European Journal of Molecular & Clinical Medicine
ISSN 2515-8260 Volume 7, Issue 11, 2020

Implementation of Efficient Data Compression Technique using Bit Reduction Burrows Wheeler Transform for Wireless Sensor Networks Environment

R. Buvanesvari[1*], A. Rijuvana Begum[2]

[1]*Research Scholar, Department of Information Technology, PRIST Deemed University, Thanjavur, India*

[2]*Research Supervisor, Department of Electronics and Communication Engineering, PRIST Deemed University, Thanjavur, India*

[1]*buvanapon@rediffmail.com*

[2]*arbbegum@gmail.com*

Abstract: ***Wireless Sensor Network (WSN) plays a significant role in Internet of Things (IoT) and it incorporated to the physical atmosphere for observing the parameters like temperature, pressure and so on. The nodes in WSN are limited interms of energy, storage, bandwidth and computation. As the communication cost is greater when compared to sensing and processing cost, a number of data compression models are applied to minimize the quantity of data being forwarded in the network. This paper introduces a new Bit Reduction with Burrows Wheeler Transform called BR-BWT based data compression technique in WSN. The presented BR-BWT model performs encoding of data in two ways namely bit reduction using codeword allocation and BWT based encoding processes. Initially, a bit reduction process takes place using predefined codeword allocation process to determine the codeword for every character in the WSN data. Besides, the BWT based compression process takes place to further compresses the bit reduced data. To validate the performance of the BR-BWT model, a real time WSN dataset is tested and the results are discussed under diverse aspects.***

Keywords: Data compression, WSN, Codeword allocation, BWT, Energy Efficiency

1. INTRODUCTION:

In recent times, the development in wireless networks and Micro-Electro-Mechanical-System (MEMS) prompts the improvement of small and inexpensive sensor nodes. WSN have developed into a trending investigation topic because of its applications in different areas, like, observing environment, monitoring boundary, managing disaster, smart building, smart cities, and so on [1]. WSN comprises an enormous number of self-governing, dedicated, portable sensors placed arbitrarily in the interested field for monitoring the physical variables in the environmental. It is generally positioned unevenly in the target area to gauge physical values such as humidity, pressure, temperature, and so on. Every sensor node comprises of 4 fundamental parts such as sensing unit, processing unit, transmission unit, and power supply [2]. A sensor node could be measured as an embedded model, which monitors the surrounding area, accumulates the information, and communicate it to the base station (BS) through single or multi-hop transmission. As the sensor nodes are operated with inbuilt batteries and need to work for a long duration, it is difficult to recharge batteries [3]. Energy efficiency is measured as a significant part in configuring WSN. Several research communities demonstrate that extreme energy is used for transmission than sensing and energy computation [4]. In WSN, the main issue to save energy depends on the efficient way of transmitting huge amount of data. There are two standard approaches to decrease energy utilization and enhance the system lifetime namely scheduling and data aggregation.

The primary process is the use of sleep state and dynamic state modes in the sensor [5], [6], [7]. A portion of the corresponding nodes could go into the sleep state and need more nodes count to stay alive for the whole area coverage and furthermore it highly preserves energy. In the subsequent process, data compression procedures are utilized for reducing data transmission to accomplish better power efficiency. It minimizes the entire bit count expected to store or transmit a huge data [8]. It is carried out by the representation of the data in its compacted manner. There are different compression methods concerning few concepts, and appropriate for various data types. For data compression, the method generally required to evaluate the data, recognizes the replication, and removes it. No general and effective compression method has been developed for extensive range of data types.

According to the value of decompressed data, data compression could be grouped into 3 different methods: lossy, lossless, and irreversible compression. The initial lossy compression is utilized for accumulating the information from where definite data could be acquired after decompression. Huffman coding [9] is one of the lossless compression methods which is applied frequently. Secondly, lossless is defined as process of eliminating a part of data [10]. The real-time data is not obtained from the decompression of static compression, like irretrievable compression. Text compression is typically accomplished through the allocation of shorter codes for normal messages whereas longer codes for unknown texts. Also, it is encoded using query tables with previous code length. Many studies have stated that, the newly deployed models are not applicable for today's applications since there are reasonable merits and equal number of demerits like storage and processing speed. As WSN data includes arithmetical values, the architecture of compression model for that information would be extremely helpful and prompts better compression execution [20-24].

Different kinds of text compression models were deployed as defined in the study. Run Length Encoding (RLE) [1] is defined as a unique character encoding. If the input character d emerges for n consecutive iterations, then n times of input sequence has been substituted by a pair nd. But, it is not as effective in data compression while the redundancy in input data is minimum. Followed by, Huffman coding [2] allocates diverse length codes for input sequence which depends upon the number of times a smaller codeword appears as well as maximum codeword for rare existence of characters. Although it can be executed simply, decoding is highly complex to find the final bit of coded character.

Arithmetic coding is mainly used for reducing the constraints involved in Huffman coding. It is accomplished by encoding the data into single character. Each message is depicted by real values from [0, 1], where $0 \leq x < 1$. However, it is composed of few constraints like the decoding task is initialized after the complete data encoding is processed. The existence of a single bit failed data leads to the failure of whole document. Lempel-Ziv-Welch (LZW) coding [3] is meant to be a dictionary relied coding model that develops a dictionary for sequentially existing patterns. Basically, LZW replaces the predicted patterns with references for a dictionary. If the size of dictionary has been improved, then the references values are also enhanced. In order to resolve the limitations involved in this model, the above mentioned methods were developed and compute the process accordingly.

