

A Wavelet-Based Sampling Algorithm for Wireless Sensor Networks Applications

Andre L. L. Aquino
Dep. Computer Science
Federal University of Ouro
Preto
Ouro Preto, MG, Brazil
alla@iceb.ufop.br

Ricardo A. R. Oliveira
Dep. Computer Science
Federal University of Ouro
Preto
Ouro Preto, MG, Brazil
rrabelo@gmail.com

Elizabeth F. Wanner
Dep. Mathematics
Federal University of Ouro
Preto
Ouro Preto, MG, Brazil
efwanner@gmail.com

ABSTRACT

This work proposes and evaluates a sampling algorithm based on wavelet transforms with Coiflets basis to reduce the data sensed in wireless sensor networks applications. The Coiflets basis is more computationally efficient when data are smooth, which means that, data are well approximated by a polynomial function. As expected, this algorithm reduces the data traffic in wireless sensor network and, consequently, decreases the energy consumption and the delay to delivery the sensed information. The main contribution of this algorithm is the capability to detect some event by adjusting the sampling dynamically. In order to evaluate the algorithm, we compare it with a static sampling strategy considering a real sensing data where an external event is simulated. The results reveal the efficiency of the proposed method by reducing the data without losing its representativeness, including when some event occurs. This algorithm can be very useful to design energy-efficient and time-constrained sensor networks when it is necessary to detect some event.

Categories and Subject Descriptors

C.2.4 [Computer-communication networks]: Distributed Systems

Keywords

Wavelets, sampling algorithms, wireless sensor network

1. INTRODUCTION

A Wireless Sensor Network (WSN) consists of spatially distributed autonomous devices which cooperatively monitor physical or environmental conditions, such as temperature, sound, vibration, pressure, motion or pollutants, at different locations [1]. In WSN applications, physical variables, such as temperature and luminosity, can be monitored continuously along the network operation. The data set representing these physical variables can be referred as data stream [10] – or sensor stream, in the WSNs context.

Despite the potential of their applications, the WSNs present some resource constraints such as low computational power, re-

duced bandwidth and limited energy source. As a result, the algorithms which process data in the sensor nodes and the quantity of temporary information that can be stored need to be limited. However, the data transmission for the offline storage, query and data analysis in a central node is an expensive procedure in WSNs, since wireless communication consumes a large amount of energy. Since the sensor nodes can process locally, part of the data processing can be performed through some sensor stream reduction algorithms. Sampling, sketching, or wavelet stream strategies, to name a few, are example of sensor stream reduction algorithms [2, 10].

Considering the aspects commented above, this work proposes and evaluates a wavelet-based sampling algorithm for WSN applications. This algorithm uses the wavelet transform with coiflets basis functions to allow a dynamical sampling [8]. This dynamical sampling permits a more detailed event detection when compared with other wavelet basis functions [8] (Haar or Daubechies). If data can be interpolated by linear or quadratic functions, then Haar and Daubechies bases are efficient. However, considering a polynomial interpolation of order $2 < n < 12$, these bases are not efficient. In this case, the usage of Coiflets basis are highly recommended. Using the proposed sampling solution, it is possible to reduce the data traffic, delay and energy consumption in a WSN. The main contribution of this algorithm is the capability to detect some event by adjusting dynamically the sampling. In order to evaluate the proposed algorithm, we compare it with a static sampling strategy considering a real sensor data where an external event is simulated. The results reveal the efficiency of the proposed method by reducing the data without losing its representativeness, even if some external event occurs.

This work is organized as follows. Section 2 presents related work about data stream reduction techniques. Section 3 discusses the wavelet transform and its main properties. Section 4 shows the problem description and the reduction strategy adopted in this work. Numerical results are given by Section 5, and Section 6 concludes this study and suggests future work.

2. RELATED WORK

Considering the data stream related work, the contributions usually focus either in improving stream algorithms or in applying the data stream techniques to specific scenarios. However, there are some works which consider WSNs as a distributed databases in which some functions (e.g. maximum, minimum, and average) can be computed in a distributed manner [4, 7].

