Using Adaptive Sampling and DWT Lifting Scheme for Efficient Data Reduction in Wireless Body Sensor Networks

Joseph Azar
FEMTO-ST Institute/CNRS
Univ. Bourgogne Franche-Comté
Belfort, France
Email: joseph.azar@univ-fcomte.fr

Carol Habib
FEMTO-ST Institute/CNRS
Univ. Bourgogne Franche-Comté
Belfort, France
Email: carol.habib@univ-fcomte.fr

Rony Darazi
TICKET Lab
Antonine University
Hadat-Baabda, Lebanon
Email: rony.darazi@ua.edu.lb

Abdallah Makhoul
FEMTO-ST Institute/CNRS
Univ. Bourgogne Franche-Comté
Belfort, France
Email: abdallah.makhoul@univ-fcomte.fr

Jacques Demerjian
LARIFA-EDST
Faculty of Sciences, Lebanese University
Fanar, Lebanon
Email: jacques.demerjian@ul.edu.lb

Abstract—In the recent years, many researches have been done on Wireless Body Sensor Networks, consisting of wearable devices that provide personalized healthcare through continuous monitoring of the patients' health condition. One of the major difficulties in WBSNs is the power consumption due to wireless transmission of sensed data. Data reduction can be considered a direct way to reduce the power consumption due to data transmission. However, most of the data reduction techniques suffer when the variation of the collected samples is high, or when the data are noisy. In this paper, we propose to enhance a data reduction scheme based on an adaptive sampling technique using dynamically adapted risk level by combining it with the Discrete Wavelet Transform lifting scheme for noise filtering. To assess our approach, we have run different series of simulation on real sensor data. The results show that combining the lifting scheme method with adaptive sampling increased the data reduction percentage by up to 50%.

Index Terms—Wireless Body Sensor Networks, Data Reduction, Energy-Efficiency, Adaptive Sampling, Discrete Wavelet Transform.

I. INTRODUCTION

A Wireless Body Sensor Network (WBSN) consists of sensor nodes attached to a patient's body, that allow the measurement of physiological parameters, such as heart rate, respiration rate, ElectroCardioGram (ECG), blood pressure, and many others. The attached sensor nodes are constituted of a battery, a sensing unit composed of a transducer and Analog to Digital (A/D) converter, a processing unit that contains a microcontroller, and a communication unit for receiving and transmitting data. These nodes send the collected measurements periodically to a coordinator, which is usually a portable device or smartphone, for fusion or aggregation. After that, the coordinator forwards the received data to the sink node or the base station [1] [2].

The interest in WBSN has been increasing recently, due to enabling an important number of innovative and interesting applications, such as distant patients monitoring and monitoring hospital and elderly patients [3] [4] [1]. However, in spite of the fact that the WBSN means to provide the perfect wireless setting to the networking of human body sensors and the setting up of inescapable health monitoring systems, there are various technical challenges that lie ahead. One of the key considerations for WBSNs is the energy consumption of sensor nodes. Usually, a battery with a limited energy budget serves as the source of power for sensor nodes. Furthermore, it is very difficult to recharge the battery, and in case of implantable setting, replacing a battery requires surgery which is difficult and costly.

In order to increase the lifetime of a sensor node, reducing the battery consumption is a must. In a WBSN, the wireless communication link is probably to be the greatest consumer of energy [5] [6]. One of the direct ways to reduce the energy consumption of the radio transceiver is to reduce the amount of data to be transmitted.

The main objective of implementing data reduction techniques on sensor nodes is to get rid of unneeded and redundant samples, thus reducing the amount of data to be transmitted. Data reduction techniques can be classified as follows [6] [7]:

- In-network processing: This technique consists of performing data aggregation to reduce the amount of data to be forwarded to the coordinator, sink, or other nodes. For example, a node can only transmit the average, minimum, or maximum of the sampled data.
- Data prediction: This technique consists of building a model describing the evolution of the data, and able to predict the sensed values within a certain margin of error. This model resides both at the sensor nodes and the

coordinator/base-station. If the predicted values satisfy the accuracy needed, there is no need to transmit the data. However, when the prediction is not accurate enough, the sensed values must be transmitted in order to update the model.

- Data compression: This technique consists of encoding the data in a way that the number of bits required to represent the compressed data is less than the number of bits required to represent the original data. Thus, reducing the number of packet transmissions [8].

One of the in-network processing techniques that is proposed to tackle the problem of energy conservation in sensor networks is the adaptive sampling [9]–[12]. The main idea behind adaptive sampling is to study the collected data between periods based on the dependence of conditional variance. Adaptive sampling techniques can optimize energy consumption in a sensor network in two ways:

- By dynamically adjusting the sampling frequency of a sensor based on to the level of variance between sensed data over a certain period of time. This approach will reduce the energy consumption of the sensing unit by preventing the sensor of collecting redundant information.
- By dynamically adjusting the rate at which a feature is calculated from the raw signal according to the level of variance between the calculated features over a certain period of time. In the case of a WBSN, vital signs data are extracted from the raw signals collected by the biosensor nodes. For example, the heart rate can be calculated using the raw ECG signal. So instead of calculating the heart rate each x seconds, it is calculated each z seconds where $z > x$ when the variance between periods is low. Thus, reducing the processing energy.

