Sparse Signal Acquisition via Compressed Sensing and Principal Component Analysis

Imrich Andráš\textsuperscript{1}, Pavol Dolinský\textsuperscript{2}, Linus Michaeli\textsuperscript{3}, Ján Šaliga\textsuperscript{4}

Technical University of Košice, Faculty of Electrical Engineering and Informatics, Department of Electronics and Multimedia Telecommunications, Letná 9, 04200 Košice, Slovakia, imrich.andras@tuke.sk

This paper presents a way of acquiring a sparse signal by taking only a limited number of samples; sampling and compression are performed in one step by the analog to information conversion. The signal is recovered with minimal information loss from the reduced data record via compressed sensing reconstruction. Several methods of analog to information conversion are described with focus on numerical complexity and implementation in existing embedded devices. Two novel analog to information conversion methods are proposed, distinctive by their computational simplicity – direct subsampling and subsampling with integration. Proposed sensing methods are intended for and evaluated with real water parameter signals measured by a wireless sensor network. Compressed sensing proves to reduce the data transfer rate by >80 % with very little signal processing performed at the sensing side and no appreciable distortion of the reconstructed signal.

Keywords: Compressed sensing, analog-to-information conversion, sparse signal, signal compression.

1. INTRODUCTION

Compressed sensing (CS) [1] appeared in recent years as a promising method of sub-Nyquist sampling, allowing signal acquisition with fewer measurements than previously thought possible. With sparse signals, CS allows for much lower sampling frequencies and lower data transfer rates. The advantage over conventional compression is very little (if any) additional digital signal processing (DSP). The compression is performed directly by the sampling process and most of the overall computational demand is moved to the reconstruction side. This is convenient, for example, in wireless sensor networks (WSN), where available computation power is expensive at sensor nodes, but virtually unlimited at central hubs. Furthermore, certain methods of CS can be implemented on existing devices with no hardware changes required [2]. CS reconstruction is also inherently immune to data loss [3] and the transferred data are naturally encrypted [4]. Since its introduction, CS was tested with WSNs performing temperature [5], EM emission [6], and ECG [7] monitoring, video streaming [8], etc. Other applications of the CS theory include WSN decentralized data storage [9], accelerated MRI [10], or digital modulation recognition [11].

This work was motivated by issues with wireless sensor network (WSN) intended for water quality monitoring of the Danube river [12]. Individual sensor nodes measure multiple water parameter signals (WPS) such as temperature, salinity, dissolved oxygen, pH, etc. [13], [14], [15]. The sensor nodes use solar power for charging the integrated battery. Around 90 % of the WSN node power consumption is drawn by the transmitter [16]. This proved to be of particular concern during winter months, when only a limited amount of solar power is available and the battery performance is reduced due to low temperatures. Significant reduction in power consumption could be achieved by decreasing the amount of data that are being transmitted. The issue with applying conventional (e.g., transform) compression methods is that they require a substantial number of signal samples in order to be effective. With the current sampling period of 1 hour this would mean a very long processing timeframe. The additional computations would also increase the power consumption, counteracting the power savings gained at the transmitter. These issues can be circumvented by utilizing CS.

This work demonstrates that CS can be used for WPS measurement and efficiently implemented in an existing WSN. Two sampling methods previously not used with CS are proposed, comprising the main novelty of this paper. Section 2 provides an overview of general CS theory and related works. Investigated sampling methods and their properties are discussed in section 3. Experimental results are shown in section 4, where the performance of the proposed methods is evaluated. An optimal CS configuration of WPS monitoring sensor node is proposed at the end of section 4, followed by the conclusion.

DOI: 10.1515/msr-2018-0025
2. COMPRESSED SENSING

The main goal of CS is condensing a signal into a few samples containing all the information, which is referred to as analog-to-information conversion (AIC). This compressed information is transferred to the receiver, where the original signal is reconstructed. No information loss and exact reconstruction is theoretically possible [1].