There are some sampling techniques which perform a “intelligent” sensor stream reduction by choosing the best elements that represent the behavior of the environment sensed. An example of a smart choice is the sampling, in each class of the histogram, built

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SAC'10 March 22-26, 2010, Sierre, Switzerland

Copyright 2010 ACM 978-1-60558-638-0/10/03 ...\$10.00.

during the sensing in online fashion. Another sampling strategy considers multivariate sensor stream reduction and, in this case, the analysis of multivariate data are used to improve the multivariate sampling [2, 6].

Dijkerman and Mazumdar [5] proposes multiresolution stochastic models of discrete wavelet coefficients as an approximations for the original time process. These models are simple due to the strong decorrelation of the wavelet transform. Assuming that a wireless channel has a temporal variation which can be modeled by wavelet, Tsatsanis and Giannakis [12] use the wavelet transforms to model systems of identification and to an efficient control of fading in wireless channels.

A more recent study presented by Martone [9] proposes new detectors for direct-sequence code-division multiple-access (CDMA) signals which outperform well-known approaches in rapidly fading multipath channels. In that work, the author proved that it is possible to represent the channel in a reduced-order dimensional space by mapping the scattering function of the multipath channel to its decomposition. As a result, it is possible to obtain an approach which is effective in fast fading environments, such as those practically found in macrocell wireless communication applications.

Regarding data reduction strategies for WSNs, current works use data fusion, aggregation, compression, or correlation techniques to help save energy and to reduce the packet delay. The closest approach to what we are showing here is the adaptive sampling, which means that the sampling strategy is modified by the phenomenon modifications. The objective of the adaptative sampling is to improve accuracy, to identify data correlation and to eliminate redundancy [11, 13].

3. WAVELET BACKGROUND

Let V'_j be a sequence of closed subspaces of $L(\mathbb{R}^2)$ and $f(x) \in L(\mathbb{R}^2)$ be the signal observed. Each V'_j represents the successive approximation of original signal, considering a resolution of 2^j . Smooth resolution occurs when the lowest value of j is used.

The details of projection between 2^j and 2^{j-1} , denoted by W_j , is defined by

$$W_j \oplus V'_j = V'_{j-1},$$

where \oplus denotes the direct sum of two vectorial spaces. In this way, V'_j can be decomposed through the direct sum of subspaces,

$$V'_j = W_{j+1} \oplus W_{j+2} \oplus \dots \oplus W_J \oplus V'_J,$$

where $j < J$ and all subspaces are orthogonal.

The discrete wavelet transform is used to analyze discrete data. Discrete filters are defined to choose the frequency levels present in data which vary in a temporal scale. Two set of functions are applied: the scale functions ($\phi(t)$); and the wavelet functions ($\psi(t)$). Both are related to small and high pass filters called, respectively, wavelet and scale vectors.

In this way, $f(x)$ can be approximated by the following expansion

$$f(x) = \sum_n s_{i_0}[n] \psi_{i_0,n}(x) + \sum_{i=i_0}^{i_1} w_i[n] \phi_{i,n}(x)$$

where $s_{i_0}[n] \in V'_j$ are the scale coefficients,

$$s_{i_0}[n] = \int f(x) \psi_{i_0,n}(x) dx$$

and $w_i[n] \in W_i$ are the wavelet coefficients,

$$w_i[n] = \int f(x) \phi_{i,n}(x) dx$$

3.1 Wavelet properties

The wavelet transform of a function $f(x)$ is a two-dimensional function $\gamma(s, \tau)$. The variables s and τ are the new dimensions, scale and translation, respectively.

The wavelet transform is composed by functions $\phi(t)$ and $\psi(t)$ which the admissibility condition $\int \psi(t) dt = 0$ holds. This condition states that the wavelet has a compact support, since the biggest part of its value is restricted to a finite interval, which means that it has an exact zero value outside of this interval. Besides, $\psi_{i,k}(t) = 2^{\frac{j}{2}} \psi(2^i t - k)$, for the dilated and translated versions of $\psi(t)$. These both cases characterize the wavelet spacial localization property.

