IoT Data Compression: Sensor-agnostic Approach

Arijit Ukil*, Soma Bandyopadhyay* and Arpan Pal*

*Innovation Labs

Tata Consultancy Services
Kolkata, India
{arijit.ukil, soma.bandyopadhyay, arpan.pal}@tcs.com

Abstract: Management of bulk sensor data is one of the challenging problems in the development of Internet of Things (IoT) applications. High volume of sensor data induces for optimal implementation of appropriate sensor data compression technique to deal with the problem of energy-efficient transmission, storage space optimization for tiny sensor devices, and cost-effective sensor analytics. The compression performance to realize significant gain in processing high volume sensor data cannot be attained by conventional lossy compression methods, which are less likely to exploit the intrinsic unique contextual characteristics of sensor data. In this paper, we propose SensCompr, a dynamic lossy compression method specific for sensor datasets and it is easily realizable with standard compression methods. Senscompr leverages robust statistical and information theoretic techniques and does not require specific physical modeling. It is an information-centric approach that exhaustively analyzes the inherent properties of sensor data for extracting the embedded useful information content and accordingly adapts the parameters of compression scheme to maximize compression gain while optimizing information loss. Senscompr is successfully applied to compress large sets of heterogeneous real sensor datasets like ECG, EEG, smart meter, accelerometer. To the best of our knowledge, for the first time 'sensor information content'-centric dynamic compression technique is proposed and implemented particularly for IoT-applications and this method is independent to sensor data types.

1. Introduction

Sensors are now pervasive and ubiquity of sensors makes IoT applications useful as well as exotic. IoT applications pervade inside human body, inside luxury (even low cost) cars to all the aspects of human life. Such huge amount of sensors generate high amount of data. Storage, transmission and processing of such high volume data pose potentially destructive performance and scalability risk; tiny sensors require large buffer, transmission of high volume data results in reduction of battery lifetime, higher processing power is needed by the tiny relay node for in-network processing. In order to harvest IoT applications like smart energy management, elderly monitoring, e-health care, data volume is to be reduced such that utility is optimized while maximizing the compression. Extensive research works already proposed various data compression methods particularly for sensor data [1, 2]. Specifically, model-based approaches that compress sensor data by establishing approximation methods outperform traditional data compression methods [1]. Piece-wise linear approximation provides an approximation model with separate linear functions per data block and seems to be a widely used lossy compression technique. Among various model-based compression technique, we consider the Chebyshev approximation due to its high capability of accurate reconstruction of

1068-0314/15 $31.00 © 2015 IEEE
DOI 10.1109/DCC.2015.66

large number of heterogeneous sensor data sets [3]. It is a nonlinear approximation model. However, Chebyshev compression works on traditional lossy compression philosophy of reducing redundant information as shown in figure 1. In fact, more compression gain → more information loss→ arbitrary information compression → may be catastrophic for sensor analytics. In order to optimize sensor data compression, human-contextual property inherent to sensor datasets due to its close proximity of human world is to be exploited. Sensor data readings provide insight of different important events. Unusual pattern in ECG record may be indicative of irregular heart function like arrhythmia or high body temperature is a sure indicative of fever or other disease. For mining and knowledge discovery such inference is of utmost importance.