In order for CS to be applicable, it is assumed that the input signal can be represented by a linear combination of known basis functions. Furthermore, the signal must be sparse, meaning that it consists of only a small number of basic functions. The signal composed of s basis functions is denoted as s-sparse. Let us define L basis vectors \( \Psi_l \in \mathbb{R}^{N \times 1}, 1 \leq l \leq L \) to represent any possible input signal components. The columns of the basis matrix \( \Psi \in \mathbb{R}^{N \times L} \) consist of these basis vectors. The input signal vector \( \mathbf{f} \in \mathbb{R}^{N \times 1} \) can then be described as

\[
\mathbf{f} = \Psi \mathbf{x}, \tag{1}
\]

where the s-sparse vector \( \mathbf{x} \in \mathbb{R}^{L \times 1} \) conducts linear combination [1].

A. Analog-to-Information Conversion

Conventional ADCs acquire signal samples at equidistant time instants, according to the Nyquist sampling theorem. With signals highly sparse at a certain domain, such sampling means great redundancy in acquired data, since the amount of information contained within a sparse signal is limited. AIC exploits this fact and only takes a limited number of samples, sufficient to represent the information content. The input signal \( \mathbf{f} \in \mathbb{R}^{N \times 1} \) is correlated with \( M < N \) measurement signals, represented by \( M \) rows of the measurement matrix \( \Phi \in \mathbb{R}^{M \times N} \). This performs the AIC, with resulting information signal vector

\[
\mathbf{y} = \Phi \mathbf{f} \in \mathbb{R}^{M \times 1}. \tag{2}
\]

There are several ways of performing the AIC, either in analog domain, or digitally after conventional sampling [17]. Analog domain implementations allow for overcoming the Nyquist limit of fast ADCs and are used mainly for high bandwidth signals. Digital domain AIC is a computationally efficient alternative to conventional compression methods. Direct digital implementation of (2) is preferable if possible, because it avoids problems associated with analog domain AIC non-idealities. Random modulation and pre-integration (RMPI), random sampling (RS), and random filtering are the most discussed in literature, but other methods may be used for specific applications.

B. Reconstruction

Let the input signal vector be s-sparse on basis \( \Psi \), and let bases \( \Psi \) and \( \Phi \) be incoherent (rows \( \phi_m \) of \( \Phi \) do not sparsely represent the columns \( \psi_l \) of \( \Psi \), and vice versa). If \( s < M \ll N \), the input signal can be reconstructed [18]. An approximation of minimal \( M \) needed is given by

\[
M_{\text{min}} = \mu \log_2(N) \tag{3}
\]

assuming zero or negligible noise. \( \mu = 1 \) for incoherent bases, but has to be increased to \( \mu = L / N \) if \( N < L \). By inserting (1) into (2) the information signal becomes

\[
\mathbf{y} = \Phi \Psi \mathbf{x}, \tag{4}
\]

with the reconstruction matrix

\[
\mathbf{A} = \Phi \Psi \in \mathbb{R}^{M \times L}. \tag{5}
\]

At the receiving end of the CS framework, the information signal \( \mathbf{y} \), and both bases \( \Psi \) and \( \Phi \) are known. The only unknown required for reconstruction of (1) is \( \mathbf{x} \). Since \( \mathbf{A} \) is a rectangular matrix, it cannot be simply inverted and an undetermined system of \( M \) equations and \( L \) unknowns is to be solved. Here the importance of sparsity is shown, because based on this requirement a unique solution can be found. Out of all the possible solutions, the right solution is the one that is the most sparse [1].

The sparsity \( s \) of vector \( \mathbf{x} \) can be characterized by its \( \ell_1 \) norm defined as

\[
\ell_1 : ||\mathbf{x}||_1 = \sum_{i=1}^{L} |x_i| . \tag{6}
\]

Sparsity is an additional information used to devise the measurement process \( \Phi \), and forming a convex optimization problem

\[
\min ||\mathbf{x}||_1 \text{ subject to } \mathbf{A} \mathbf{x} = \mathbf{y} . \tag{7}
\]

This guarantees a correct optimal solution, if such exists. The original signal estimate can be found more efficiently as

\[
\hat{\mathbf{f}} = \Psi (A^T A)^{-1} A^T \mathbf{y} , \tag{8}
\]

where \((*)^{-1}\) denotes the pseudoinverse matrix of \((*)\). (8) incorporates solving an \( \ell_2 \) minimization problem, but in practical scenarios (7) and (8) yield the same solutions within a noise related uncertainty [19].