Another property is the wavelet smoothness. Consider the expansion of $\gamma(s, \tau)$, around $\tau = 0$, in a Taylor serie of n order,

$$\gamma(s, 0) = \frac{1}{\sqrt{s}} \left[\sum_{p=0}^n f^{(p)}(0) \int \frac{t^p}{p!} \psi(t/s) dt + O(n+1) \right],$$

where $f^{(p)}$ is the p^{th} derivative of f and $O(n+1)$ represents the rest of the expansion. A wavelet, $\psi(t)$, has L null moments if

$$\int_{-\infty}^{\infty} t^k \psi(t) dt = 0,$$

to $0 \leq k < L$. If

$$M_k = \int t^k \psi(t) dt,$$

we have

$$\gamma(s, 0) = \frac{1}{\sqrt{s}} \left[\frac{f^{(0)}(0)}{0!} M_0 s^1 + \dots + \frac{f^{(n)}(0)}{n!} M_n s^{n+1} + O(s^{n+2}) \right]$$

The vanishing moments are

$$M_n(0, l) = 0, \text{ for } l = 0, 1, \dots, L-1$$

where L is the number of moments of the wavelet. From the admissibility condition, we have that the 0^{th} moment, M_0 , is equal to zero. If the other M_n moments are zero, then $\gamma(s, \tau)$ will converge to a smooth function $f(t)$. So, if $f(t)$ is described by a polynomial function of degree up too $(L-1)$, the term $O(s^{n+2})$ will be zero and small values will appear as a linear combination of ψ in the function $\gamma(s, \tau)$.

The regularity degree and decreasing rate of wavelet transforms are related to its number of null moments. This property is important to infer the approximations property in multiresolution space. When a wavelet has various null moments, there will be coefficients with low values. In regions where $f(x)$ is a smooth function, the wavelet coefficients with thin scales in $f(x)$ are null.

3.2 Coiflets basis

The periodic wavelet transforms are applied to the limited interval of functions. In order to apply the periodic transform, it is necessary to consider that the limits of target function are repeated. With the goal of avoiding this condition, the coiflets basis has the null moments property in its scale function. The Coiflets basis is a extension of the Daubechies, having the following properties

$$\begin{cases} \int \psi(x) dx = 1 \\ \int x^l \psi(x) dx = 0, \quad l = 0, 1, \dots, L-1 \\ \int x^l \phi(x) dx = 0, \quad l = 0, 1, \dots, L-1 \end{cases}$$

Due to these properties, ψ and ϕ have null moments. The function ϕ is smoother and more symmetric than the ψ considered in Daubechies family. Moreover, this function is a better approximation of polynomial functions having interpolation properties. Considering $2L$ the number of coiflets moments and $f(x)$ in $[p, q]$, we

have

$$\int_p^q f(x)\phi(x) = \int_p^q f(0) + f'(0) + \dots + \frac{f^{2L-1}(0)x^{2L-1}}{(2L-1)!} + \dots \approx f(0)$$

Let $f(x)$ be a polynomial function of order $p \leq 2L - 1$. Then, there is $g(x)$ with the same degree so that

$$f(x) \approx \sum_{\tau} g(\tau)\phi_{s,\tau}(x).$$

This property states that, during the transformation, some terms can be obtained, directly or sampled, considering the following approximation error

$$\left\| f(x) - \sum_{\tau} g(\tau)\phi_{s,\tau}(x) \right\| = O(2^{s \cdot 2L}), \quad (1)$$

This error has the δ Dirac property $\int f(x)\delta(x)dx = f(0)$ only if it has infinite null moments. Due to this property, the coefficients in $\gamma(s, \tau)$ are sparse representations of $f(x)$ and, consequently, only few coefficients are necessary to approximate $f(x)$. With this, the scale coefficients in $\gamma_{s,\tau}$ can be approximated by sampling of $f(x)$, so that

$$\gamma(s, \tau) = f(2^{-(s)}\tau) + O(2^{s \cdot 2L}).$$

Considering the L -th order coiflets basis, it is possible to define a fast algorithm, using the low error property, when the function $f(t)$ is a smooth function. Each element in $V'_j[t]$ can be approximated by