Lightweight Temporal Compression (LTC) is mainly applied for eliminating the faults in all readings, by managing a knob [4]. By compression, future savings are carried out with maximum errors where K-RLE is an extended method of RLE and defined as a string of N values from [K-d, K+d] as a pair (N,d), in which K denotes the accuracy level while d shows a data item [5]. In specific, for resource composed of WSN [6], a lightweight compression technology is an expanded version of LZW dictionary relied coding so called Sensor LZW (S-LZW) has been deployed. However, the major issue in this approach is that, increasing dictionary size and to enhance the compression task. Encoding 528 bytes of data with extra interdependent packets which is a tedious process. Lossless Entropy Compression (LEC) [7] is defined as a predictive coding model. LEC evaluates the variations from successive sensor values and classifies it into diverse classes where the volumes are improved drastically. Among S-LZW, LEC has accomplished better compression; however, it is static and inapplicable for the modifications of source information statistics.

Sequential LEC (S-LEC) model has been developed [8] as independent compression technique for solving the limitations. The data is compressed in WSN by applying frequent context data from nearby residues. The demerits can be avoided and reduced using adaptive lossless data compression (ALDC) [9]. In order to process a best compression, ALDC model applicable to apply the changes present in source data statistics. Even though, the ALDC is independent, still it is a worst performer. The detection errors were compressed using an entropy encoder [10] while ALF has been applied to examine future M instances in Adaptive Linear Filtering Compression (ALFC). The changes made in the source is highly applicable in reducing the necessity of co-efficient extraction in adaptive prediction model.

Predictive coding used in 2 modal transmission (TMT) is deployed [11], that contains 2 modules namely compressed as well as non-compressed. In case of compressed mode, the range of [−R, R] is forwarded by encoded bits while in non-compressed mode, the actual data of error norms exist from [−R, R]. In order to reduce the function of maximum error terms detection, it solves the constraints of reduced code efficacy. For reaching the lower storage and rapid compression model, the Fast and Efficient Lossless Adaptive Compression Scheme (FELACS) has been deployed [12]. Thus, the energy consumption and compression effectiveness has been improved using compression process effectively.

In resource-based networks, where the model is independent for data accuracy; else, it is operated in exception state, data minimization approach is highly applicable. In case of sensors [13], an adaptive sampling technique has been presented for determining the best sampling frequencies. For snow-monitoring domains, sensor devices have been applied. With respect to classical fixed-rate model, the final outcome depicted that adaptive models have minimum amount of instances. For remote as well as expanded aquatic state, a novel adaptive sampling model for energy control in robotic observation of water quality has been initialized. For the enhancement of sampling task, a multivariate sampling (MuSA) model [14] with the application of component analysis (CA) technique. The final outcomes show that MuSA decreases the power consumption and delay in data transmission process. In order to reduce the consolidation of data transmission, a data-based model is proposed in [15].

A data-driven adaptive sampling algorithm (DDASA) has been utilized for maximizing the energy efficiency [16]. A one step decoding is reformed for original data, [17] examine the compressed sensing against network coding. Therefore, it demonstrates the coding and reformation model that activates accurate reconstruction. In order to reach effective power utilization in data collection, sparsest random sampling model for cluster-based compressive data gathering (SRS-CCDG) in WSNs has been presented [18], in which the sparsest random sampling method have been combined into WSNs. Furthermore, explanatory methods are applied to develop a correlation over power cost as well as cluster size at the time of using diverse inter-cluster as well as intra-cluster transmission approaches. Finally, the simulation outcome pointed that the SRS-CCDG enhances the independency of model and reduces the power cost.

European Journal of Molecular & Clinical Medicine
ISSN 2515-8260 Volume 7, Issue 11, 2020

This paper introduces a new BR-BWT based data compression technique in WSN. The presented BR-BWT model performs encoding of data in two ways namely bit reduction using codeword allocation and BWT based encoding processes. Initially, a predefined codeword allocation process is carried out to reduce the bit count needed for every character in the WSN data. The allocation of codeword significantly reduces the data size into 50% by assigning a predefined 4-bit codeword to 8-bit data. Moreover, the BWT based compression process takes place to further compress the encoded data in WSN. To validate the performance of the BR-BWT model, original WSN dataset is tested and the final outcomes are discussed under diverse aspects.

2. RESEARCH ELABORATIONS

The newly deployed BR-BWT model depends upon a single bit, dictionary relied single character encoding method exploits a 4-bit code allocation dictionary (CAD) for allocating codewords for input series. The entire performance of BR-BWT method is depicted in Fig. 1.