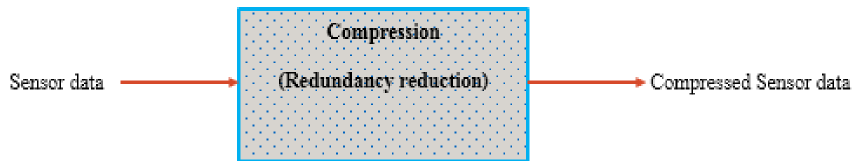


Figure 1. Traditional data compression

We propose SensCompr that extracts the useful or interesting information from sensor data and adapts the parameters like threshold selection, block-size estimation for Chebyshev compression to yield maximum compression gain while sacrificing insignificant information loss. This philosophy is also supported by classical information theory. Actually, Chebyshev compression due to its fixed block size parameter violates the maximum error bound feature of model-based compression [1]. . In fact, fixed block size is a demerit of Chebyshev compression. We overcome this problem by introducing dynamic block size Chebyshev. We focus on practical deployment of sensor data compression, where data compression can be seamlessly achieved during capture phase (sensor device), in-network processing phase (edge gateway) and sensor data storage and analytics phase (IoT platform/ cloud) independent of the sensor data types. Obviously such horizontal model incurs the cost of generalization but avoids costly implementation and maintenance of number of sensor-specific compression methods. Instead of physical modeling (which is sensor specific), our method is based on signal property modeling typically through statistical analysis. This approach is indeed of immense benefit for edge gateway and IoT platform that handle heterogeneous sensor datasets. SensCompr has the added advantage of being a unsupervised learning method that again eliminates training and setup costs. We show that our approach incurs significant compression gain while the loss at the reconstruction does not invite practical feature extraction error. We consider ECG signal, a critical human health indicator to demonstrate this claim.

This paper is organized as follows. In Section 2, we describe the sensor information extraction scheme. We briefly introduce Chebyshev approximation and describe our proposed scheme SensCompr in Section 3. Results are shown in Section 4. In Section 5, we discuss extension of SensCompr for quasi-periodic signals like ECG and validates that proposed modeling is as good as physical modeling in terms of retaining practical feature extraction is concerned. Finally, we conclude in Section 6.

2. Sensor Data Importance Score Evaluation: The Phenomena of Interesting Events

Sensor data contains significant information about physical world like one's location information through GPS (results tracking, location based services)), health information through wearable sensors (results health insurance scheme revocation), energy consumption pattern though smart meter (results in detection of in-house activities) and many more. However, every event derived from smart meter data is not private and does not invoke interest to the malicious agents. Particularly, when sensor data reading from an individual is normal with respect to its usual activities or expected activities, knowledge-discovery agents or analytics engines may not be interested, as they cannot fetch extra information and they certainly look for anomalous sensor data to identify a person from crowd or to capture the misalignment with the trend (e.g. new premium is charged if anomalous health condition or different life style is detected). Continuing this philosophy, we apply robust outlier detection method for the detection of important information-bearing points to capture the unusual pattern or anomalous events that is part of the more useful information particularly in sensor data analytics. Next, we quantify the amount of useful information as sensor importance score. Analogically, MPEG audio coding like layer III coding exploits the psychoacuastic model of 'useful information' conceived by human auditory system. Instead of conventional compression wisdom of "how well a random variable can be represented in bits under rate-distortion constraint", we conceive the notion 'does all the random variables are at all to be represented". Our approach is to maximize the (contextual) information content between the (important) original sample space to its rate-distorted representation.

- ### Extraction of Sensor Importance Score

In order to extract the interesting events (events that are out of ordinary), robust outlier detection is employed, which requires masking effect and swamping effect to be minimized. Masking effect causes one or more outlier points to get undetected, while normal points are depicted as outlier by swamping effect. Current literature is mainly focused to minimize one of the false alarms. In order to fairly accommodate heterogeneous (sensor) datasets, we need to minimize both the effects according to the dataset statistical properties [5]. Through kurtosis measurement (κ), we are able to estimate the spread of probability distribution $\mathcal{D}$ of sensor dataset $\varepsilon(t)$, which helps us to apply appropriate outlier detection technique, where, $\kappa(\varepsilon) = \frac{E\left[\left(\varepsilon(t) - \overline{\varepsilon(t)}\right)^4\right]}{\sigma^4} - 3$. Formally, we define our region of interest $\mathbb{S}$ of ε with underlying distribution $\mathcal{D}$ for outlier-detection function $\mathcal{A}$ as [7, 8]: $\inf_{\varepsilon} \mathcal{A}(\varepsilon, \mathcal{D}) = 0, \quad \sup_{\varepsilon} \mathcal{A}(\varepsilon, \mathcal{D}) = 1$.