$$V'_j[t] = 2^{-\frac{t}{2}} f(2^i t)$$

and each approximated wavelet coefficient, $W_t[t]$ at scale 2^i , are computed as

$$W_j[t] = \sum g[2t - n]V'_{j+1}[t].$$

4. PROBLEM DESCRIPTION

Based on Aquino et.al. [3], a WSN system can be represented by the diagram shown in Figure 1. In this diagram, \mathcal{N} denotes the environment and the process to be measured, P is the phenomena of interest, with V^* their space-temporal domain. If true, complete and uncorrupted observation was possible, a set of ideal rules (R^*) can be devised and it leads to an ideal decision set (D^*). With these characteristics, we consider V^* as data stream.

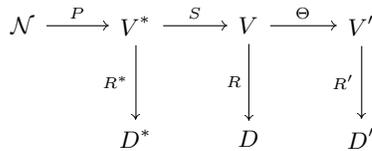


Figure 1: Representation of a WSN system, where the ideal ($\mathcal{N} \rightarrow V^* \rightarrow D^*$), sensed ($\mathcal{N} \rightarrow V^* \rightarrow V \rightarrow D$) and reduced ($\mathcal{N} \rightarrow V^* \rightarrow V \rightarrow V' \rightarrow D'$) behaviors are shown

Instead of the ideal situation, we have a set of n sensors $S = (S_1, \dots, S_n)$, in which each sensor reads the phenomena and produces a report in the domain V_i , $1 \leq i \leq n$. Using these information, we have a set of rules (R) which leads to a set of decisions (D). Therefore, differently from V^* , V is a sensor stream.

Require: Vector V and filters h and g of Coiflets basis;

Ensure: Reduced function V'_j ;

```

1: Generate filters  $h(i)$  and  $g(i)$ ;
2: for  $t \in [0, (|V|/2 - 1)]$  do
3:    $u = 2t + 1$ ;
4:    $V'_{jt} = g_1 V_{u+1}$ ;
5:   for  $n \in [1, (|h| - 1)]$  do
6:      $u = u - 1$ ;
7:     if  $u \leq 0$  then
8:        $u = |V| - 1$ ;
9:     end if
10:     $V'_{jt} = V'_{jt} + g_{n+1} V_{u+1}$ ;
11:  end for
12:   $V'_{j(t+1)} = V'_{jt}$ ;
13: end for
  
```

Figure 2: Pseudo-code of the wavelet-based sampling algorithm.

Dealing with V may be too expensive in terms of, for instance, power, bandwidth and computer resources usage. In most situations, since the level of redundancy is not negligible, it is necessary to reduce the volume of these information. Sensor stream reduction techniques, denoted by Θ , transform the complete domain V into a small one V' . New rules which use V' , denoted by R' , lead to a set of decisions D' .

Considering this model, the problem addressed in this work can be stated as follows:

Problem statement: *What is the impact over the decisions D , when we use a Θ wavelet-based sampling reduction over V generated by S ?*

For addressing this problem, the scope of this work consider a WSNs temperature monitoring application with following sensed behavior,

$$\mathcal{N} \rightarrow V_t^* \xrightarrow{S_m} \bigcup_{24h} \{V_t\}_{4min} \rightarrow D,$$

where V_t^* is the ideal temperature sensed, S_m is a real sensor node TMote Sky¹, and $\bigcup_{24h} \{V_t\}_{4min}$ represent the sensing data gathering at intervals of approximately four minutes during a day.

Before applying the reduction, it is important to define the stream item which will be used in Θ reduction. In the sensed behavior we have the item stream

$$V = \bigcup_{24h} \{V_t\}_{4min} = V^1, V^2, \dots, V^{360},$$

where $V^i \in \mathbb{R}$, $-100 \leq V^i \leq 100$. The stream item V is processed using 360 readings per day, every 4 minutes. Then, the reduced behavior is

$$\mathcal{N} \rightarrow V_t^* \rightarrow V_t \xrightarrow{\Theta} V'_t \rightarrow D,$$

where Θ is the wavelet-based sampling reduction algorithm detailed in Figure 2. In this algorithm, h and g are the discrete forms of ψ and ϕ respectively; V'_j represents the reduced data, in which j represents the resolution, i.e., the sampling rate scale, for example, 2^j , 4^j and so forth.