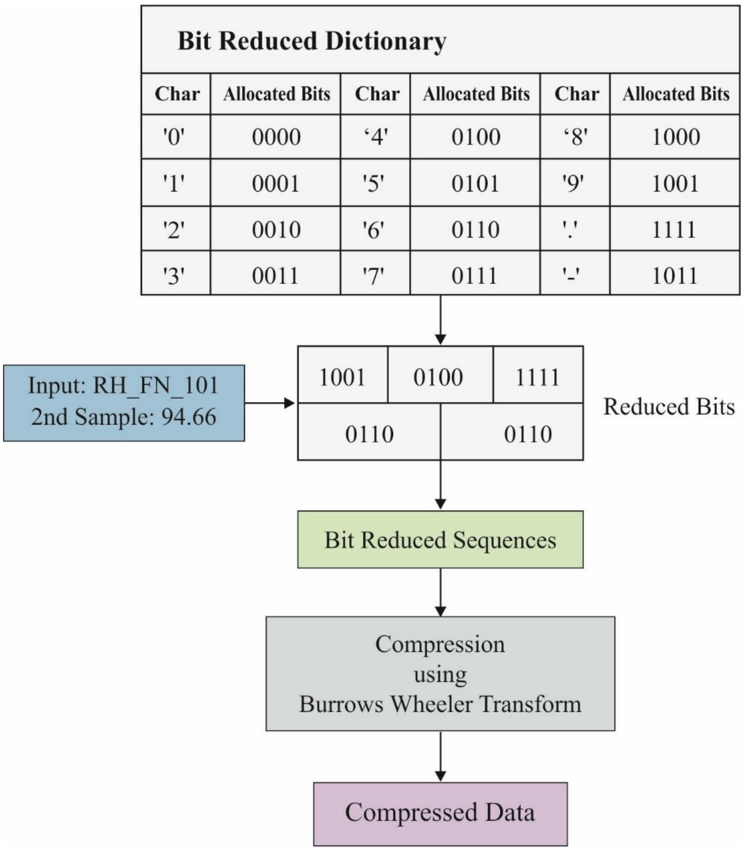


Fig. 1 Overall Process of Proposed BR-BWT model

The exclusive feature of BR-BWT approach is that it uses 4-bit codewords for all characters. The BR-BWT model requires a lower C_{bits}which stores compressed file as calculated in Eq. (1):

$$C_{bits} = \sum_{i=1}^{N} N_{CAD}(i) \tag{1}$$

where N_{CAD} shows the bit count and i represents the data instance. In specific, the bits required to save a character is around 4. Then, the bits which are essential for storing each character in BR-BWT model is estimated by Eq. (2).

$$CAD_{ch\_av} = \frac{C_{bits}}{N} \approx 4 \quad (2)$$

Eq. (2) states that, the average count of bits required for storing a character results in effective compression operation.

A. Bit Reduction Process using Codeword Allocation

The complete process of BR-BWT compression as well as decompression task and the optimal CAD is provided. Initially, the BR-BWT method retains a BR-BWT that captures the codewords of 12 characters. As the BR-BWT model is deployed for WSN data is composed of arithmetic characters and dot characters. It is mainly applied for reducing the difficulty of a model. The BR-BWT is predetermined, compression and decompression contains a BR-BWT apriori. After receiving the input sequence, the BR-BWT method applies BR-BWT and designates a codeword to them. Consequently, the final outcome of codewords are encoded and integrated for producing the compressed file with actual size which is then offered to receiver end.

Here, newly deployed BR-BWT method applies symmetrical compression in which decompression task is an opposite function of compression process. Since the BR-BWT model contains similar BR-BWT as an encoding device, there is no requirement to send extra data with the compressed file for reformation task. At the initial stage, the BR-BWT method learns the compressed file with binary codewords. Here, compressed data is divided as 4 bits where the BR-BWT is applied for codeword mapping. Once the codewords are identified, decoded characters are integrated for reconstructing actual data. In the next stage, BWT is applied to further compress the data.

B. BWT based Compression Process

In BWT, the text is assumed as blocks. An effective lossless source code is said to be the sequence of source code that obtains optimal function for each source. The exact function is constrained with the following. Suppose the class of $\{P_\theta : \theta \in \Lambda\}$ of stationary ergodic sources with definite source alphabet $\mathcal{X}$. For all $\theta \in \Lambda$, let $H_\theta(X^n)$ and $H_\theta(\mathcal{X})$ be nth order entropy and entropy rate of $P_\theta$which has been depicted as follows.

$$H_\theta(X^n) = \sum_{u^n \in \mathcal{X}^n} [-P_\theta(u^n) \log P_\theta(u^n)] \quad (3)$$

and

$$H_\theta(\mathcal{X}) = \lim_{n\to\infty} \frac{1}{n} H_\theta(X^n) \quad (4)$$

for all $\theta \in \Lambda$. The applied variable-rate lossless source coding principle for coding n-sequences from $\mathcal{X}$, for $u^n = (u_1 \dots , u_n) \in \mathcal{X}^n$, let $\ell_n(u^n)$ defines the certain length applied in the lossless description of u^n with decided coding principle. For all $\theta C \Lambda, \delta_n(\theta)$ shows the essential redundancy for coding the samples from distribution P_θ. Thus, $\delta_n(\theta)$ demonstrates the difference over the target rate for each symbol $E_\theta \ell_n(X^n)/n$ using block length-n code and better rate for all symbols $H_\theta(X^n)/n$ for coding $n - vectors$ from P_θ; hence,

$$\delta_n(\theta) = \frac{1}{n} E_\theta l_n(X^n) - \frac{1}{n} H_\theta(X^n) \quad (5)$$

The sequence of coding principles are described by the redundancy functions $\{\delta_n(\cdot)\}_{n=1}^{\infty}$, showcases a periodical min-max universal lossless source code on Λ if $\delta_n(\theta) \to 0$ for all $\theta \in \Lambda$ and rapid min-max universal lossless source code on Λ if the convergence is same in θ. These estimated bounds show the code function on sequence X^n by means of "empirical entropy" of X^n related to distribution method which is same as basic sources.

