2016 IEEE 41st Conference on Local Computer Networks

Wavelet–Based Analysis of Interference in WSNs

Aikaterini Vlachaki, Ioanis Nikolaidis and Janelle Harms
Department of Computing Science,
University of Alberta,
Edmonton, AB, Canada
E-mail: {vlachaki, nikolaidis, janelleh } @ualberta.ca

***Abstract*—Motivated by the computational, bandwidth and energy restrictions of wireless sensor network nodes and their need to, collectively, determine the presence of exogenous interference that could impair their communication, we consider schemes that could support the task of interference classification as a first step towards interference mitigation strategies. In particular, we examine the effectiveness of the Discrete Wavelet Transform (DWT) to communicate to other nodes the state of the channel, as sampled by a node, in a compressed, denoised form. We examine the suitability of different wavelet filters and thresholding methods in order to: (a) preserve key features of the interference, (b) denoise the noisy interference samples, and (c) reduce the amount of information that needs to be communicated to describe the interference.**

I. Introduction

In this paper, we explore whether wavelet compression is an effective means to convey sampled background noise information collected by wireless nodes. The motivating application is that of distributed decision-making by wireless sensor nodes regarding the state of the channel with respect to the presence (or not) of interference on the channel. The nodes could, subsequently, adopt mitigation strategies particular to the interference at hand. Here, we are only concerned with developing a low overhead means to communicate the sampled background noise information, among the nodes in a local wireless network. The assumption is that the nodes are battery–powered and any additional processing and transmissions, beyond what is needed by the applications, should be minimized.

Our previous work in this area, as well as that of a number of other authors, suggested that there are a few particular classes of interference [1]. In [2] we identified five classes which we name: quiet (q), quiet-with-spike (qs), quiet-with-rapid-spikes (qrs), high-end-level (hl), and shifting-mean (sm). Note that the main difference between qs and qrs is that qs are seemingly random impulses whereas qrs has strong periodic characteristics. A receiver at the mercy of qrs, sm, and hl is likely to be on occasion unable to properly decode a packet received during high level incursions of the interference. Note that, as in our previous work, we assume an inexpensive approach to sampling the channel, i.e., by collecting RSSI (Received Signal Strength Indicator) measurements, as they are reported by the node's RF front-end.

In order to reach an agreement (or not) among nodes in a network regarding the class of interference that is present, there exist two main strategies: (a) *local classification:* where each node samples the background noise and classifies it independently of the rest, sending its classification result to the rest, or (b) *global classification:* where each node sends the entire time series of sampled background noise to the remaining nodes, allowing each node to perform a comparison, and if deemed similar, run a classifier. Option (a) entails the risk that, even if the class inferred is the same, two or more nodes may be seeing a completely different temporal pattern e.g. a different phase of a qrs-like time series. Option (b) allows a thorough sample-by-sample analysis, e.g., via cross-correlation calculation of the time series from different nodes, but the transmission and reception energy cost involved to collect the samples by all nodes is prohibitive for battery-powered wireless nodes.

Our earlier work on characterizing the agreement between nodes with respect to the interference seen on the channel [3] was performed assuming case (b), i.e., complete access to the time series as sampled by each node, and was neither distributed nor energy-efficient. In this paper we still adopt strategy (b) but use an intermediate (compressed) representation of the sampled data via a Discrete Wavelet Transform (DWT). DWT is used for its ability to both perform denoising and compression at the same time. The denoising eliminates the effect of small noise components while retaining the overall trend and pattern of the background noise. Note that the application of DWT is a delicate matter when we wish to have the noise patterns "survive" somehow the transformation but at the same time we wish to eliminate some of the less helpful (from the point of deciding the class) facets of noise. Hence, in our work we examined various DWT alternatives and their effects.

Note that for strategy (b), a node has to store its own time series as well as the time series samples sent from other nodes (which we assume is received in DWT form, and uncompressed locally via Inverse Discrete Wavelet Transform (IDWT)). For this reason, the time series segments are not very long (or the sampling rate to acquire them is not very high). Operationally, we would expect that the nodes take, possibly periodically, some time off their regular operation to simultaneously sample the background noise of the channel in order to collectively determine the state/class of the channel. In Section IV we explain how the comparison of the time series is performed based on the cross-correlation across different nodes' time series.