These techniques work well when the collected time series are smooth. In the case of noisy data, or jagged time series, these techniques perform poorly.

In this paper, we propose to enhance the performance of an adaptive sampling technique using dynamically adapted risk level proposed in [11] on jagged time series of vital signs. To do so, we combine this technique with two data processing techniques characterized by their simplicity and speed of computation:

1) The Discrete Wavelet Transform (DWT) lifting scheme based Haar wavelet transform used for noise removal.
2) The Differential Pulse Code Modulation (DPCM) used for data compression and to provide a higher level of data reduction.

In our experimentation, we consider that the monitored vital signs are calculated based on raw signals, and the adaptive sampling technique adapts the sampling rate at which a vital sign is calculated from the digital signal. Thus, affecting communication resources and processing costs.

The following of the paper is organized as follows: Section II presents an Early Warning Score (EWS) system used to evaluate the critical level of a collected vital sign. Section III explains the adaptive sampling technique used in our approach.

Section IV gives some background information about the DWT lifting scheme and DPCM algorithms and discusses our proposed data reduction scheme. Section V details and analyses the experimental results. Section VI concludes the paper.

II. EARLY WARNING SCORE SYSTEM

Acute care teams have long empowered the utilization of an EWS system to enable a more auspicious reaction to, and evaluation of, intensely ill patients. An EWS system is based on a straightforward scoring system that assigns a score to each physiological measurement monitored from acute patients. For each vital sign, we define a normal healthy range, and the assigned score shows how extreme the measured value varies from the norm. The more a measured value is outside the normal healthy range, the bigger is the value of the assigned score.

PHYSIOLOGICAL PARAMETERS	3	2	1	0	1	2	3
Respiration Rate	≤8		9 - 11	12 - 20		21 - 24	≥25
Oxygen Saturations	≤91	92 - 93	94 - 95	≥96			
Any Supplemental Oxygen		Yes		No			
Temperature	≤35.0		35.1 - 36.0	36.1 - 38.0	38.1 - 39.0	≥39.1	
Systolic BP	≤90	91 - 100	101 - 110	111 - 219			≥220
Heart Rate	≤40		41 - 50	51 - 90	91 - 110	111 - 130	≥131
Level of Consciousness				A			V, P, or U

Fig. 1: National Early Warning Score (NEWS)

In our work, we use the scoring template for the National EWS (NEWS) used in U.K [13] illustrated in Figure 1 to assess the severity level of the collected vital signs.

III. ADAPTIVE SAMPLING

The adaptive sampling technique that is proposed in [11] allows the sensor node to adjust its sampling rate periodically according to the dynamic evolution of the monitored vital sign while taking into consideration the vital sign's monitoring importance regarding the patient's health condition. For instance, if a vital sign is unstable then the sensor node's sampling rate should better be set to a maximum to avoid missing any important changes and to capture all variations. However, if the vital sign is somewhat stable then the sensor node's sampling rate should better be set to a minimum so that the energy level is preserved. Furthermore, we take another parameter into consideration: the monitoring importance assigned to a vital sign. The latter changes over time with the patient's changing health condition. A vital sign can be judged to be very essential/less essential at a certain disease stage (e.g. influenza, asthma attack, etc.) or for a specific population

(e.g. elderly, hypertension patients, heart risk failure at risk patients, athletes, etc.) or after a surgical intervention or for a recovering patient. As a consequence, the more the sensor node is judged to be critical, the higher its sampling rate gets in order to keep track of all variations and the less the sensor node is judged to be critical, the lower its sampling rate gets in order to conserve its energy level. In the following, the Fisher test with one-way ANalysis Of VAriance (ANOVA) is firstly presented, then the behavior function we have proposed for the sampling rate adaptation is described and finally, the proposed risk level evaluation function is presented.

A. Fisher Test with one-way ANOVA

The Fisher test with one-way ANOVA is used to check whether the expected values of a quantitative variable in several pre-defined groups differ from each other. We propose to test the following null hypothesis: The means of the measurements of the last h consecutive periods are equal. In order to do so, we compute the Fisher test with one-way ANOVA statistic test using the following formula:

$$F = \frac{\frac{SF}{(h-1)}}{\frac{SR}{(N-h)}} \tag{1}$$

where SF is the between period variation, SR is the within period variation, h is the total of consecutive periods and N is total of measurements. SF and SR are calculated as follows:

$$SF = \sum_{j=1}^{h} n_j \times (\overline{Y_j} - \overline{Y})^2 \tag{2}$$

$$SR = \sum_{j=1}^{h} \sum_{i=1}^{n_j} (y_{ji} - \overline{Y_j})^2 \tag{3}$$

where y_{ji} is the ith measurement of the jth period, n_j is the total of measurements of the jth period, $\overline{Y_j}$ is their mean and $\overline{Y}$ is the mean of all the measurements taken during the h consecutive periods. When the hypothesis is accepted the Fisher statistic follows the F-distribution with (h-1, N-h) degrees of freedom. If F is greater than the critical value $F_t = F_\alpha(h-1, N-h)$ that is defined for a given Fisher risk α (false-rejection probability), then the hypothesis is rejected. Otherwise the hypothesis is accepted.