A work of unifilar, ergodic and finite-state-machine (FSM) modules are useful as provided. FSM is referred as definite alphabet $\mathcal{X}$, a limited set of states $\mathcal{S}, |\mathcal{S}|$ conditional probability score $\{p(\cdot\,|s)\}_{sC\mathcal{S}}$, and a next-state function $f: \mathcal{S} \times \mathcal{X} \to \mathcal{S}$. It is offered with FSM data source as well as initial statess_0, a conditional probability of string $u^n = u_{1,} \dots, u_n \in \mathcal{X}^n$ with s_0 that is expressed as,

$$Pr(u^n|s_0) = \prod_{i=1}^{n} p\,(u_i|s_{i-1}) \tag{6}$$

where $s_i = f(s_{i-1}, u_i)$ for every $1 \leq i \leq n.$ The FSM X sources, termed as FSM sources are defined as a subset of FSM sources in which the integer M where $i \geq M$, the M symbols $u^{i}_{i-\Lambda J+\perp}$ estimate the state s_i at time i. In FSMX sources, the set $\mathcal{S}$ is meant to be minimal suffix set of strings from $\mathcal{X}^*$ with the help of feature in which $s \in \mathcal{S}$and all $u \in \mathcal{X}$ in which $p(u|s) \neq 0$, the string su contains an individual suffix in $\mathcal{S}$. Therefore, for FSMX source, $f(s_{i-1}, \mathrm{u}_i) = suf(s_{i-1}u_i)$ for all i, where $\mathrm{suf}(su)$ is a suffix of a string accomplished by applying a symbol u for entire string s.

FSMX sources were obtained from FSM sources using recent as well as previous condition $(s_i = f(s_{i-1}, u_i))$. As a result, the drawbacks are limited, by using normalized FSMX sources, termed as finite-memory sources. For FSM, lower suffix set S of strings from $\mathcal{X}^*$ and integer M is represented as,

$$Pr\big(u^n\big|u^0_{-(M-1)}\big) = \prod_{i=1}^{n} p\,(u_i|s_{i-1}) \tag{7}$$

and $s_{i-1} = suf(u_{i-M}, u_{i-(M-1)\}} \dots, u_{i-1})$, for every $i.$ The state variables $\{s_i\}$ are defined as variable-length strings. For stationary, the symbols $X_{-M+1}, X_{-M+2}, \dots, X_0$ should to be attained from a stationary distribution on $\mathcal{X}^M$ provided by FSM is defined as,

$$Pr(u^n) = p(u^M) \prod_{i=M+\perp}^{n} p\,(u_i|s_{i-\perp}) \tag{8}$$

where $p(u^M)$ implies the stationary distribution on $\mathcal{X}^M$ promoted by FSM.

The FSM states that, there is no requirement of contexts with length k to hold conventional k symbols for a data string. The improved $|\mathcal{S}|$ scores are not arranged while $|\mathcal{S}|$ is modified by length of previous data applied in conditional distributions. The effective maximization in $|\mathcal{S}|$ results in performance degradation, since the convergence measures tends in Section V with $|S|$.In this approach, $\theta = (p(1|s): sC\mathcal{S})$ describes the distribution P_θ, and thus, $K = |\mathcal{S}|$ and $\Lambda \underline{C} \mathbb{R}^K$, with familiar $(K/2)\log n/n + O(1/n)$. In general, $K = |S|(|\mathcal{X}| - 1)$ offers the number of parameters should be defined the conditional probabilities $p(u|s)$ for every $u \in \mathcal{X}$ and values of $s \in \mathcal{S}$. The BWT is described as reversible block-sorting transform which depends upon a series of n data for permuted data series of identical symbols and single integer in {1, ...} $n\}$. Let

$$BWT_n: \mathcal{X}^n \to \mathcal{X}^n \times \{1, \dots, n\} \tag{9}$$

When n-dimensional BWT function is provided below:

$$BWT_n^{(-1)}: \mathcal{X}^n \times \{1, \dots, n\} \to \mathcal{X}^n \tag{10}$$

Where it is an inverse of BWT_n. The sequence length n is accomplished from a source argument with functional transcript that is limited as,

$$(v^n, u) = BWT(u^n) and BWT^{(-1)}(v^n, u) = u^n. \tag{11}$$

The performance of BWT_u and $BWT_{\mathbb{N}}$ represents the character and integer portions of BWT. The front BWT is processed by the development of n cyclic shifts for original data string and sorting the cyclic shifts in lexicographical fashion.

3. RESULTS

In order to investigate the performance of BR-BWT model, it is sampled on WSN datasets such as humidity as well as temperature dataset. With the application of newly developed models namely ALDC, LEC, SLZW, FELACS methods, the BR-BWT's function is associated interms of energy consumption and compression operation. The presented model is related with the classical compression approaches called as gzip, bzip2, Huffman as well as arithmetic coding. The proposed BR-BWT model has been simulated using Adruino board, sensors and Python Programming language.