Also, despite some similarities in the current work with that of cognitive networks, e.g., the identification of which channel

DOI 10.1109/LCN.2016.127

is used to [4], our approach is more general, specifically: (i) we do not ascribe a role to a channel user (e.g., primary or not), and in fact, (ii), there is no requirement that the channel is used by a communicating user as the interference can emanate from non-communicating sources as well (e.g., microwave oven, elevator, internal combustion engine, etc.), and, consequently, (iii), the interference does not obey structure, such as modulation, that would allow it to be detected as such (e.g., via cyclostationarity [5]), and even if it was the result of modulation, (iv), the sampling rate is too low to be able to identify modulation characteristics with certainty.

II. Related Work

The role of wavelets, to represent time series of network activity has been explored in the past in the context of multifractal wavelet models for broadband traffic analysis [6], i.e., to describe actual packet data traffic in a wired environment rather than interference patterns in a wireless communication setting. Also, the general idea is to have large collections (long runs) of samples, whose ultimate objective is to elicit a single characterization for the entire network traffic process (rather than what we want, which is the compression of the time series). To the best of our knowledge, there has been no work on the short-term wireless interference sample compression using wavelets. In DWT, a wavelet prototype function is chosen, called a basis wavelet or mother wavelet. Temporal analysis is performed with a contracted, high-frequency version of the prototype wavelet; while frequency analysis is performed with a dilated, low-frequency version of the same wavelet. Hence, the original signal or function can be represented in terms of the wavelet expansion using coefficients in a linear combination of the wavelet functions. If the optimal mother wavelet is chosen and adapted to the examined data and if/when the coefficients are truncated below a threshold, the data is sparsely represented which makes wavelets suitable for data compression [7]. Towards this end, Ngui et al. [8] enlisted mother wavelet selection methods based on the similarity between the analyzed signal and the candidate mother wavelet, in contrast to selecting a mother wavelet on its properties alone. We follow in this paper the same principle of selecting an application–specific mother wavelet. In previous work, for example, Singh et al. [9] proposed a quantitative approach based on the maximum cross-correlation coefficient criterion, to select the optimal wavelet basis function in order to denoise ECG signals. Patrick P.C. Tsui et al. [10] proposed a technique for automatic ultrasound non-destructive Foreign Body (FB) detection and classification based on an information measure (via Shannon and relative entropy). We paid attention to the compression effect of the DWT on our data, combining in essence steps such as those followed by Singh's et al. [9], along with a form of a compression ratio inspired by the work of Chompursi et al. [11].

III. Mother Wavelet and Thresholding Selection

Using visual inspection of the decompressed time series we find significant differences compared to the original time series. Those differences depended less crucially on the mother wavelet selection and more on the used thresholding technique. Thresholding allows coefficients of magnitude close to zero (corresponding to minor signal details) to be set to zero and be omitted, still allowing the signal to be recovered adequately. Several thresholding techniques have been studied and are now standard in numerical computing environments. In our work, seven thresholding methods were implemented in MATLAB, using both 'hard' and 'soft' thresholding [12]. When an estimation of level noise is needed for these implementations, it was approximated based on first-level wavelet coefficients, since our real world RSSI data do not follow an ideal Additive White Gaussian Noise (AWGN) model.

A. The RSSI Traces

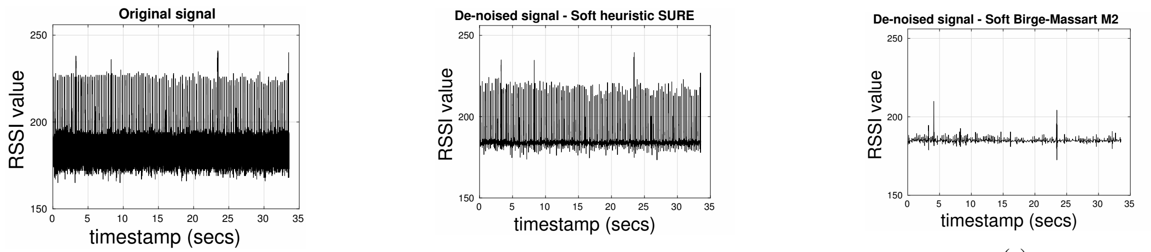
We use the RSSI traces collected by Boers et al. [2], across 256 channels in an indoor urban environment. The nodes were placed in a four-by-four grid with 1.84 m spacing. The sensors were just recording the RSSI values without transmitting. All 16 nodes would switch in unison to a new channel, and simultaneously sample the RSSI. In total, 256 channels were examined producing a total of 256×16=4096 traces. The frequencies were from a base frequency of 904 MHz to a maximum of 954 MHz. Each trace was a sequence of 175000 successive RSSI samples, representing 35 seconds (of the node's local clock) at 5000 samples per second.