B. Behavior Function

Bezier curves are parametric curves used to model shapes by having knowledge about some points of interest [10]. In this work, we used quadratic bezier curves that are represented using three points. These curves are delimited by the two points P_0 (start point) and P_2 (end point) and their curvature is determined by the point P_1. We define the coordinates of the three points as follows: $P_0(0; l_y)$, $P_1(b_x; b_y)$ and $P_2(h_x; h_y)$ such as $0 < b_x < h_x$ and $l_y < b_y < h_y$.

Since the position of P_1 varies on the diagonal [AB] of the behavior rectangle where $A(0; h_y)$ and $B(h_x; l_y)$, the

coordinates b_x and b_y of P_1 satisfy the equation of [AB] which is defined as follows:

$$y = \frac{l_y - h_y}{h_x} \times x + h_y \tag{4}$$

In our approach, the sensor node's risk level r determines the position of P_1 and thus the curvature of the BV function:

$$B_r : [0; 1] \longmapsto [0; h_x] \times [l_y; h_y]$$
$$r \longmapsto (b_x; b_y)$$

Hence, the following equations can be derived to find the coordinates b_x and b_y of the behavior point P_1:

$$B_r(r) = \begin{cases} b_x = (1-r) \times h_x \\ b_y = l_y + r \times (h_y - l_y) \end{cases} \tag{5}$$

The higher is the value of r, the more the vital sign is considered to be critical and its monitoring importance is increased.

Finally, the BV function curve can be represented using the following quadratic bezier curve functions:

$$BV(F) = \begin{cases} \frac{h_x + l_y - 2b_y}{4b_x^2} F^2 + \frac{h_y - l_y}{h_x} F + l_y, & h_x = 2b_x \\ \\ (h_y + l_y - 2b_y)\alpha(F)^2 + \\ 2(b_y - l_y)\alpha(F) + l_y, & otherwise \end{cases}$$

with $\alpha(F) = -b_x + \frac{\sqrt{b_x^2 - 2b_x F + h_x F}}{hx - 2bx}$, such as $0 \le b_x \le h_x$, $0 \le F \le h_x$ and $h_x > 0$.
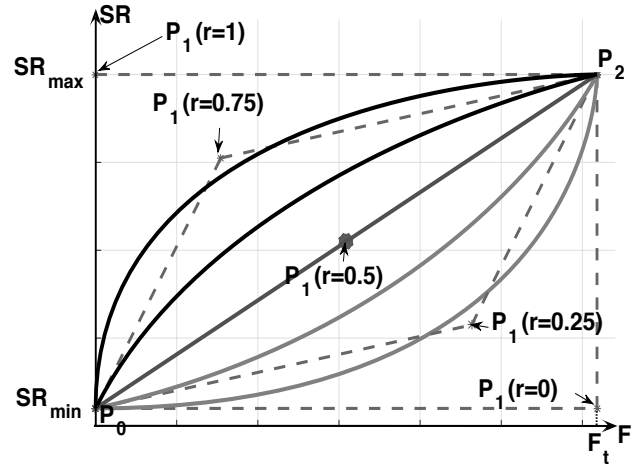

Fig. 2: Behavior Function: Quadratic Bezier Curves

The sampling rate of a sensor node is updated based on the BV function shown in Figure 2. This function takes as parameters the Fisher test result F and the risk level r of the vital sign given the patient's health condition. The result of Fisher test F is represented by the x-axis and the sampling rate

SR is represented the y-axis. The start point P_0 corresponds to $F = 0$ and $SR = SR_{min}$ and the end point P_2 corresponds to $F = F_t$ and $SR = SR_{max}$ where F_t is the critical value given by the Fisher Test for N measured samples during h periods. Based on the application's requirements, we set a maximum sampling rate SR_{max} and a minimum sampling rate SR_{min}. Depending on the values of SR_{max}, SR_{min}, and the critical value F_t given by the Fisher Test, the BV function is then defined as follows:

$$BV(SR_{max}, SR_{min}, r, F, F_t) = SR \qquad (6)$$
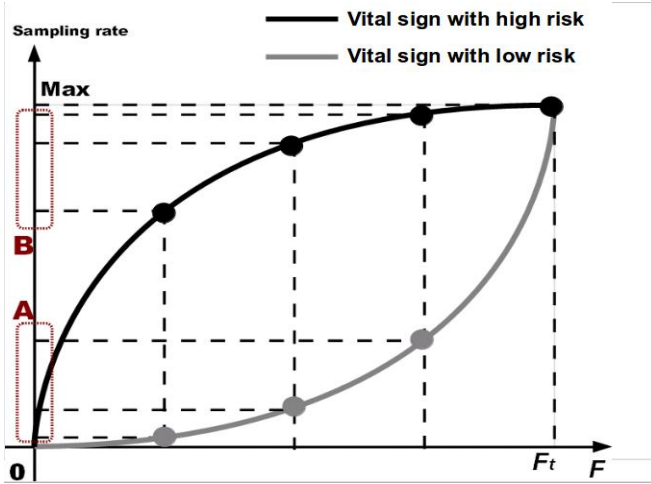
Fig. 3: Sampling rate adaptation according to the vital sign's risk level

When the collected measurements in h periods show high variations (F is close to F_t), we set a high sampling rate to the sensor node to avoid missing important variations. Otherwise, when F is close to 0, we set a low sampling rate to the sensor node to preserve its energy level.