A. Dataset used

SensorScope, the ecological observation WSN dataset has been applied. Firstly, temperature and humidity measures from 3 SensorScope were sampled: HES-SO Fish-Net Deployment, Le Gènèpi Deployment as well as LUCE deployment [19]. The size of a dataset may vary from 12,652 to 64,913 instances. The deployment scenarios utilize a TinyNode node which comprises a TI MSP430 microcontroller, a Xemics XE1205 radio and a Sensirion SHT75 sensor module. The relative humidity and temperature sensors are connected to a 14 bit analog to digital converter (ADC). The outcome of the ADC for raw relative humidity (raw_h) and raw temperature (raw_t) are represented by the resolutions of 12 bits and 14 bits correspondingly. The actual output raw_h and raw_t is transformed to compute h and t in percentage (%) and degree Celsius correspondingly. The dataset published on SensorScope deployments defines the physical measures h and t. Since the compression technique operates on raw_h and raw_t, the physical measures h and t are converted to raw_h and raw_t utilized the inverted versions of the conversion function earlier to data compression. The details of the dataset are provided in Table 1.

Table 1 Dataset description

Deployment name	Node ID	Symbolic name	No. of samples	Time Interval	
				From day	To day
LUCE	84	LU_84	64,913	23-Nov-2006	17-Dec-2006
HES-SO FishNet	101	FN_101	12,652	09-Aug-2007	31-Aug-2007
Le Gènèpi	20	LG_20	21,523	04-Sep-2007	03-Oct-2007

B. Performance Measures

A set of measures used to examine the performance of the proposed methods are given as follows.

$$Compressionratio = \left(\frac{No. of bitsincompresseddata}{No. of bitsinuncompresseddata}\right) \quad (12)$$

$$Compressionfactor = \left(\frac{No. of bitsin\ uncompresseddata}{No. of bitsincompresseddata}\right) \quad (13)$$

$$Spacesavings = 100 * \left(1 - \frac{No. of bitsincompresseddata}{No. of bitsinuncompresseddata}\right) \quad (14)$$

$$PCR = 100 * \left(1 - \frac{No. ofpackets\ required\ totransmitcompresseddata}{No. ofpackets\ required\ totransmituncompresseddata}\right) \quad (15)$$

$$Compression\ Rate = \frac{No. of\ bits\ in\ essential\ for\ compressed\ data}{No. of\ Characters\ in\ Sensed\ Data} \quad (16)$$

$$Power\ Savings = \frac{1 - Compression\ Rate}{Total\ No. of\ bits} * 100 \quad (17)$$

C. Discussion

Table 2 provides an analysis of the results by the BR-BWT model interms of different aspects. The table indicated that the presented BR-BWT model has achieved maximum compression efficiency. On the applied LU_84Temp, the BR-BWT model has compressed the dataset into a maximum extent with the minimum CR of 0.130375 and CF of 7.670202. On the given FN_101Temp, the BR-BWT method has compressed the dataset to a certain extent with the lower CR of 0.187275 and CF of 5.339738. On the provided LG_20Temp, the BR-BWT model has compressed the dataset into a higher extent with the least CR of 0.219614 and CF of 4.553452. On the applied LU_84RH, the BR-BWT approach has compressed the dataset into a higher extent with the minimal CR of 0.183520 and CF of 5.449003. On the projected LFN_101RH, the BR-BWT model has compressed the dataset into a greater extent with the lesser CR of 0.237570 and CF of 4.209280. On the presented LG_20RH, the BR-BWT model has compressed the dataset into a certain extent with the lower CR of 0.258909 and CF of 3.862363.

Table 2 Result Analysis of BR-BWT on various WSN dataset in terms of Compressed Size, CR, CF

Dataset	**Original Size (Bits)**	**Compressed Size (Bits)**	**Compression Ratio**	**Compression Factor**
LU_84 Temp	3135824	408832	0.130375	7.670202
FN_101 Temp	680368	127416	0.187275	5.339738
LG_20 Temp	1043032	229064	0.219614	4.553452
LU_84 RH	4096168	751728	0.183520	5.449003
FN_101 RH	696704	165520	0.237570	4.209280
LG_20 RH	1358872	351824	0.258909	3.862363

Table 3 portrayed the packet size analysis of the BR-BWT model on the applied set of dataset. The proposed BR-BWT model has reached to a minimum number of compressed packets compared to original packet sizes. On the test LU_84 Temp dataset, the BR-BWT model compresses the original packet size of 13516 bits into 1762 bits. On the test FN_101 Temp dataset, the BR-BWT method compresses the actual packet size of 2932 bits into 549 bits. On the test LG_20 Temp dataset, the BR-BWT approach compresses the original packet size of 4495 bits into 987 bits. On the test LU_84 RH dataset, the BR-BWT scheme compresses the real packet size of 17655 bits into 3240 bits. On the test FN_101 RH dataset, the BR-BWT technique compresses the actual packet size of 3003 bits into 713 bits. On the test LG_20 RH dataset, the BR-BWT approach compresses the original packet size of 5857 bits into 1516 bits.