Case I: Criteria for Masking effect Minimization

Suppose $\mathcal{A}$ fails to identify ρ_M number of outliers in $out(\mathbb{S}, \mathcal{D})$ and with α_M as the smallest value of α that confirms existence of a constellation of missed β observations in $out(\mathbb{S}, \mathcal{D}, \beta)$, i.e.

$$\alpha_M\left(\varepsilon, \rho_{M,}\ \beta, \mathcal{D}\right) = \inf\left\{\begin{matrix} \alpha > 0\text{: } there\ exists\ \beta\ outliers\ in\ \varepsilon\ identified\ by \\ out(\mathbb{S}, \mathcal{D})\ missed\ in\ \mathcal{A} \end{matrix}\right\}$$

Case II: Criteria for Swamping effect Minimization

Conversely,

$$\alpha_S(\varepsilon, \rho_{S,}\ \beta, \mathcal{D}) = \inf\left\{\begin{matrix} \alpha > 0\text{: } there\ exists\ \beta\ outliers\ in\ \varepsilon\ not\ identified\ by \\ out(\mathbb{S}, \mathcal{D})\ included\ in\ \mathcal{A} \end{matrix}\right\}$$

Accordingly, it can be computed that the masking breakdown point ξ_M that when added on regular observations of ε causes $\mathcal{A}$ to breakdown. For Case I: $\xi_M(\varepsilon, \rho_{M,}\ \beta) = \frac{\gamma_M}{(\gamma_M + |\varepsilon|)}$, where $\gamma_M = min\{\gamma\text{: } \alpha_M(\varepsilon, \rho_{M,}\ \beta, \mathcal{D}) = 0\}$. Similarly, swamping breakdown point $\xi_S$in Case II is defined as: $\xi_S(\varepsilon, \rho_{S,}\ \beta) = \frac{\gamma_S}{(\gamma_S + |\varepsilon|)}$, where, $\gamma_S = min\{\gamma\text{: } \alpha_M(\varepsilon, \rho_{S,}\ \beta, \mathcal{D}) = 0\}$. In order to minimize anomaly detection error ξ_M, ξ_S are to be optimized and $\xi_M(\varepsilon, \rho_{M,}\ \beta) = \xi_S(\varepsilon, \rho_{S,}\ \beta) = 0.5$ is the optimum result.

- Minimizing Masking (False Negative) Effect:

  For minimizing masking effect, masking breakdown point ξ_M is to be close to 0.5 as described above. It is shown in [8], that for a large class of identifier Hampel identifier, a nonlinear local outlier detector based on Median Absolute Deviation (MAD) scale estimation provides outlier detection with $\xi_M \to 0.5$, particularly when $\mathcal{D}$ can be approximated to random observation, which we interpret as $\kappa(\varepsilon) < 3$. Specifically, $\mathcal{A}$ for $\kappa(\varepsilon) < 3$ is derived as follows:

  $MAD(\varepsilon_N) = median(|\varepsilon_1 - median(\varepsilon_N)|,\ \dots,\ |\varepsilon_N - median(\varepsilon_N)|), |\varepsilon| = N.$

  An observation ε_n is detected as outlier when: $|\varepsilon_n - median\ (\varepsilon_N)| > \Omega(N, \Lambda)MAD(\varepsilon_N)$, where Ω is defined in [7] for particular type-I error Λ [8].

- Minimizing Swamping (False Positive) Effect:

  When $\xi_S \to 0.5$, $\kappa(\varepsilon)\ \geq 3$, we consider Rosner filter for outlier detection. Rosner filter a backward selection method-based outward outlier testing using generalized (Extreme Studentized Deviate) ESD test is used to detect one or more outliers in a univariate data set [5, 9]. Given the upper bound, φ, it performs q separate test and compute $Q_i = \frac{max_i(\varepsilon_i - \overline{\varepsilon_N})}{s.d(\varepsilon_N)}$. Observations that maximizes $(\varepsilon_i - \overline{\varepsilon_N})$ are removed and repeat until q observations have been removed. This results in the *q* test statistics *Q1, Q2, ..., Qq*. Accordingly, following q critical values are computed as: $\eta_i = \frac{(n-i)t_{p,n-i-1}}{\left(\left(n-i-1+t_{p,n-i-1}^2\right)(n-i+1)\right)^{1/2}}$, i=1, 2, .., q

  $t_{p,\nu}$ is the 100p percentage point from the student's t distribution with ν degrees of freedom and $p = 1 - \frac{\theta}{2(n-i+1)}$. Outlier points in ε are the largest i such that $Q_i > \eta_i$. In fact, $\xi_S \to \xi_S\,|_{optimum}$ when $\xi \to$*student's t distribution.*

- Deriving Sensor Importance Score:

  Consider ν be the useful information part of ε and λ be the normal (uninteresting) part; $\varepsilon = \nu \cup \lambda$. Our idea is given $\mathcal{S}$, i.e. the information leakage transfer function $\Upsilon_{\mathcal{S},\nu}$: $\mathcal{S} \to \nu$ and $\phi_M = \frac{\sum_{i=1}^{|\nu|} Pr(\nu_{i.}) \log_2 \frac{1}{Pr(\nu_{i.})}}{\sum_{i=1}^{|S|} Pr(S_{i.}) \log_2 \frac{1}{Pr(S_{i.})}} = \frac{H(\nu)}{H(S)}$ [10]. In order to enhance the measurement accuracy, statistical compensation due to statistical relation between

ε and ν is computed using two-sample Kolmogorov-Smirnov (KS) test [10]. We propose L1-Wasserstein metric between $\varepsilon, \nu(\mathcal{w}_{\varepsilon,\nu})$ to estimate statistical misfit or compensation $\phi_S = \mathcal{w}_{\varepsilon,\nu}$. Logically, sensor importance score$(\phi_\varepsilon) = \phi_M \wedge \phi_S$. Algebraically, $\phi_\varepsilon = \phi_M \times \phi_S$. With $\phi_\varepsilon[0,5]$, we scale ϕ_ε as:

$$\phi_\varepsilon \longmapsto [\phi_\varepsilon \times 5]: \phi_\varepsilon = [1,5], \phi_{MAX} = 5 \tag{1}$$

with high magnitude of ϕ_ε signifies more useful information in ε.

3. SensCompr: Our Proposed Dynamic Sensor Data Compression

We follow the intuition from classical compression concept to retain the useful information as much as possible and abandon the non-sensitive part of the information. In fig. 2, we depict the functional architecture of SensCompr. We specifically optimize the two adaptable parameters of Chebyshev compression, viz. threshold value and block size. The adaptation module consists of: 1. Block size Optimization and 2. Threshold adaptation.

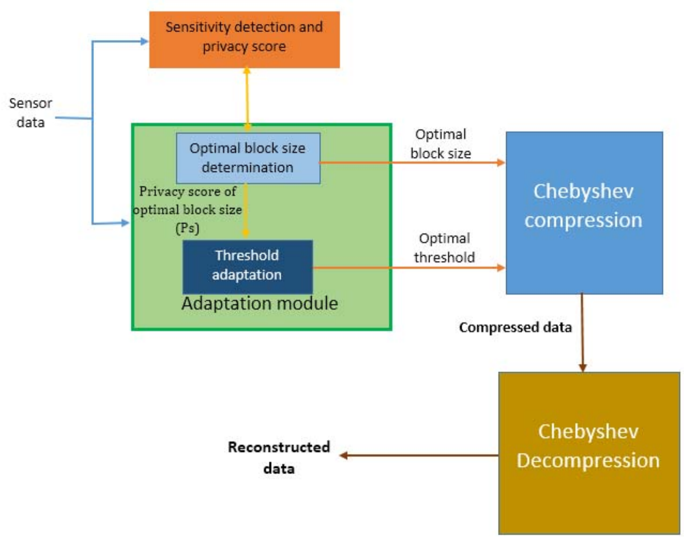


Figure 2.Functional architecture of ***SensCompr***, showing different functional modules and schematic data flow along with parameter optimization.

▪ Chebyshev Compression: Brief Overview

Compressed sensing is the recent trend in sensor data compression research, which is optimized for sensor signal that can be represented by sparse matrix [11]. But, many real-life sensor signals like smart meter data, accelerometer are not sparse. So, we consider Chebyshev compression to ensure better compression performance for dense matrix [1]. For a (fixed) block size *B* consisting of *N* elements, compressed data ε_δ is represented as linear combination of Chebyshev polynomials [1]: $\varphi(i) = \sum_{i=0}^{i=N} \alpha_i . \eta_\vartheta(i)$, where: $\vartheta = \left(i - \frac{N+1}{2}\right) . \frac{2}{N-1}$, normalized to [-1, 1], simply, cosine look-up table, α_i is the Chebyshev polynomial co-efficient at degree i. For a defined threshold Γ [12]:

$$\varepsilon_\delta = \begin{cases} \varphi(i), & \varphi(i) \geq \Gamma \\ 0, & else \end{cases} \quad (2)$$

Quantization (digitization) is done for non-zero ε_δ for storage, transmission purpose.