The issue with CS is determining the base matrices \( \Psi \) and \( \Phi \). AIC is performed in such a way that the resulting measurement matrix \( \Phi \) shows random attributes, ensuring bases incoherency. The basis matrix \( \Psi \) can be based on signal model if the sparse domain is known, e.g. frequency or wavelet. For other signals there is not a specific domain for this purpose, and a suitable dictionary must be learned. Correctly constructed set \( \Psi \) of basic functions ensures high compression ratio and low reconstruction error. For WPS basis extraction the principal component analysis (PCA) [20]
is proposed. PCA is convenient with signals that would require complicated models or their analytical model is unknown. The PCA is particularly prominent in biomedical applications [21]-[24], and has been utilized for basis extraction in various CS applications [3], [7], [10], [25]. The CS framework with both sensing and decoding side is summarized in Fig. 1.

3. INVESTIGATED MEASUREMENT METHODS AND MATRICES

A. Random modulation and pre-integration

The functional diagram of analog RMPI is shown in Fig. 2. Input signal \( f(t) \) is multiplied by sequence \( p(t) \) of \( \pm 1 \)'s from the random sequence generator (RSG). Output signal of the mixer is being integrated and the integrator output is periodically sampled by a conventional ADC. After each sample is taken, the integrator is reset. Sampling frequency of the ADC is several times lower than the input signal Nyquist frequency. The only component operating at Nyquist rate is the mixer, any further DSP is performed after compression at a lower rate. The compression itself is performed in the analog domain without DSP. This allows for real-time energy-efficient compression of very high bandwidth signals.

B. Random sampling

The RSAIC architecture is among the easiest to implement, yet it is rarely studied. RS can be achieved either by triggering an ADC according to RSG (Fig. 5.), or simply by discarding samples of a free-running ADC [30]. Both of these approaches are suitable for high bandwidth signals. There is no special analog signal processing and the compression is performed with virtually no DSP. The RS may also be implemented via firmware upgrade on existing devices. Despite all of these advantages the RMPI is by far more common in literature. The reason is probably the limited versatility of RS. RMPI yields low sampling rate, but the input information present between sampling instants is preserved via integration. With RS all of this information is discarded, which practically limits the RS AIC to frequency sparse signals with little expected variance [30].

After the input signal is sampled by a Nyquist ADC into its discrete representation \( f_m \), the signal is multiplied by a random sequence \( p_n \) of \( \pm 1 \)'s and summed. The summator is being periodically reset after outputting the accumulated value. All of this DSP runs at Nyquist rate in real time, so the digital RMPI is more suitable for slow signals. Elements \( \Phi_{mn}, 1 \leq m \leq M, 1 \leq n \leq N \) of the measurement matrix \( \Phi \in \mathbb{R}^{M \times N} \) representing an ideal or digital RMPI can be described as

\[
\Phi_{mn} = \left\{ \begin{array}{ll}
p_n, & (m-1)C < n \leq mC \\
0, & \text{otherwise}
\end{array} \right.
\]

where

\[
C = \frac{N}{M} \in \mathbb{N}
\]

is the compression ratio.

The main disadvantage of analog RMPI is the need for specialized non-ideal analog circuitry. There may be a discrepancy between the expected measurement process \( \Phi \) and the actual RMPI behavior [26] resulting in improper reconstruction (8). Accurate modeling of physical circuitry [27] and stochastic phenomena mitigation [28], [29] may be a problem. If the acquired signal is not of excessively high bandwidth, RMPI may be implemented digitally as is shown in Fig. 3. The advantage is that the reconstruction error can be significantly lower, since the only non-ideal component is the input ADC. With slow signals like WPS, digital RMPI could be implemented on an existing device solely via firmware upgrade.
A typical WPS (Fig.4.) contains a large slowly changing DC component, a faster AC component with relatively small amplitude and a small amount of noise [32]. The AC component is a distorted sine wave with basic period of approximately 1 day and a few higher harmonics. The WPS thus carries only limited information content and can be classified as frequency sparse [33], meeting the criteria for RS application.