Figure 3 illustrates the relation between the risk level r and the sensor node's sampling rate for the same value of F. If the value of r is low, the sensor node's sampling rate will take a lower value in order to preserve the energy level. In contrast, the sensor node, which is monitoring an essential vital sign regarding the patient's health condition, will be assigned a higher sampling rate in order not to miss important events and measurements. As a result, the number of samples transmitted to the coordinator at the end of a time period p can be increased or decreased depending on the sensor node's sampling rate.

C. Risk Level Evaluation

The patient's health condition may change throughout the life of the WBSN. Thus, we propose to adapt the risk level r of the monitored vital signs. For any vital sign, we calculate its monitoring importance (risk level) using the following risk level evaluation function:

$$Eval : (r_{global}, \bar{S}) \longmapsto [0; 1]$$
$$Eval(r_{global}, \bar{S}) = \alpha \times r_{global} + \beta \times \frac{\bar{S}}{S_{max}} \qquad (7)$$

where r_{global} represents the patient's overall health condition in terms of severity level and the score ratio $\frac{\bar{S}}{S_{max}}$ represents the situation of a vital sign (its criticality) compared to the worst case (highest criticality level). $\bar{S}$ is the average score of the measurements during one round R, such as $R = h \times p$ where p is a time period and $h \in \mathbb{N}$, and S_{max} is the maximum score that a measurement can have according to the used EWS.

We refer r_{global} as the patient's global risk level such as $r_{global} \in [0; 1]$ and $r_{global} \in \mathbb{R}$. It is assessed by the coordinator of the WBSN based on the multi-sensor data fusion approach proposed in [14]. The higher its value is, the more the patient's health condition is assessed to be severe.

The coefficients α and β are weight coefficients such as : $0 \leq \alpha \leq 1$ and $0 \leq \beta \leq 1$ and $\alpha + \beta = 1$.

The values of these coefficients are defined by the healthcare experts. If the overall health condition of the patient is more essential, thus α better have a higher value than β. On the other hand, if the status of the vital sign itself is more essential, then β better have a higher value than α. For example, some viruses such as influenza are accompanied by fever. We should then give a higher monitoring importance to the temperature since it has a high impact on these types of sickness. Thus, assigning higher value to β give more importance to the score ratio of the temperature instead of the overall health condition regrouping all the monitored vital signs.

IV. PROPOSED DATA REDUCTION SCHEME

In this section, we start by giving a brief introduction on the DWT lifting scheme and how it can be used for noise removal. Then we introduce the DPCM algorithm and finally we discuss our proposed approach, in which we consolidate the adaptive sampling, DWT lifting scheme, and DPCM techniques.

A. DWT lifting scheme for noise filtering

The lifting scheme is a procedure that implements the DWT, and compute its coefficients using means and differences. This procedure was introduced by Sweldens in [15].

The monitored vital signs values are usually a set of integer numbers, while wavelet transforms results in floating point numbers. In effect, we choose to transform the vital signs data using the integer lifting scheme, that converts integers to integers. Note that an integer dataset can be represented using fewer bits than a floating point dataset.

In the following, we present the integer version of the lifting scheme procedure step by step. Consider a set of data x_j with 2^n elements, the lifting scheme performs the transform in three operations: split, predict, and update as shown in Figure 4.

$a)$ Split operation: In this operation, the data are divided into two sets: the even and odd sets, containing each 2^{n-1}
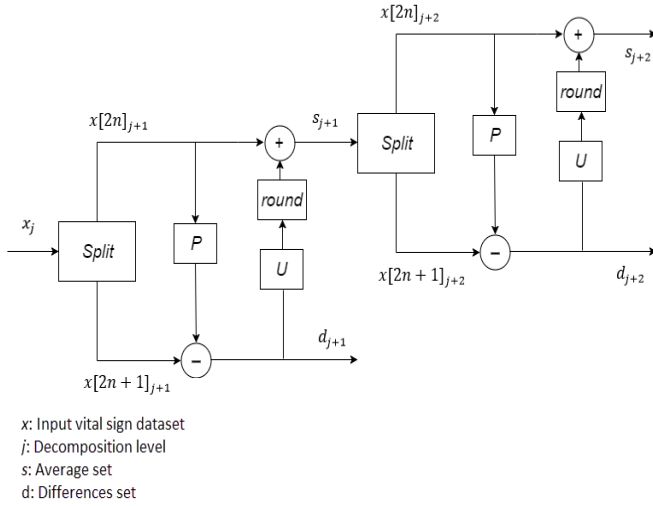
Fig. 4: The integer lifting scheme forward transform

x: Input vital sign dataset
j: Decomposition level
s: Average set
d: Differences set

elements as shown in Figure 4. Afterwards, the odd set is transformed into the differences set denoted by d_{j+1} (predict operation), and the even set is transformed into the average set denoted by s_{j+1} (update operation).