- **Optimization of Compression Parameter: Block Size Adaptation**

We propose following algorithm to find the optimal block size B_o, which is dynamically adapted based on useful information differential ($\Delta\phi_\varepsilon$) as:

$$B_o = \begin{cases} recursion(B_{init} = 2.B_{init}), & \Delta\phi_\varepsilon \neq 0 \\ & B_{init} else \end{cases} \quad (3)$$

Where, B_{init} is the initial block size and $\Delta\phi_\varepsilon = \left|\phi_{|B|} - \phi_{|B|+1}\right|$. Whenever, differential useful information is non-zero for different block sizes, we consider larger block size. It is known that sensor events are highly correlated in some domain (temporal or spatial or spatio-temporal), i.e. $H(\chi^{t+1}|\chi^t) \ll H(\chi^{t+1})$, where $H(.)$is the information entropy andχ is the random variable corresponds to sensor data. Consider without loss of generality, the probability of a block (n) of data having different useful information score with adjacent block be ω_n, $n = 1,2,..,N$ and ω_n are in i.i.d.$\sim p(\omega)$: $\left|\frac{1}{n} log_2 p(\omega_1, \omega_2, \ldots \omega_n)\right| \rightarrow H(\omega)$.

- **Optimization of Compression Parameter: Adaptation of Threshold**

Threshold acts as a clipping parameter as described in (1) to maximize the compression gain factor, which also negatively impacts on the information retention property of reconstructed signal. Optimal determination of threshold is required for compression gain- information loss trade-off. Let compression code $\varphi: \Theta \rightarrow \cup_{\pi \geq 1} \{0,1\}^\pi$ assigns each data points $\varepsilon\ (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N) \in \Theta\ (\Theta = \{0,1\}^n)$, a finite sequence of 1s and 0s to create the (Chebyshev) code word $\varphi(\varepsilon)$. The condition for decompressibility of $\varphi: \sum_{\varepsilon \in \Theta} \mathrm{P}(\varepsilon)|\varphi(\varepsilon)| \geq -\sum_{\varepsilon \in \Theta} \mathrm{P}(\varepsilon) log_2 \mathrm{P}(\varepsilon)$

Where, $\mathrm{P}(\varepsilon)$ is the probability distribution on the compression space Θ. From theory of large deviation, the optimality condition is :

$\sum_{\varepsilon \in \Theta} \mathrm{P}(\varepsilon)|\varphi(\varepsilon)| = -\sum_{\varepsilon \in \Theta} \mathrm{P}(\varepsilon) log_2 \mathrm{P}(\varepsilon)$, when $\mathrm{P}(\varepsilon) = 2^{-NH(\varepsilon)} \approx 2^{-|\nu|H(\nu)}$, νbe the data samples of the assigned block. Large magnitude of Γ results the decompressed sequence would have type outside the set ε, i.e. the probability of information loss $(P_e) \rightarrow 1$. Let's consider threshold (Γ): $\Gamma = c.2^{(\phi_{MAX} - \phi_{\varepsilon \in B_o})}$(4); $log_2\Gamma = -c' log_2(\phi_{\varepsilon \in B_o})$, ϕ is the importance information score (3), which follows KL divergence asymptotically when $\Gamma = c.2^{(\phi_{MAX} - \phi_{\varepsilon \in B_o})}$ is the optimal choice as shown in fig. 3. However, dispersion (Fano) factor needs to be compensated to satisfy the condition $P_e \rightarrow 0$. We consider $c = ceil(\mathcal{F})$, where $\mathcal{F}$ (Fano factor) = $\frac{\sigma^2}{\bar{\varepsilon}}$. The dynamic compression algorithm for SensCompr is captured below:

1. Find Fano factor $\mathcal{F}$ for the optimal block of data $(\varepsilon \in B_o)$ and $c = \mathrm{ceil}(\mathcal{F})$, B_o is computed as per (3).
2. Find importance score of $\phi_{\varepsilon \in B_o}$(1) with optimal block size (B_o).

3. Find optimal threshold (Γ_o) for optimal block of data $(\varepsilon \in B_o)$ is derived as: $c.2^{(\phi_{MAX}-\phi_{\varepsilon\in B_o})}$ (4).
4. Do Chebyshev compression with Γ_o and B_o for $(\varepsilon \in B_o)$ as per (2) and go to step.1 until all the sample data points are covered.