From all the Nyquist samples of input signal (1) a subset of \( M \) samples is selected. \( \Xi \) contains indices of selected samples

\[
\Xi = \{ \xi_1, \xi_2, \ldots, \xi_M \}. \tag{11}
\]

Entries of \( \Xi \) are random numbers, following the rules

\[
\forall m \in \mathbb{N}, 1 \leq m \leq M : \xi_m \in \mathbb{N}, 1 \leq \xi_m \leq N, \tag{12}
\]

\[
\forall i, j \in \mathbb{N}, 1 \leq i < j \leq M : \xi_i < \xi_j. \tag{13}
\]

Elements of measurement matrix representing RS AIC can be obtained using (13) as

\[
\varphi_{mn} = \begin{cases} 1, n = \xi_m \\ 0 \end{cases}, \tag{14}
\]

and the information signal vector \( y \) (2) elements become

\[
y_n = f_{\xi_n}. \tag{15}
\]

C. Direct subsampling

Coherent sub-Nyquist sampling without antialiasing filter is not yet covered by CS related literature for an obvious reason. In general, proper reconstruction with such AIC is not possible because of aliasing. However, the basic functions of WPS obtained by PCA are not strictly periodic functions. Elements of \( \Psi \) are general waveforms, as is shown in Fig.6. Only some of the basic functions resemble sinewaves with a period of approximately 1 day. These have a common composition of higher harmonics and interharmonics already incorporated. As such, higher harmonics are not subject to irreversible aliasing. The original signal can be reconstructed provided that the basis harmonic component is sampled at Nyquist rate. Irreversible input information loss occurs if the period of a specific higher harmonic component is exactly equal to the sampling period or its integer multiple.

The presented study considers coherent subsampling via decimation of a Nyquist record, hereafter referred to as direct subsampling (DS). Corresponding measurement matrix consists of elements

\[
\varphi_{mn} = \begin{cases} 1, n = mC \\ 0 \end{cases}, \tag{16}
\]

resulting in information signal (2)

\[
y_n = f_{ac}. \tag{17}
\]

DS is proposed as application-specific AIC for WPS because of its simplicity. The implementation on existing devices is even easier than RS, since no RSG is needed. If the existing Nyquist ADC has a fixed clock, samples may be discarded. Better energy efficiency would be achieved if the ADC clock could be reprogrammed to a sub-Nyquist rate, eliminating the need to discard samples.

D. Subsampling with integration

Subsampling with integration (SI) is proposed as application-specific simplified alternative to RMPI. The modulator can be omitted because WPS represents absolute values of chemical properties and is non-alternating. Input signal is integrated over the sampling period, as is shown in Fig.7. The integrator is being periodically reset after the integrated signal is sampled.

---

**Fig.5. RS AIC.**

**Fig.6. The most significant WPS basis components obtained by PCA.**

**Fig.7. Analog SI AIC.**
Digital implementation of SI AIC as shown in Fig.8. may be preferable due to the reasons discussed in section 3.4. The measurement matrix changes from (11) to

\[
\varphi_{mn} = \begin{cases} 
1, & (m-1)/C < n \leq mC \\
0 &
\end{cases}.
\]  

(18)

Fig.8. Digital SI AIC.

4. RESULTS

Digital implementations of AIC methods described in section 3 were tested using WPS data obtained during the pilot operation of the WSN [13]. All of the tested methods can be implemented as a software add-on on existing sensor nodes. Testing was performed by means of simulations, but it sufficiently represents implementation on the physical WSN.

\[ L' = 200 \text{ signal segments with length of } N = 120 \text{ (5 days) were randomly chosen from the WPS database and normalized to 1. This training set was used for extraction of the basis matrix } \Psi \in \mathbb{R}^{N \times L} \text{ by the PCA. Another set of 200 randomly chosen WPS segments was used for testing. The signal-to-deviation ratio } \text{SDR} \text{ was used for evaluation. SDRs of each of the test signals were averaged. With RS and RMPI, } 10^3 \text{ measurements of each test signal were performed with a new measurement matrix generated every time.}

4A. Number of principal components used for reconstruction

Out of all of the \( L' \) extracted basis components, only a limited number has appreciable energy. The rest of the components can be left out of \( \Psi \) with little impact on SDR. Achievable SDR with respect to the number of components in model (1) was found by using an identity measurement matrix \( \Phi = I_N \) (no resampling) and performing reconstruction (8) with varied \( L \).

\[ \text{SDR} = 10 \log \left( \frac{\sum_{n=1}^{N} \hat{f}_n^2}{\sum_{n=1}^{N} (f_n - \hat{f}_n)^2} \right), \]  

(19)

where \( \hat{f}_n \) is the reconstructed signal, was used for evaluation. SDRs of each of the test signals were averaged. With RS and RMPI, 10^3 measurements of each test signal were performed with a new measurement matrix generated every time.