b) Predict operation: In this operation, the even set is used to predict the odd set. This is based on the fact that a correlation exists between an odd element and its even neighbors. Each element in the odd set is replaced by the difference between its actual value and the predicted value. The even elements are left unchanged and become the input for the update operation. The prediction, which is the difference, is denoted by Equation 8:

$$d[n]_{j+1} = x[2n+1]_j - P \times (x[2n]_j). \tag{8}$$

where P is the prediction operator. The significance of the prediction operation is that it takes advantage of the correlation between the data and the small difference between an odd element and its prediction.

In this paper, we are using the integer lifting scheme version of the Haar transform due to its simplicity and speed of computation. The Haar transform is based on the calculation of averages and differences (details). Consider two consecutive samples a and b, the idea is to calculate the average $s = (a+b)/2$ and the difference $d = b - a$. In this paper, we use the Haar prediction as in Equation 9

$$d[n]_{j+1} = x[2n+1]_j - x[2n]_j. \tag{9}$$

c) Update operation: The update operation replaces the even entry by an average. The importance of this operation is that it reflects our knowledge of the data and results in a smoother input for the next decomposition level. The update operation is defined by Equation 10:

$$s[n]_{j+1} = x[2n]_j + U \times (d[n]_{j+1}). \tag{10}$$

In the Haar version of the integer lifting scheme, the update operation replaces an even element with the average of the even/odd pair as in Equation 11

$$s[n]_{j+1} = x[2n]_j + \lfloor d[n]_{j+1}/2 \rfloor. \tag{11}$$

The lifting scheme reverse transform uses the same operators, P and U similarly to the forward transform, but with a reversal of sign at each step. The reverse transform is based on the three operations: reverse update, reverse prediction, and merge. These three operations are defined in Equations 12, 13, and 14 respectively.

$$x[2n]_j = x[2n]_{j+1} - U \times (d[n]_{j+1}). \tag{12}$$

$$x[2n+1]_j = x[2n+1]_{j+1} + P \times (x[2n]_j). \tag{13}$$

$$x_j = Merge(x[2n+1]_j, x[2n]_j). \tag{14}$$

The described operations (split, predict, and update) can be repeated n times on a set of data of length 2^n. The final result is a set s_n of one average (the mean value of all entries) and n sets of differences.

In this work, we are interested in using the DWT lifting scheme for noise filtering and data smoothing. To do so, we separate the signal into approximation and detail coefficients as show in Figure 5.

Figure 5 illustrates 1-level wavelet decomposition of 64 heart rate samples. Each decomposition level l_j breaks down the signal into two sets: approximation coefficients set (averages) s_j and detail coefficients set (differences) d_j. The information of the original signal is conserved in the approximation coefficients. However, some of the detail coefficients, which represent detail movements in data, may be considered as noise. We can set these coefficients to zero prior to the reconstruction process in order to remove noise from the original signal. In other words, the reconstruction requires rebuilding the signal from every component but noise. Figure 6 illustrates the smoothed signal after discarding the detail coefficients (considered as noise), and reconstructing the signal from Level-1 approximation coefficients.

B. Differential Pulse Code Modulation

The DPCM, also known as Delta coding, is a lossless compression technique that takes advantage of the correlation between two consecutive samples. The differences between correlated samples are small, resulting in compression. The DPCM operates by keeping the first sample as a reference and replacing every next sample with the difference between this sample and the previous one:

$$d(l) = \begin{cases} x(l), & \text{if } l = 0 \\ x(l) - x(l-1), & \text{if } l > 0 \end{cases} \tag{15}$$