B. Wavelet Compression Evaluation

We considered seven thresholding methods (both soft and hard versions) and Daubechies, Symmlet, Coiflet filters of varying orders. Initially, we kept the wavelet filter constant and implemented all the thresholding methods. We visually (qualitatively) inspected the recovered signals to verify which patterns of each class remain untampered, as the quantitative facets we used[1] were unfortunately insufficient. Visually speaking, we are interested in the thresholding method that guarantees consistent preservation of spikes for qs and qrs classes, as well as a uniform behavior of the noise present in the signals (i.e. noise preserved or suppressed throughout the duration of each signal and on the same scale across different classes). Also, the more levels that were used in the transform, the more intense the denoising and compression effects. Moreover, the possibility of losing a pattern increases in higher levels and depends on the principles applied by the thresholding method. Note that the qs exhibits a similar behaviour to qrs, while the q class is similar to the hl.

An example of acceptable vs. unacceptable results of thresholding in a qrs class series is presented in Figures [1a-1c]. It shows that the recovered signal after soft heuristic SURE (Stein's Unbiased Risk Estimate [13]) thresholding and denoising over 10 levels using a db2 wavelet filter limits the amplitude of the noise in the base of the signal, while preserving the periodic spikes. It should be noted that the soft heuristic SURE method is preferred over the hard heuristic

[1]We considered three metrics: (1) Root Mean Square Error (RMSE), (2) an idealized form of compression ratio, and (3) an energy–based ratio.



(a) Channel 54, node 6.

(b)

(c)

Figure 1: Channel 54, Node 6, 10 DWT Levels

SURE method, because it shrinks the noise amplitude at the base of the signal better than hard. A poor choice of a thresholding method is evident in Figure 1c, where the soft Birge-Massart thresholding will completely suppress the periodic spikes (leaving behind only two in this example). For a sm class series good results are obtained from soft heuristic SURE method and they also turn out to be more forgiving to the use of the hard heuristic SURE and even the soft version of the Birge-Massart thresholding. We have also studied the case of hl class time series, and a distinct observation is that the hard Birge-Massart thresholding results in undesired effects, as it does not uniformly suppress the noise, leaving spikes of noise randomly interfering with the dominant "clean" pattern.

In short, based on the results we collected, we can claim that the soft heuristic SURE thresholding method has the acceptable and desired effects of denoising and compression for all classes. After selecting the best thresholding method, we kept it constant while various types of wavelet basis functions were evaluated with the three metrics. However, all base functions produced similar results in terms of absolute values and did not suppress the characteristics of any class.

IV. Time Series Comparison

Pairwise Analysis of Time Series

With the wavelet basis function and thresholding decided (and assumed known to all nodes), the next task is to find the similarity across the time series collected from the various nodes, once decompressed via IDWT, via pairwise cross–correlation computation. Time series of different nodes, that we know are very similar, should exhibit a very high cross-correlation. In an ideal globally–synchronized distributed clock experiment, we would only care about the presence of strong cross–correlation at lag zero. Given our observations about the node clock drift and the possible impact of buffering and processing at the nodes and the data collection host, we conjectured that, as long as the cross–correlation is maximized at a lag within a small range around lag zero, it is very likely that the nodes indeed observe the same channel behavior at the same point in natural time, and it is only the reporting of their data that is skewed with respect to timestamp values. We rather arbitrarily set the "acceptable" lag range to correspond to timestamp discrepancies of $+/-10msec$, so that these results are comparable to the ones from our previous work [3]. In this section we compare DWT-based results for 2, 5, and 10 levels of compression, against the cross–correlation results reported in [3]. We investigate how the different levels affect this aggregate analysis.