If M represents the number of decomposition level, the total number of operations in the data vector with size $|V|$ has a num-

¹Some information about TMote Sky can be found in http://www.capsil.org/capsilwiki/index.php/TELOSB/TMote_Sky. Unfortunately, the TMote project was discontinued.

ber of operations of order $O(M|V|\log|V|)$. The result of wavelet transform is the signal decomposition in different subspaces, or sub-bands, which different resolutions in time and frequency. Considering the temporal series, they are decomposed in other series which compose the original one.

Finally, the decision set D considered is the data representativeness including when some external event occurs, i.e., after the reduction will be possible to detect the presence of any event. This decision is the approximation rate considered by the proposed algorithm and can be calculated directly from Equation 1.

The main aspects of the proposed sensor stream algorithm can be summarized as follows: (i) the stream item considered is the sensor data buffered V ; (ii) the online behavior of the algorithm considers one pass in V but the information of the oldest processing is not used; and (iii) considering the data representativeness as the more important aspect in our applications, the approximation rate is calculated from Equation 1 through the decision D .

5. NUMERICAL RESULTS

In this section, we evaluate the impact of the data stream wavelet algorithm Θ to reduce the data sensed V . In order to tackle the problem stated in Section 4, we consider two cases: scenario with and without an event. Figure 3 illustrates the scenario without event.

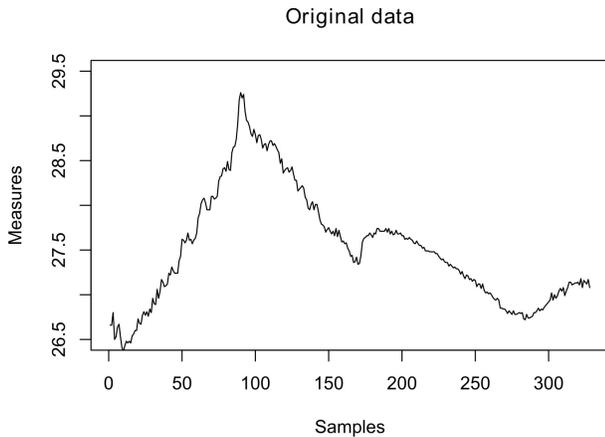


Figure 3: Data gathering without event.

Figure 4 illustrates the scenario with event. To simulate this event, we cause a random perturbation in V so that the values change drastically.

Initially we consider the wavelet transform with $j = 2$, using only the V_2' values in which the samples are performed in 2^2 steps, and it is based on coiflets filters and $|V'| = |V|/4$. In order to compare our strategy we use a simple static sampling with the same steps, i.e., the sampling considers 2^2 static steps instead the coiflets filters. Figure 5 shows the results when there is not any external event. The results were sampled of data showed in Figure 3 and the proposed strategy is represented by continuous line which indicates $|V_j'| \approx |2^{i/2}V_i|$, where $2^{i/2}$ is the normalization factor at the $\psi_{i,k}(t) = 2^{i/2}\psi(2^i - k)$.

In a different way, as showed in Figure 6, when we consider the presence of some event (Figure 4) the wavelet-based sampling is able to detect it. This occurs because the coiflets transform is used here to detect the changes in the sampling. It is important to highlight that the error in booth cases, Figures 5 and 6, can be obtained from Equation 1.

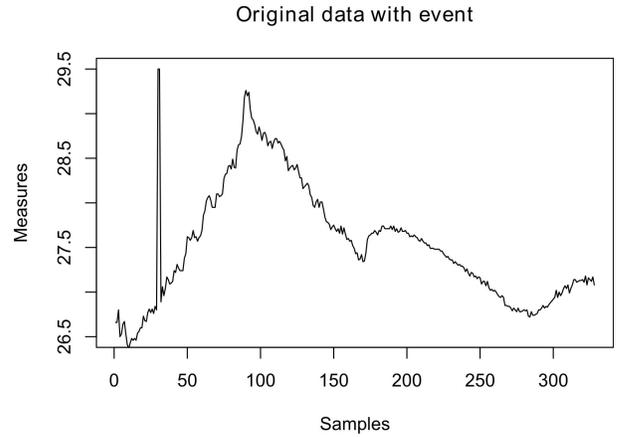


Figure 4: Data gathering with event.