In this work, we intend to apply the DPCM technique on the level-1 approximation coefficients resulting from the DWT
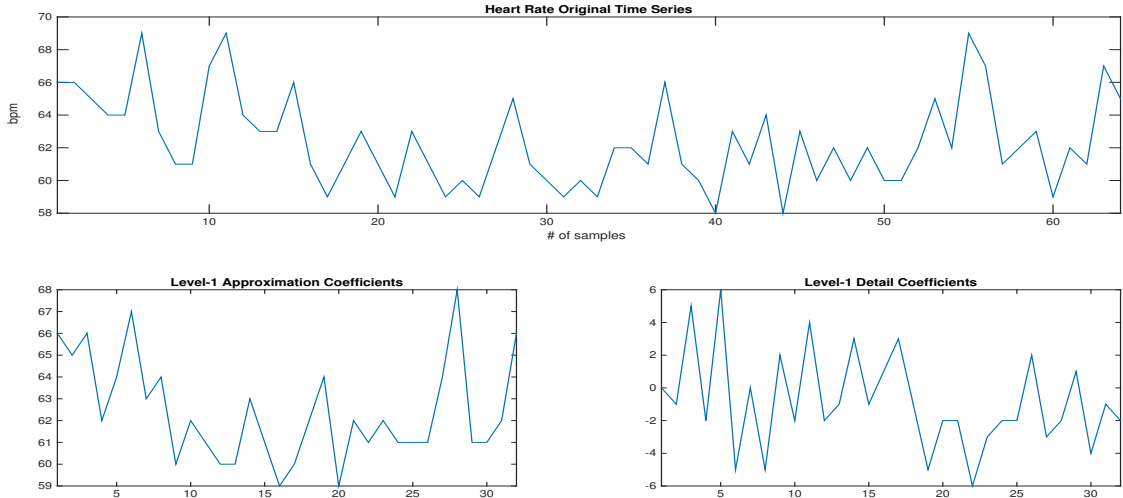
Fig. 5: Heart Rate signal wavelet decomposition

lifting scheme transformation as discussed in the previous section. The resulted signal consists of smaller values that are close to zero, and need smaller number of bits to be represented.

C. Proposed data reduction scheme
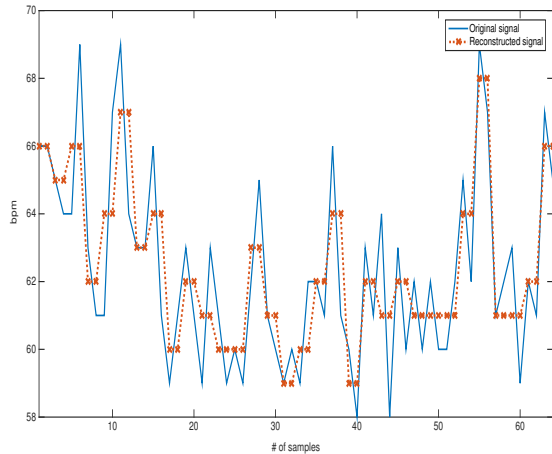
Fig. 6: Original heart rate signal vs reconstructed heart rate signal

To gain more data reduction, we propose to combine the adaptive sampling using dynamically adapted risk level with the integer lifting scheme based Haar wavelet transform and the DPCM techniques.

In the first place, the adaptive sampling reduces the number of samples by exploiting the variances of the collected measurements over time while taking into consideration the monitoring importance of a given vital sign. Given that the lifting scheme processes datasets of length 2^n, where n is a positive integer. Thus, we broke up the sampled dataset into

subsamples each of which has a power of two number of samples. After that, we apply a 1-level wavelet decomposition using the integer lifting scheme to each of these subsamples, and transforming each one of them into two sets: approximation coefficients and detail coefficients. In the next step, we discard the detail coefficients sets (will be replaced by zeroes when reconstructing the data) and apply the DPCM on the approximation coefficients sets. Finally, we encode the compressed coefficients and transmit the resulted bitstream to the coordinator. The described process is illustrated in Figure 7.

The encoding is done by converting the entries of the compressed signal X' into binary words. We use two word sizes to encode X'. Equation 16 shows the calculation of the word sizes. The smaller word size fits the absolute value of the compressed signal's median $med(X')$, and thus it is used to encode the entries in X' that are smaller than or equal $med(X')$. The larger word is used to encode the entries greater than $med(X')$ as it fits the entry with the largest absolute value in X'.

To differentiate between these two sizes, we add a prefix bit before each code. A value of "0" of this bit is used to indicate the small word size, and a value of "1" to indicate the large word size. We add another prefix bit for the numbers sign.

$$W(X'[i]) = \begin{cases} \lceil \log_2 med(X') \rceil, & X'[i] \leqslant med(X') \\ \lceil \log_2 max(X') \rceil, & X'[i] > med(X') \end{cases} \quad (16)$$

V. EXPERIMENTAL RESULTS AND ANALYSIS

A custom Java-based simulator has been used to test the following data reduction techniques: adaptive sampling (AS), DPCM, adaptive sampling + DPCM (AS+DPCM), and the proposed data reduction scheme (AS+LS+DPCM).
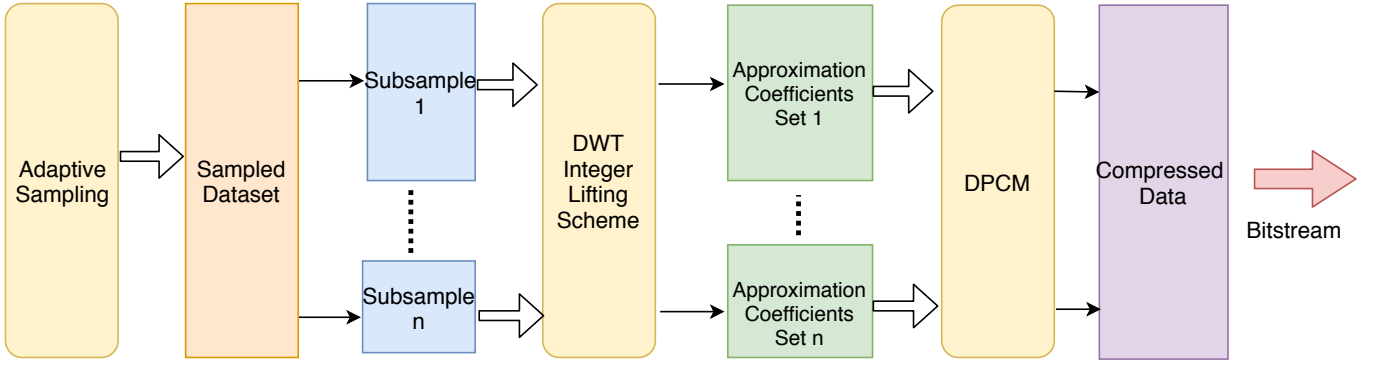
Fig. 7: Proposed data reduction scheme