Table 3 Result Analysis of BR-BWT on various WSN dataset in terms of Packet Size

Dataset	**Original packet size (bits)**	**Compressed packet size (bits)**
LU_84 Temp	13516.48276	1762.20689
FN_101 Temp	2932.655172	549.206897
LG_20 Temp	4495.827586	987.344828
LU_84 RH	17655.89655	3240.20689
FN_101 RH	3003.103448	713.448276

LG_20 RH	5857.206897	1516.48275

Fig. 2 provides a comparison of SS offered by the BR-BWT model and recently presented methods. The table values indicated that the S-LZW model has exhibited ineffective compression efficiency and attained minimum SS. Next, the LEC model has shown better compression than S-LZW model by achieving slightly higher SS. Also, the ALDC and FELACS models have exhibited even higher compression efficiency by attaining closer SS. At the same time, the BCAT model has tried to show effective compression performance over the earlier models. However, the presented BR-BWT model has reached to a maximum SS on all the applied dataset.

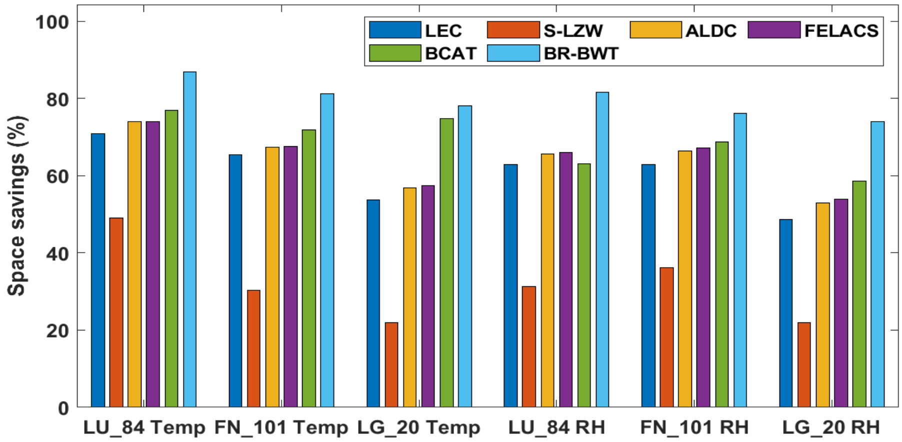
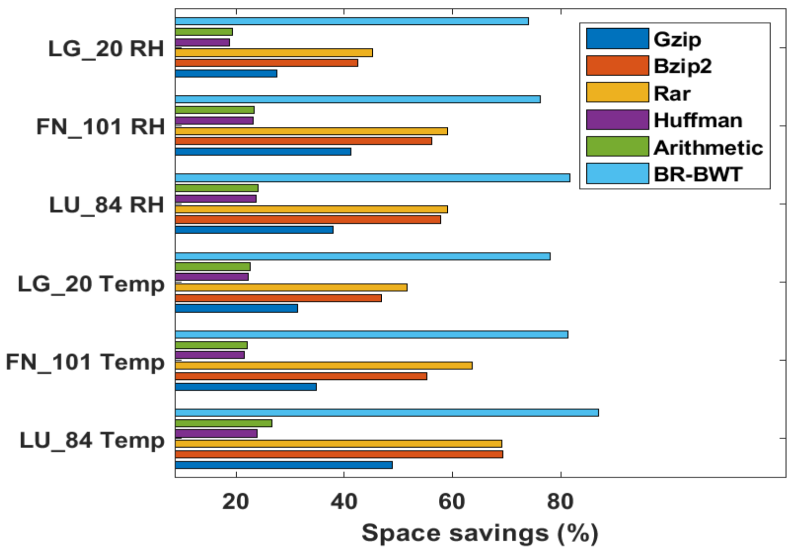


Fig. 2 Comparative analysis of BR-BWT models with recent methods in terms of Space savings

Fig. 3 offers a comparison of SS provided by the BR-BWT method with the help of conventional compression technology. The table measures pointed that the Huffman method has showcased worst compression efficiency and accomplished lower SS. Besides, the Arithmetic model has implied moderate compression when compared to Huffman model by accomplishing better SS. Additionally, the Gzip and Rar methodologies have showcased gradual compression efficiency by achieving nearer SS. Meanwhile, the Bzip2 model has attempted to exhibit efficient compression performance over the previous methods. Thus, the proposed BR-BWT approach has accomplished greater SS on all the applied dataset.



European Journal of Molecular & Clinical Medicine
ISSN 2515-8260 Volume 7, Issue 11, 2020

Fig. 3 Comparative analysis of BR-BWT models with traditional compression algorithms in terms of Space savings

Fig. 4 signifies a comparison of power saving offered by the BR-BWT method and newly projected models. The table values showed that the S-LZW model has showcased poor compression efficiency and reached lower power saving. Followed by, the LEC method has implied moderate compression when compared to S-LZW method by accomplishing maximum power saving. Furthermore, the ALDC and FELACS approaches have showed better compression efficiency by achieving closer power saving. Simultaneously, the BCAT model has attempted to exhibit productive compression performance than previous models. Therefore, the projected BR-BWT method has attained to a higher power saving on all the given dataset.