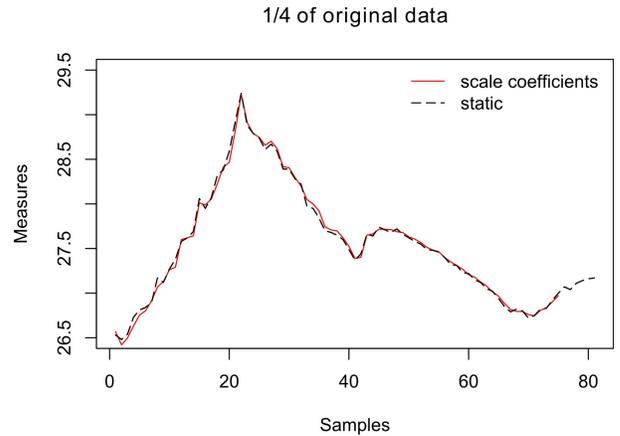


Figure 5: Sampling without event presence. The reduction is 1/4 of original data.

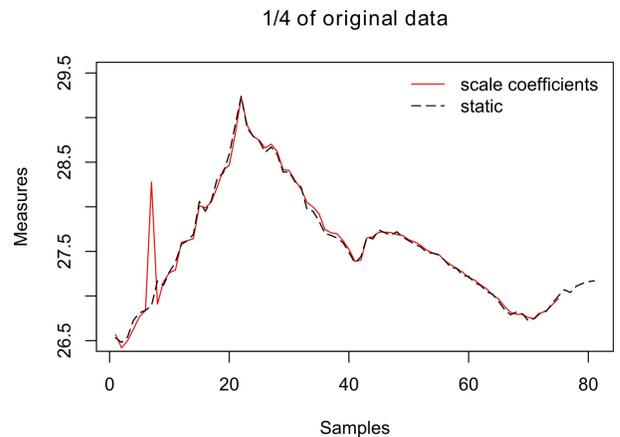


Figure 6: Sampling with event presence. The reduction is 1/4 of original data.

Considering the wavelet transform with other values to j , if the reduction is increased the representativeness is decreased. This fact can be illustrated by Figure 7, in which the samples performed in 2^3 steps are based on coiflets filters and $|V'| = |V|/8$. Again we use a simple static sampling with the same steps and, in the same way,

the wavelet-based sampling is able to detect the external event.

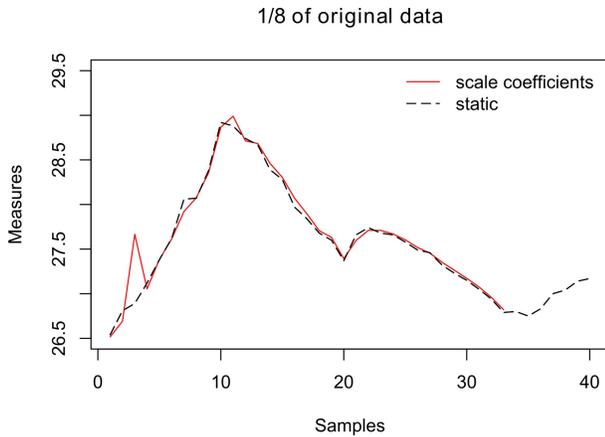


Figure 7: Sampling with event presence. The reduction is 1/8 of original data.

It is important to highlight that through the reduction technique some WSN requirements, such as the energy consumption, delay, and data lost, are achieved. These requirements can be better explored if the data is reduced.

6. CONCLUSIONS AND FUTURE WORK

WSNs are energy constrained, and the extension of their lifetime is one of the most important issues in the design of such networks. Usually, these networks collect a large amount of data from the environment. In contrast to the conventional remote sensing — based on satellites which collect large images, sound files, or specific scientific data — WSNs tend to generate a large amount of sequential small and tuple-oriented data from several nodes, which constitutes data streams.