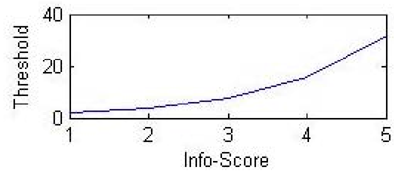


Figure 3. Threshold value as function of 'useful information score'

4. SensCompr Results

Firstly, we depict the performance of SensCompr by experimenting with publicly available sensor datasets like REDD [13], BLUED [14], Physionet [15]. We consider smart meter datasets for initial experimentations due to its more random-like samples. Five independent household smart meter data is chosen for first set of experimentation. Fig. 4 shows considerable compression gain by SensCompr_1 (without block size adaptation, considering fixed block size ($B_o = 512$) with optimal threshold Γ_o for fixed block size) and lesser loss factor. We observe that block size optimization along with optimal threshold (Senscompr) yields more gain yet lesser loss factor.

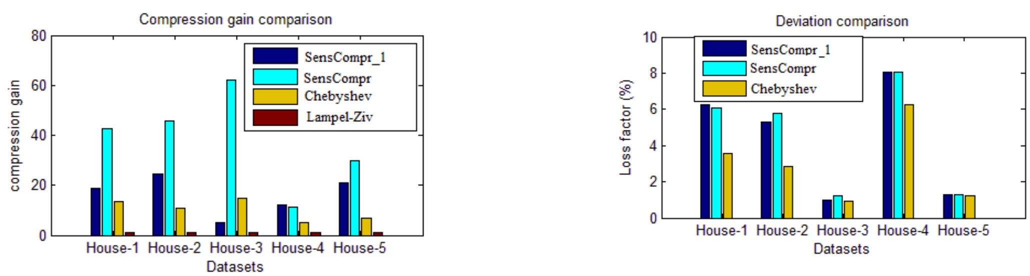


Figure 4. Compression gain and deviation loss for SensCompr of five independent smart meter datasets.

We extend our experiment with other different sensor signals from heterogeneous vertices like healthcare, transportation. In fig. 5, original and reconstructed ECG is shown. Performance comparison considering EEG, ECG, accelerometer sensor data is shown in fig. 6, where we observe that except ECG signal (Red marked, signal indexed 6, 7, 8), others show good performance gain. Mean performance gain, which is the aggregated compression performance for a system (edge gateway/ IoT platform) that handles multiple sensor data is shown in fig. 7.

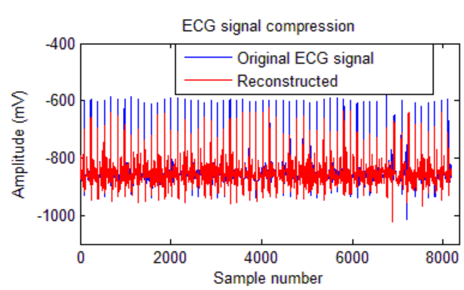
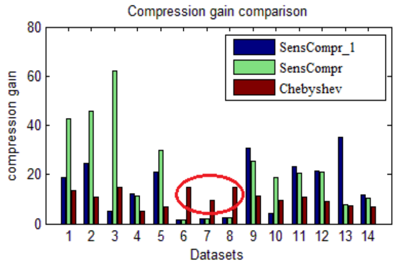


Figure 5. Original ECG and reconstruction.



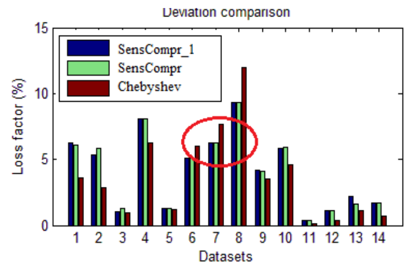


Figure 6. Performance gain comparison of SensCompr for heterogeneous sensor data, showing poor performance for ECG signal.

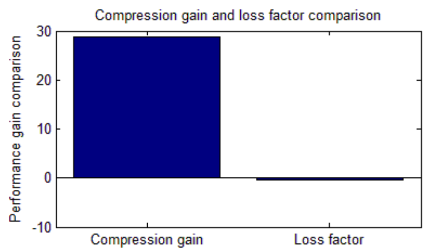


Figure 7. Mean compression gain and loss factor with heterogeneous sensor data, showing high (+ve) compression gain and –ve loss factor.