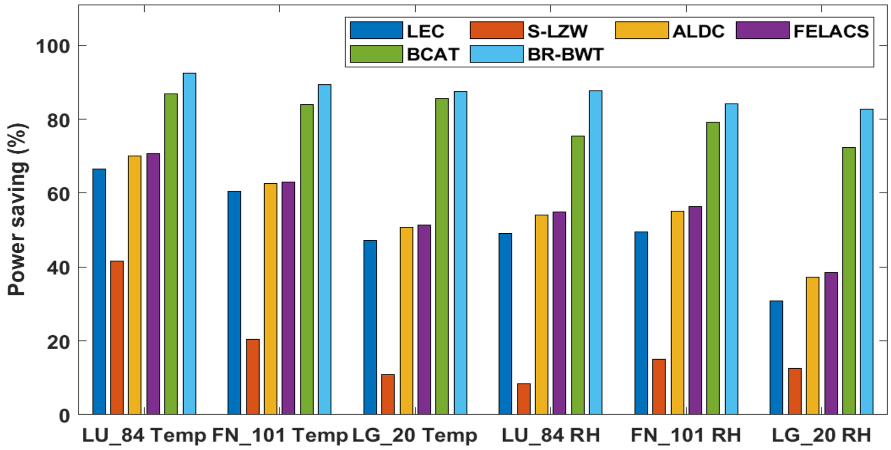


Fig. 4 Comparative analysis of BR-BWT models with compression algorithms in terms of Power saving

Fig. 5 provides a comparison of bit rate offered by the BR-BWT model and recently presented methods. The table values indicated that the S-LZW model has exhibited ineffective compression efficiency and attained maximum bit rate. Next, the LEC model has shown better compression than S-LZW model by achieving slightly lower bit rate. Also, the ALDC and FELACS models have exhibited even higher compression efficiency by attaining closer bit rate. At the same time, the BCAT model has tried to show effective compression performance over the earlier models with low bit rate. However, the presented BR-BWT model has reached to a minimum bit rate on all the applied dataset.

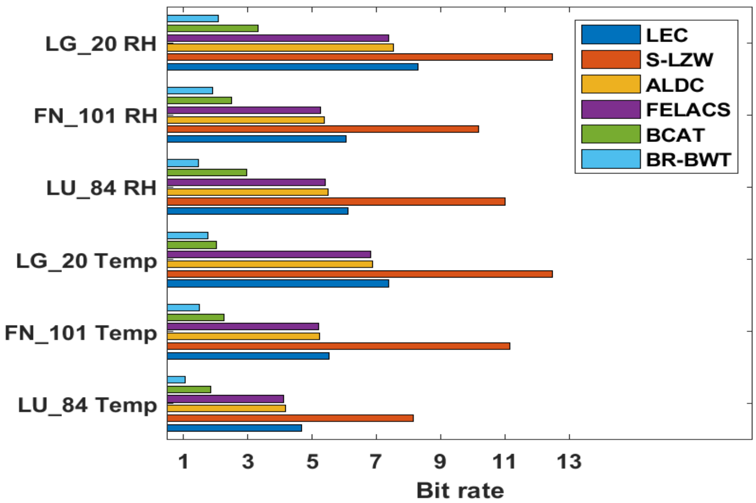


Fig. 5 Comparative analysis of BR-BWT models with existing methods in terms of Bit rate

Table 4 Result Analysis of BR-BWT for Real Time Dataset in terms of various measures

Dataset	**Original**	**Compressed**	**CR**	**CF**	**SS**	**Bit Rate**	**Power savings**
Real Data 1	2224	1328	0.59712	1.67470	40.29	4.7770	65.88
Real Data 2	1944	1480	0.76132	1.31351	23.87	6.0905	56.50
Real Data 3	2168	1032	0.47601	2.10078	52.40	3.8081	72.80

Table 4 investigates the results of the BR-BWT model on the real time dataset. On the applied real data 1, the proposed BR-BWT model compresses the 2224 bits into 1328 bits with the CR of 0.59712, CF of 1.67470, SS of 40.29%, bit rate of 4.777 and power saving of 65.88%. On the provided real data 2, the presented BR-BWT approach compresses the 1944 bits into 1480 bits with the CR of 0.76132, CF of 1.31351, SS of 23.87%, bit rate of 6.0905 and power saving of 56.50%. On the given real data 3, the presented BR-BWT model compresses the 2168 bits into 1032 bits with the CR of 0.59712, CF of 1.47601, SS of 52.40%, bit rate of 3.8081 and power saving of 72.80%.

4. CONCLUSIONS

This study has established a novel Bit Reduction with Burrows Wheeler Transform named as BR-BWT based data compression model in WSN. The proposed BR-BWT technique has performed data encoding in 2 ways like bit reduction using codeword allocation as well as BWT based encoding task. At the initial stage, previous codeword allocation process is conducted to compute the codeword for each character present in the WSN data. Followed by, the BWT based compression process is performed to minimize the volume of data transmission. For performance validation of the BR-BWT technique, the actual WSN dataset is sampled and the final outcomes are defined under various factors. The experimental outcome stated the superior characteristics of the BR-BWT model compared to traditional and recently proposed models. In future, the performance of the proposed model can be improvised by the use of other dictionary based coding techniques.