The following sections present the results of applying the aforementioned data reduction techniques on a heart rate time series collected from a subject that have high blood pressure (hypertension) at Femto-ST institute using Polar M600 wearable. We consider that the heart rate samples are accumulated in the memory and transmitted after each period $p = 320$ sec.

Three metrics are discussed in the following: characteristics of the collected data, loss of information, and data reduction.

A. Characteristics of the collected data

Table I shows some statistical characteristics of the heart rate time series. Specifically, we have computed the min, max, mean $\bar{s}$, and the standard deviation $\sigma_{\bar{s}}$ of the samples, in addition to the mean $\bar{d}$ and the standard deviation $\sigma_{\bar{d}}$ of the differences between consecutive samples. It is noticeable that the standard deviation and the range of the data give high values, which correspond to jagged time series.

TABLE I: Statistical characteristics of the heart rate (HR) time series

	$\bar{s}$	$\sigma_{\bar{s}}$	$\bar{d}$	$\sigma_{\bar{d}}$	min	max
HR	71.8	12.2	2.7	3.4	40	147

B. Adaptive sampling vs Loss of information

In this section, we first define the parameters settings of the adaptive sampling technique, then we discuss its effect on the loss of information.

The adaptive sampling chosen parameters settings are defined as follows:

- Round $R = 2 \times p$.
- Total number of periods = 78.
- Fisher Risk $\alpha_{fisher} = 0.03$.
- Minimum sampling rate $SR_{min} = 1 samples/16sec$, corresponding to 20 samples per period, which is the minimum number of samples required for the Fisher Test and ANOVA.
- Maximum sampling rate $SR_{max} = 1 samples/5sec$ corresponding to 64 samples per period.
- We set the initial risk level to $r° = 0.5$, revealing that the vital sign is slightly critical.

In order to track the impact of the adaptive sampling and DWT lifting scheme techniques on the loss of information, we used the Root Mean Square Error (RMSE) error measurement statistic. For each period, we measure the difference between the mean value of the reconstructed dataset, and the mean values of the actually observed dataset as denoted in Equation 17. Then, we measure the difference between the score of the mean value of the reconstructed dataset and the score of the mean value of the original dataset as denoted in Equation 18.

$$RMSE = \sqrt{\frac{\sum\limits_{i=1}^{n} \left(A_i - B_i \right)^2}{n}} \qquad (17)$$

$$RMSE_{scores} = \sqrt{\frac{\sum\limits_{i=1}^{n} \left(S(A_i) - S(B_i) \right)^2}{n}} \qquad (18)$$

Where n is the total number of periods during the simulation, A is the mean value of the ith period in the original dataset, B is the mean value of the ith period in the reconstructed dataset, and S is the function that allocates a score for each vital sign based on the EWS.

Table II shows the Root mean squared error between the original and sampled datasets' measurements/scores. It is clear that when using the DWT lifting scheme for noise filtering, the RMSE value has been increased. This is due to discarding a significant amount of coefficients (detail coefficients) which will increase the data reduction percentage.

What is most vital to us is to not miss the critical events or changes in the data. In other words, the severity level of the vital sign being monitored, which reflects a patient's health status. The S-RMSE of the proposed data reduction scheme is very close to 0.2, thus the loss of information is insignificant.

TABLE II: Root mean squared error between the original and sampled datasets' measurements/scores

	AS	AS+DPCM	AS+LS+DPCM
RMSE	0.016	0.016	0.546
S-RMSE	0	0	0.022

C. Data reduction

In this section, we study and compare the data reduction performed by the adaptive sampling using dynamically adapted risk level, the DPCM, and the proposed data reduction scheme over 78 periods.