In this work, from the 23438 (based on the ground truth found [2]) that agree on the class, we found that 11169 and 10741 pairs exhibit maximum cross–correlation at lags corresponding to timestamp discrepancies of $+/-10msec$ that indicate synchronization *and* correct classification of the signals, at 2 and 5 levels respectively. For 10 levels we have 9927 such pairs. As a result, we observe that node pairs are more likely to be considered to belong to the same class in comparison to the ones in the previous work that were 9696 [3], according to the cross–correlation statistic. This is very encouraging, since the cross–correlation is expected to be more sensitive to the detection of dissimilarities in the denoised signals, as the distinctive characteristics of the classes now become the distinctive characteristic of the time series due to the level of denoising.

Aggregate Analysis of Node Pairs

Table I summarizes the results, presenting the percentage of pairs of time series, that were known to belong to the same class, whose maximum cross-correlation falls within the corresponding cross-correlation "bin" (bins are 0.2 units wide). Clearly, the larger the percentage in the bins with the higher values, the better. The table has one column dedicated to each class. For every class we compare these maximum cross–correlation results (per bin) for four cases: Results from our previous work labeled (Orig.) [3], and results for soft heuristic SURE thresholding with db2 filter in 2, 5 and 10 levels.

A trend we observe is that as the levels of wavelet decomposition increase the used maximum cross-correlation, being a metric of similarity, increases consistently, shifting 'binwise' to higher bins. This is more evident for the q and qs classes, followed by qrs and sm. It is also noteworthy that for the lowest wavelet decomposition level, i.e., 2, we have results very similar to those found in [3]. This is expected, since the lower the decomposition level, the coarser the denoising and compression, resulting in signals resembling a signal under low pass filtering. As a result, for both cases, the qrs and sm channels class characterizations can be trusted as depicting accurately the same channel state. The hl classification is

Table I: Percentages of the maximum cross-correlation, $r_{xy}(k)$, between time series known to be of the same class, occurring in lags $[-10, 10]\,ms$ with a maximum falling within specific bins. "Orig." are the results reported in [3], while 2, 5, and 10 are the DWT levels under soft heuristic SURE thresholding, and db2 filter.

max $r_{xy}(k)$	q				qs				qrs				sm				hl			
	Orig.	2	5	10	Orig.	2	5	10	Orig.	2	5	10	Orig.	2	5	10	Orig.	2	5	10
$[0, 0.2)$	48.0	71.8	23.0	1.0	50.2	60.5	9.9	2.9	1.7	3.4	0.6	0.6	0.5	1.2	0.3	0.3	42.9	57.1	50.0	16.7
$[0.2, 0.4)$	27.3	19.6	37.0	18.4	39.0	34.5	39.8	23.8	26.8	28.8	8.4	8.5	2.7	5.6	4.7	4.1	14.3	0.0	12.5	33.3
$[0.4, 0.6)$	16.1	6.6	21.6	33.5	9.8	4.7	32.1	31.4	47.9	45.8	27.4	24.4	8.6	9.5	6.9	6.6	28.6	28.6	12.5	25.0
$[0.6, 0.8)$	7.7	2.0	13.1	27.6	0.9	0.3	17.4	30.1	22.9	21.4	48.1	47.4	13.5	15.3	11.9	11.6	14.2	14.3	25.0	25.0
$[0.8, 1)$	0.9	0.0	5.3	19.5	0.1	0.0	0.8	11.8	0.7	0.6	15.6	19.1	74.7	68.3	76.1	77.4	0.0	0.0	0.0	0.0

debatable as a non-trivial percentage corresponds to low maximum cross–correlation and the q and qs classifications are the most problematic because of the very low cross-correlation. The most dramatic 'binwise' shifts are observed for the q and the qs classifications for the highest wavelet decomposition level, however, the previous does not apply to the hl case which now becomes the problematic case.

Finally, since the corresponding statistics about node pairs that are known to not agree on the channel classification are also available, we were able to determine how well relying on a high maximum cross–correlation can avoid false positives. We observed that as the levels increase from 2 to 5 to 10, the false positives (i.e. identifying a pair of time series as being in the same class while they are not) increase from 8.4% to 24.6% to 48.4%. A possible justification is the really high suppression of random noise behavior whose contribution to the numerical value of the cross–correlation would cancel out, compared to the filtered and smoothed time series. Overall, level 5 may be seen as a good compromise.