5. SensCompr Accurate

We analyze that the *useful information extraction method* is not appropriate for quasi-periodic signals like ECG. The outlier detection algorithm proposed is mainly for random like signals. We consider adaptive window-based discord discovery (AWDD) as the outlier detection technique *out* [16] for quasi-periodic signals like ECG. The subsequence $\mathscr{s}$ of length n of time-series $\varepsilon_t$is discord in$\varepsilon_t$if $\mathcal{S}$ has the largest distance to its nearest non-self match [16], i.e., ∀subsequence $\mathcal{S}$of ε_t non-self match $\lambda_{\mathscr{s}}$of $\mathscr{s}$, and non-self match $\lambda_{\mathscr{s}*}$of $\mathscr{s}^*$, min $(Dist(\lambda_{\mathscr{s}}, \ \mathscr{s}) > min\ (Dist(\lambda_{\mathscr{s}*},\mathscr{s}^*)$. Below in fig. 8, we show the significant improvement of compression performance while comparing SensCompr Accurate and SensCompr. It is to be noted from fig. 8 that substantial compression gain is achieved in ECG data compression by SensCompr accurate while least impact of information loss or loss factor. The complete SensCompr Accurate scheme is illustrated below:

SensCompr Accurate

5. Get the basic signal property of the sensor dataε.
6. If quasi-periodic -> use adaptive window-based discord discovery and find sensor importance score.
7. Else (as others are mostly non-stationary random process), consider point outlier detection using Hampel/Rosner filtering and find sensor importance score ϕ_ε (1).
8. Find optimal block sizeB_o recursively (3).
9. Find optimal threshold Γ_o (4).

10. Compress $\varepsilon \xrightarrow{Chebyshev,\, B_o, \Gamma_o} \varepsilon^{compressed}$ (2).

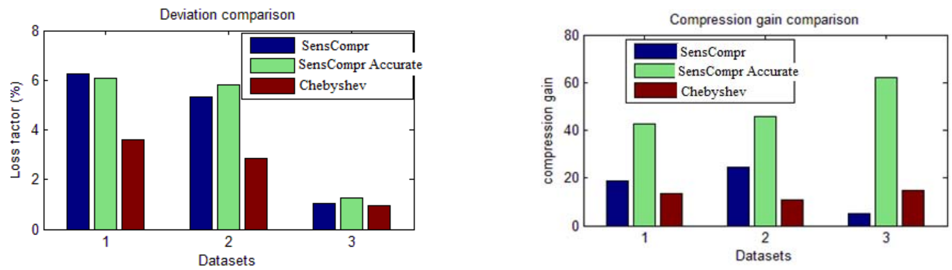


Figure 8. Improved compression performance with the help of extended SensCompr Accurate for quasi-periodic signals like ECG.

- **Practical Validation: Accuracy in ECG Feature Extraction**

One of the basic interpretation derived from ECG signal for human cardio health diagnosis, is the R wave characterization, particularly the R-R interval computation in the detection of QRS complex, which plays vital role in the determination of heart rate variability [17]. In fig. 9, it is shown that the error in R-R interval detection from reconstructed ECG signal is < 2%, which does not have impact in the diagnosis process.

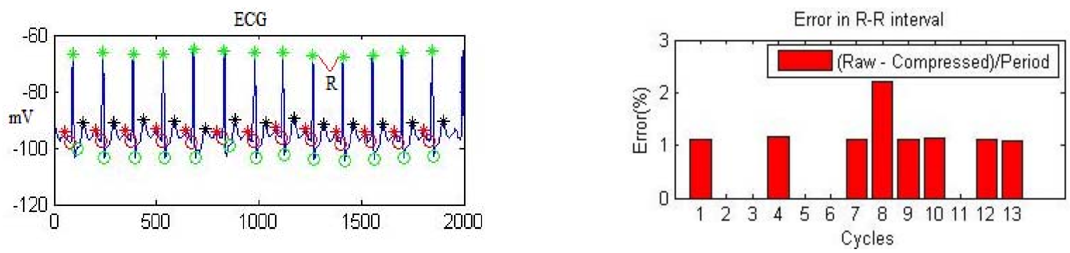


Figure 9. ECG signal compression using SensCompr, its impact to feature space analysis.