5. REFERENCES

[1] J. Capon, “A probabilistic model for run-length coding of pictures,” IRE Trans. Inf. Theory, vol. 100, pp. 157–163, 1959.
[2] D. A. Huffman, “A Method for the Construction of Minimum-Redundancu Codes,” Proc. IRE, vol. 40, no. 9, pp. 1098–1102, 1952.
[3] T. A. Welch, “A technique for high-Performance Data Compression,” IEEE, pp. 8–19, 1984.
[4] T. Schoellhammer, B. Greenstein, E. Osterweil, M. Wimbrow, and D. Estrin, “Lightweight temporal compression of microclimate datasets,” UCLA Cent. Embed. Netw. Sens., 2004.
[5] E. P. Capo-Chichi, H. Guyennet, and J. M. Friedt, “K-RLE: a new data compression algorithm for wireless sensor network,” in in Proceedings of the 3rd International Conference on Sensor Technologies and Applications, (SENSORCOMM’09), 2009, pp. 502– 507.

[6] Sadler and Martonosi, “Data compression algorithms for energy-constrained devices in delay tolerant networks,” in In: Proceedings of the 4th international conference on embedded networked sensor systems – SenSys ’06, 2006, pp. 265–78.
[7] M. Francesco and M. Vecchio, “An efficient lossless compression algorithm for tiny nodes of monitoring wireless sensor networks,” Comput. J., vol. 52, no. 8, pp. 969–987, 2009.
[8] Y. Liang and Y. Li, “An efficient and robust data compression algorithm in wireless sensor networks,” IEEE Commun. Lett., vol. 18, no. 3, pp. 439–442, 2014, doi: 10.1109/LCOMM.2014.011214.132319.
[9] J. G. Kolo, S. A. Shanmugam, D. W. G. Lim, L.-M. Ang, and K. P. Seng, “An Adaptive Lossless Data Compression Scheme for Wireless Sensor Networks,” J. Sensors, 2012, doi: 10.1155/2012/539638.
[10] A. Kiely, M. Xu, W.-Z. Song, R.Huang, and B. Shirazi, “Adaptive Networks, linear filtering compression on realtime sensor networks,” Comput. J., vol. 53, no. 10, pp. 1606–1620, 2010.
[11] Y. Liang and W. Peng, “Minimizing energy consumptions in Wireless Sensor Networks via two-modal transmission,” Comput. Commun. Rev., vol. 40, no. 1, pp. 13–18, 2010.
[12] J. Gana, S. A. Shanmugam, D. Wee, G. Lim, and L. Ang, “Fast and efficient lossless adaptive compression scheme for wireless sensor networks q,” Comput. Electr. Eng., vol. 41, pp. 275–287, 2015.
[13] C. Alippi and G. Anastasi, “An adaptive sampling algorithm for effective energy management in wireless sensor networks with energy-hungry sensors,” IEEE Trans. Instrum. Meas., vol. 59, no. 2, pp. 335–344, 2010, doi: 10.1109/TIM.2009.2023818.
[14] A. L. L. Aquino, O. S. Junior, A. C. Frery, E. L. Albuquerque, and R. A. F. Mini, “MuSA: Multivariate Sampling Algorithm for Wireless Sensor Networks,” IEEE Trans. Comput., vol. 63, no. 4, pp. 968–978, 2014, doi: 10.1109/TC.2012.229.
[15] A.L.L. Aquino, E.F. Nakamura, Data centric sensor stream reduction for real-time applications in wireless sensor networks, Sensors 9 (12) (2009) 9666–9688.
[16] T. Shu, M. Xia, J. Chen, C. de Silva, An energy efficient adaptive sampling algorithm in a sensor network for automated water quality monitoring, Sensors 17 (11) (2017) 2551.
[17] M. Taghout, A.K. Chorppath, T. Waurick, F.H.P. Fitzek, On the design of a joint compressed sensing and network coding framework, in: European Wireless 2018; 24th European Wireless Conference, 2018, pp. 113–119.
[18] P. Sun, L. Wu, Z. Wang, M. Xiao, Z. Wang, Sparsest random sampling for cluster-based compressive data gathering in wireless sensor networks, IEEE Access 6 (c) (2018) 36383–36394.
[19] https://lcav.epfl.ch/page145180-en.html.
[20] Sujitha, B., Parvathy, V. S., Lydia, E. L., Rani, P., Polkowski, Z., & Shankar, K. (2020). Optimal deep learning based image compression technique for data transmission on industrial Internet of things applications. Transactions on Emerging Telecommunications Technologies, e3976.
[21] Uthayakumar, J., Elhoseny, M., & Shankar, K. (2020). Highly Reliable and Low-Complexity Image Compression Scheme Using Neighborhood Correlation Sequence Algorithm in WSN. IEEE Transactions on Reliability.
[22] Geetha, K., Anitha, V., Elhoseny, M., Kathiresan, S., Shamsolmoali, P., & Selim, M. M. (2020). An evolutionary lion optimization algorithm ‐ based image compression technique for biomedical applications. Expert Systems, e12508.
[23] Krishnaraj, N., Elhoseny, M., Thenmozhi, M., Selim, M. M., & Shankar, K. (2020). Deep learning model for real-time image compression in Internet of Underwater Things (IoUT). Journal of Real-Time Image Processing, 17(6), 2097-2111.