data loss from erasure channels by viewing it as a low encoding cost, proactive, error correction scheme. We employed the RIP to illustrate for two channel models, that even extreme stochasticity in losses can be handled cheaply and effectively.

IV. Acknowledgments

This research was sponsored by the U.S. Army Research Laboratory and the U.K. Ministry of Defence and was accomplished under Agreement Number W911NF-06-3-0001. The views and conclusions contained in this document are those of the author(s) and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory, the U.S. Government, the U.K. Ministry of Defence or the U.K. Government. The U.S. and U.K. Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

References

[1] K. Ni and G. Pottie, "Bayesian selection of non-faulty sensors," in *IEEE International Symposium on Information Theory, ISIT*, 2007.

[2] E. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?," *Trans on Info Theory*, 2006.

[3] E. Candès, "The restricted isometry property and its implications for compressed sensing," *Comptes rendus-Mathématique*, 2008.

[4] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, 2008.

[5] K. Srinivasan, M. Kazandjieva, S. Agarwal, and P. Levis, "The β-factor: measuring wireless link burstiness," in *SenSys*, 2008.

[6] E. Elliott, "Estimates of error rates for codes on burst-noise channels," *Bell Syst. Tech. J*, 1963.

[7] M. Mushkin and I. Bar-David, "Capacity and coding for the Gilbert-Elliot channels," *IEEE Transactions on Information Theory*, 1989.