However, despite the increased level of denoising and compression the hl time series result in maximum cross-correlation in the lowest bins, i.e., there is weak evidence of their similarity if we resort to cross-correlation. Hence, nodes that see very similar hl time series would have trouble reaching agreement that they are observing the same time series. Finally, as as an indication of compression effectiveness, the DWT representation results in the following average number of non-zero coefficients and (in parentheses) average maximum integer value per coefficient, for 2, 5, and 10 levels respectively: 43009 (422), 7296 (1142), and 3088 (6115). Clearly the reduction of non-zero coefficients as we increase levels, has to be seen against the trade off of increasing bits per coefficient for increasing levels.

V. Conclusion

We have studied whether DWT is a suitable method to compress and denoise RSSI time series in a wireless network. We are aiming to communicate the state of a channel from the perspective of a single node among the WSN nodes, in a new compressed and denoised version. We have examined the suitability of different wavelet filters and thresholding methods aiming to compress while preserving the noise patterns and find an appropriate compression level. We have found that for higher decomposition levels the accuracy of the similarity measure improves. However, there are more false positives. Therefore, there is a compromise to be made in choosing an appropriate level, taking into consideration that the number of non-zero coefficients decreases with increased level.

References

[1] H. Lee, A. Cerpa, and P. Levis, "Improving wireless simulation through noise modeling," in *Proceedings of the 6th international conference on Information processing in sensor networks*, ser. IPSN '07. New York, NY, USA: ACM, 2007, pp. 21–30.

[2] N. Boers, I. Nikolaidis, and P. Gburzynski, "Patterns in the rssi traces from an indoor urban environment," in *Computer Aided Modeling, Analysis and Design of Communication Links and Networks (CAMAD), 2010 15th IEEE International Workshop on*, 2010, pp. 61–65.

[3] A. Vlachaki, I. Nikolaidis, and J. J. Harms, "A study of channel classification agreement in urban wireless sensor network environments," in *SENSORNETS 2014 - Proceedings of the 3rd International Conference on Sensor Networks, Lisbon, Portugal, 7 - 9 January, 2014*, O. Postolache, M. van Sinderen, F. H. Ali, and C. Benavente-Peces, Eds. SciTePress, 2014, pp. 249–259.

[4] I. F. Akyildiz, B. F. Lo, and R. Balakrishnan, "Cooperative spectrum sensing in cognitive radio networks: A survey," *Phys. Commun.*, vol. 4, no. 1, pp. 40–62, Mar. 2011.

[5] G. Gelli, L. Izzo, and L. Paura, "Cyclostationarity-based signal detection and source location in non-gaussian noise," *IEEE Transactions on Communications*, vol. 44, no. 3, pp. 368–376, Mar 1996.

[6] R. H. Riedi, M. S. Crouse, V. J. Ribeiro, and R. G. Baraniuk, "A multifractal wavelet model with application to network traffic," *IEEE Trans. Inf. Theor.*, vol. 45, no. 3, pp. 992–1018, Sep. 2006.

[7] A. Graps, "An Introduction to Wavelets," *IEEE Comput. Sci. Eng.*, vol. 2, no. 2, pp. 50–61, Jun. 1995.

[8] W. K. Ngui, M. S. Leong, L. M. Hee, and A. M. Abdelrhman, "Wavelet Analysis: Mother Wavelet Selection Methods," *Applied Mechanics and Materials*, vol. 393, pp. 953–958, 2013.

[9] B. N. Singh and A. K. Tiwari, "Optimal Selection of Wavelet Basis Function Applied to ECG Signal Denoising," *Digit. Signal Process.*, vol. 16, no. 3, pp. 275–287, May 2006.

[10] P. P. C. Tsui and O. A. Basir, "Wavelet basis selection and feature extraction for shift invariant ultrasound foreign body classification," *Ultrasonics*, vol. 45, no. 14, pp. 1 – 14, 2006.

[11] Y. Chompusri, K. Dejhan, and S. Yimman, "Mother wavelet selecting method for selective mapping technique ECG compression," in *Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2012 9th International Conference on*, May 2012, pp. 1–4.

[12] D. L. Donoho, M. Vetterli, R. A. DeVore, and I. Daubechies, "Data compression and harmonic analysis," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2435–2476, Oct. 1998.

[13] C. M. Stein, "Estimation of the Mean of a Multivariate Normal Distribution," *The Annals of Statistics*, vol. 9, no. 6, pp. 1135–1151, 1981.