6. Conclusion

In this era of IoT and omnipresence of sensor devices, the storage, transmission, processing of high volume of sensor data pose potentially destructive performance and scalability risk. Arbitrary (conventional) compression may result in significant intelligence loss. In this paper, we proposed a novel sensor data compression scheme that is effective to multitude of sensor data types and yield considerable compression gain. Our unsupervised, information-centric approach of extracting the interesting phenomena, which is embedded with sensor dataset maximizes the compression gain while optimizing the data reconstruction loss independent of sensor data types. As, no single lossy compression provides considerable performance for every kinds of sensor datasets, our dynamic method of signal property and information theoretic approach is proved to be simplistic, practical as well effective. We intend to establish the universal appeal of this proposed scheme to ensure that larger gamut of sensor data is covered and the outcome is consistent.

Acknowledgement

We acknowledge the contribution of Chetanya Puri for his contribution towards development of AWDD code in Matlab.

References

[1] N. Hung, H. Jeung, and K. Aberer, "An evaluation of model-based approaches to sensor data compression," IEEE Trans. on Knowl. and Data Eng., pp. 2434 - 2447, 2013.
[2] V. Misra, V. K Goyal, and L. R. Varshney, "Distributed scalar quantization for computing: High-resolution analysis and extensions," IEEE Trans. Inf. Theory, vol. 57, pp. 5298–5325, Aug. 2011.
[3] S. E. Hawkins III, and E. H. Darlington, "Algorithm for compressing time-series data," NASA Tech Briefs, pp. 7-8, June 2012.
[4] J. Wang. et al. " Detecting and disposing abnormal signal outliers with masking effect by using data accumulated generating operation" IEEE CISP, pp. 426-430, 2008.
[5] A. Ukil, S. Bandyopadhyay, and A. Pal, "Sensitivity Inspector: Detecting privacy in smart energy applications, " IEEE ISCC, 2014.
[6] R. Serfling, and S. Wang, "General foundations for studying masking and swamping robustness of outlier identifiers," Elsevier Statistical Methodology, August 2013.
[7] L. Davies, and U. GatherSource, "The identification of multiple outliers," Journal of the American Statistical Association, vol. 88, no. 423, pp. 782-792, Sep. 1993.
[8] H. Liu, S. Shah, W. Jiang, "On-line outlier detection and data cleaning," Elsevier Computers and Chemical Engineering, pp. 1635–1647, 2004.
[9] B.Rosner, "Percentage points for a generalized ESD many-outlier procedure," Technometrics, vol. 25, issue. 2, pp. 165 - 172, 1983.
[10] A. Ukil, S. Bandyopadhyay, and A. Pal, "IoT privacy: To be private or not to be private," IEEE INFOCOM, 2014.
[11] F. Krzakala, M. Mezard, and L. Zdeborova, "Compressed sensing under matrix uncertainty: optimum thresholds and robust approximate message passing," IEEE ICASSP, pp. 5519 - 5523, 26 - 31 May, 2013.
[12] R. Banerjee, A. Sinha and K. Chakravarty, "Gait based people identification system using multiple switching kinects," ISDA, December 8 - 10, 2013.
[13] Z. Kolter, and M.J. Johnson, "REDD: A public data set for energy disaggregation research," ACM SustKDD, 2011.
[14] K. Anderson, et al. "BLUED: A fully labeled public dataset for event-based non-intrusive load monitoring research," ACM SustKDD, 2012.
[15] G.B. Moody, R.G. Mark, "The impact of the MIT-BIH Arrhythmia Database," IEEE Engg in Med and Biol., vol. 20, issue. 3, pp. 45-50, 2001.
[16] F.C. Chuah, and F. Fu, "ECG anomaly detection via time series analysis," ACM ISPA, pp. 123 – 135, 2007.
[17] H. Rabbani, M.P. Mahjoob, E. Farahabadi, and A. Farahabadi, "R peak detection in electrocardiogram signal based on an optimal combination of wavelet transform, Hilbert transform, and adaptive thresholding," Journal of Medical Signals and Sensors, pp. 91 – 98, 2011.