In this work, a wavelet-based sampling algorithm is proposed and evaluated. This algorithm uses the wavelet coefficients to adjust the sampling performed. The main contribution of this algorithm is the capability of event detection by adjusting dynamically the sampling rate. The results reveal the efficiency of the proposed method by reducing the data without losing its representativeness, including when some event occur.

However, some important sensor stream algorithm characteristics can be highlighted: (i) the stream item definition; (ii) the online behavior by considering one pass over V ; and (iii) based on Coiflets wavelet properties, the approximation rate can be calculated.

As future work, we intend to apply the proposed method to process sensor streams along the routing task and in clustered networks. Thus, not only the data from a source can be reduced, but also similar data from different sources can be reduced, resulting in an efficient WSN. However, we intend to use the information of the oldest processed data to increase the sampling performed.

Acknowledges

This work is partially supported by the Brazilian National Council for Scientific and Technological Development (CNPq) under the grant number 477292/2008-9. The conference participation is partially supported by the Research Foundation of the State of Minas Gerais (FAPEMIG).

7. REFERENCES

- [1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. A survey on sensor networks. *IEEE Communications Magazine*, 40(8):102–114, August 2002.
- [2] A. L. L. Aquino, C. M. S. Figueiredo, E. F. Nakamura, L. S. Buriol, A. A. F. Loureiro, A. O. Fernandes, and C. N. C. Junior. Data stream based algorithms for wireless sensor network applications. In *21st IEEE International Conference on Advanced Information Networking and Applications (AINA'07)*, pages 869–876, Niagara Falls, Canada, May 2007. IEEE Computer Society.
- [3] A. L. L. Aquino, C. M. S. Figueiredo, E. F. Nakamura, A. C. Frery, A. A. F. Loureiro, and A. O. Fernandes. Sensor stream reduction for clustered wireless sensor networks. In *23rd ACM Symposium on Applied Computing 2008 (SAC'08)*, pages 2052–2056, Fortaleza, Brazil, March 2008. ACM.
- [4] M. Cammert, J. Kramer, B. Seeger, and S. Vaupel. A cost-based approach to adaptive resource management in data stream systems. *IEEE Transactions on Knowledge and Data Engineering*, 20(2):202–215, February 2008.
- [5] R. W. Dijkerman and R. R. Mazumdar. Wavelet representations of stochastic processes and multiresolutionstochastic models. *IEEE Transactions on Signal Processing*, 42(7):1640–1652, July 1994.
- [6] O. S. Junior, A. L. L. Aquino, R. A. F. Mini, and C. M. S. Figueiredo. Multivariate reduction in wireless sensors networks. In *IEEE Symposium On Computers and Communications (ISCC'09)*, Sousse, Tunisia, July 2009. IEEE Computer Society.
- [7] X. Lian and L. Chen. Efficient similarity search over future stream time series. *IEEE Transactions on Knowledge and Data Engineering*, 20(1):40–54, January 2008.
- [8] S. Mallat. *A Wavelet Tour of Signal Processing (Wavelet Analysis & Its Applications)*. Academic Press-Elsevier, Los Angeles, 1998.
- [9] M. M. Martone. Wavelet-based separating kernels for array processing of cellular ds/cdma signals in fast fading. *IEEE Transactions on Communications*, 48(6):979–995, June 2000.
- [10] S. Muthukrishnan. *Data Streams: Algorithms and Applications*. Now Publishers Inc, Hanover, MA, USA, January 2005.
- [11] E. F. Nakamura, A. A. F. Loureiro, and A. C. Frery. Information fusion for wireless sensor networks: Methods, models, and classifications. *ACM Computing Surveys*, 39(3):9/1 – 9/55, April 2007.
- [12] M. K. Tsatsanis and G. B. Giannakis. Principal component filter banks for optimal multiresolution analysis. *IEEE Transactions on Signal Processing*, 43(8):1766–1777, August 1995.
- [13] K. Yuen, B. Liang, and B. Li. A distributed framework for correlated data gathering in sensor networks. *IEEE Transactions on Vehicular Technology*, 57(1):578–593, January 2008.