In order to compare the performance of the data reduction techniques, we computed the Compression Ratio (CR) and the percentage reduction denoted as:

$$CR = \frac{orig\_data}{red\_data}. \tag{19}$$

$$Red(\%) = 100 \times \left(1 - \frac{1}{CR}\right). \tag{20}$$

Where red_size and $orig\_size$ are the sizes in bits of the reduced and the original bitstreams respectively. We consider that a heart rate sample is represented using 9 bits. Table III compares the performance obtained by the four approaches. The adaptive sampling approach was able to achieve a reduction percentage by up to 17%, while the DPCM achieved a reduction percentage by up to 28%. The reason behind the poor performance of the adaptive sampling and DPCM techniques is that these techniques highly depends on the temporal correlation among the collected data. Since the collected data do not show high temporal correlation, the results achieved by these techniques were not as expected. When we combined the adaptive sampling with the DPCM, we were able to increase the reduction percentage by up to 40%. On the other hand, applying 1-level wavelet decomposition on the sampled data using the DWT lifting scheme was able to achieve a reduction percentage by up to 69%, thus increasing the performance achieved by the adaptive sampling by more than 50%.

The reasons behind the significant improvement of the data reduction percentage when using the DWT lifting scheme for noise filtering are:

1) A significant amount of data considered as noise (detail coefficients) has been discarded and replaced by zero at the sink level.
2) The DPCM is now applied on the smoothed approximation coefficients. Hence, the differences between consecutive values are now smaller.

TABLE III: Average percentage reduction of the deployed data reduction approaches

	# of bits	CR	Red(%)
Original Data	44928.0		
AS	36900	1.218	17.86
DPCM	32153	1.397	28.43
AS+DPCM	26860	1.673	40.21
AS+LS+DPCM	14041	3.2	68.74

VI. Conclusion

Given that the communication unit in a sensor node is the primary consumer of energy in WBSNs, data reduction can be considered as an efficient and direct way to increase the lifetime of a WBSN. In this paper, we presented a data reduction scheme in which we increased the performance of an adaptive sampling technique using dynamically adapted risk level on jagged time series of vital signs by combining it with the discrete wavelet transform lifting scheme used for noise removal and the differential pulse code modulation used for data compression. The results show that the proposed data reduction scheme was able to increase the data reduction percentage by up to 50%.

References

[1] X. Lai, Q. Liu, X. Wei, W. Wang, G. Zhou, and G. Han, "A survey of body sensor networks," *Sensors*, vol. 13, pp. 5406 – 5447, 2013.
[2] N. Boudargham, J. B. Abdo, J. Demerjian, and C. Guyeux, "Exhaustive study on medical sensors," in *SENSORCOMM 2017, The Eleventh International Conference on Sensor Technologies and Applications*. IARIA, 2017.
[3] R. Negra, I. Jemili, and A. Belghith, "Wireless body area networks: Applications and technologies," *Procedia Computer Science*, vol. 83, pp. 1274 – 1281, 2016.
[4] T. Yilmaz, R. Foster, and Y. Hao, "Detecting vital signs with wearable wireless sensors," *Sensors*, vol. 10, pp. 10 837 – 10 862, 2010.
[5] G. Z. Yang, *Body Sensor Networks*. Springer, 2006.
[6] G. Anastasi, M. Conti, M. D. Francesco, and A. Passarella, "Energy conservation in wireless sensor networks: A survey," *Ad Hoc Networks*, vol. 7, pp. 537 – 568, 2009.
[7] T. Rault, A. Bouabdallah, and Y. Challal, "Energy efficiency in wireless sensor networks: A top-down survey," *Computer Networks*, vol. 67, pp. 104 – 122, 2014.
[8] T. Srisooksai, K. Keamarungsi, P. Lamsrichan, and K. Araki, "Practical data compression in wireless sensor networks: A survey," *Journal of Network and Computer Applications*, vol. 35, pp. 37 – 59, 2012.
[9] D. Laiymani and A. Makhoul, "Adaptive data collection approach for periodic sensor networks," in *2013 9th International Wireless Communications and Mobile Computing Conference (IWCMC)*. IEEE, 2013.
[10] A. Makhoul, D. Laiymani, H. Harb, and J. M. Bahi, "An adaptive scheme for data collection and aggregation in periodic sensor networks," *International journal of sensor networks*, vol. 18, pp. 62–74, 2015.
[11] C. Habib, A. Makhoul, R. Darazi, and R. Couturier, "Real-time sampling rate adaptation based on continuous risk level evaluation in wireless body sensor networks," in *2017 IEEE 13th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*. IEEE, 2017.
[12] G. B. Tayeh, A. Makhoul, D. Laiymani, and J. Demerjian, "A distributed real-time data prediction and adaptive sensing approach for wireless sensor networks," *Pervasive and Mobile Computing*, vol. 49, pp. 62 – 75, 2018.
[13] "National early warning score (news), royal college of physicians, london, u.k., may 2015." https://www.rcplondon.ac.uk/projects/outputs/national-early-warning-score-news, accessed: 2017-12-07.
[14] C. Habib, A. Makhoul, R. Darazi, and R. Couturier, "Multisensor data fusion for patient risk level determination and decision-support in wireless body sensor networks," in *Proceedings of the 19th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*. ACM, 2016.
[15] W. Sweldens, "The lifting scheme: A construction of second generation wavelets," *SIAM Journal on Mathematical Analysis*, vol. 29, pp. 511–546, 1998.