4B. Required sampling rate

It can be seen that PER is never equal to 1, since some of the test signals are noisy. Exact reconstruction cannot be expected even with noiseless signals - no training set can be large enough to account for every possible signal variation. With available training set the principal component dictionary \( \Psi \) saturates around \( L = 100 \). Increasing the number of principal components further brings no appreciable improvement.

Based on results shown in Fig.9. and Fig.10., the number of principal components used was chosen to be \( L = 50 \), which applies to all further evaluations. \( L < 200 \) was chosen on purpose in order to simulate an unsaturated principal component dictionary expected in practical applications. Smaller \( L \) also results in faster reconstruction due to smaller size of matrices while having no impact on AIC.

Fig.9. Achievable SDR vs. number of principal components used.

Fig.10. PER vs. number of principal components used.
Any analytical derivation of $M_{\text{min}}$ is probabilistic. Choosing $M < M_{\text{min}}$ does not mean that the reconstruction is impossible, but rather that there is a high risk of picking particular signal values that result in unsuccessful signal recovery. The most reliable way of choosing an appropriate $M$ for certain application is an experiment, where $M$ is varied and reconstruction error is observed. Unknown $s$ on a particular $\Psi$ base can be determined retroactively by observing $x$ in successful reconstructions. SDR as a function of compression ratio (12) for investigated AIC methods is shown in Fig.11.

The performance of all AIC methods is similar within approximately 6dB SDR range. SI achieves the highest SDR at both small and high compression ratios, while RS and RMPI perform slightly better in the intermediate range. DS may be preferable due to its simplicity, despite worse performance at high $C$. The final choice of AIC method implemented in WSN would strongly depend on the required SDR.

C. Resistance to quantization

High ADC resolution, denoted in number of bits (NOB), is easily achievable with signals as slow as WPS. However, the total amount of data payload that is to be transferred is proportional to NOB. Lower NOB is beneficial for overall sensor node power consumption, and can be achieved by rescaling existing high-resolution samples. Influence of quantization on SDR was tested using a simulated ideal unipolar ADC with a range of 2. Results for $C = 6$ and varying resolution are shown in Fig.12.

RS and DS show slightly greater robustness against quantization than RMPI and SI. The highest SDR is achievable by SI, which is consistent with results from the previous section. Increasing NOB above 12 yields virtually no improvement in SDR with all evaluated AIC methods.

D. Requirements on the reconstruction error

Results presented above show that the SDR is dependent on compression ratio and quantization noise. The specific combination of $C$ and NOB defines the power consumption, which is in trade-off with the SDR. How high the SDR should be, and whether the introduced misinformation is of any concern, remains to be decided by the end user. Experiments show that SDR above 30 dB is enough to capture the slowly changing DC component of any available WPS. In order to preserve the smallest AC components present in the database, SDR should remain above 50 dB.

The authors propose to use SI AIC with $C = 6$ and NOB = 12. The power consumption with such configuration would be reduced by approximately 80% while maintaining SDR ≥ 50dB. An example of input and reconstructed signal obtained with this configuration and $SDR = 51.43$dB is shown in Fig.13. Further reduction of power consumption is possible at the cost of lower SDR.

5. CONCLUSION

CS was originally promoted as near-optimal and lossless on orthonormal bases. However, practical WPS results show that the reconstruction is not exact if appreciable compression ratio is achieved. This was also concluded in some of the cited works on other CS applications, e.g. ECG [7]. Considering the reasonable compression at minimal computation effort, CS is still a reasonable choice for WPS monitoring WSN. RS and RMPI methods of AIC have been previously used in WSNs and proved to be efficient also with WPS. The presented DS and SI methods pose even lower computational demand on the sensor node. The SI method at the same time outperforms previously presented AIC methods in achievable SDR. AIC configuration that would reduce the sensor node power consumption by approximately 80% was proposed and remains to be implemented on existing WSN nodes.
ACKNOWLEDGMENT
The work is a part of the project supported by the Science Grant Agency of the Slovak Republic (No. 1/0722/18).

REFERENCES


Received April 13, 2018
Accepted September 03, 2018