Post-Processing in Wireless Sensor Networks: Benchmarking Sensor Trace Files

Martin C. Bor
Delft University of Technology
Netherlands
m.c.bor@tudelft.nl

Ioannis Chatzigiannakis
Research Academic Computer
Technology Institute, Greece
ichatz@cti.gr

Stefan O. Dulman
Delft University of Technology
Netherlands
s.o.dulman@tudelft.nl

Panagiotis Kikiras
Dept. of Computer & Comm.
Engineering, University of
Thessaly, Greece
kikirasp@inf.uth.gr

Evangelos Theodoridis
Research Academic Computer
Technology Institute, Greece
theodori@cti.gr

Otto W. Visser
Delft University of Technology
Netherlands
o.w.visser@tudelft.nl

ABSTRACT

Wireless sensor network research usually focuses on the reliable and efficient collection of data. Here, we address the next step in the traces lifetime: we aim at investigating and evaluating, by qualitative and quantitative means, data repositories of already collected measurements. We propose the use of a set of new metrics, which enable reliable evaluation of algorithms using traces (both in average cases and "stressful" setups) removing the need for running algorithms in a real testbed, at least in the development stage.

Categories and Subject Descriptors

C.2.4 [**Distributed Systems**]: Distributed applications; D.2.8 [**Software Engineering**]: Metrics—*complexity measures, performance measures*

General Terms

Experimentation, Performance

Keywords

Benchmark,Sensor Trace Files, Wireless Sensor Networks

1. INTRODUCTION

Aside from consuming the data at the moment it is produced (i.e., as in security applications targeting real-time event detection), there are a number of cases in which traces present interest also at a later time (e.g., optimizations of existing algorithms). To the best of our knowledge there is no metric currently in use, in the wireless sensor network field, characterizing the degree of "difficulty" present in the existing traces. We propose, as the main contribution of this paper, a new set of metrics describing the network status and the temporal and spatial correlation, which can be used to characterize the traces. A classification of the traces becomes thus feasible and this enables, in turn, straightforward benchmarking of different algorithms based on commonly accepted traces. Most of WSN algorithms usually have different behavior in different WSN settings with different metric characteristics.

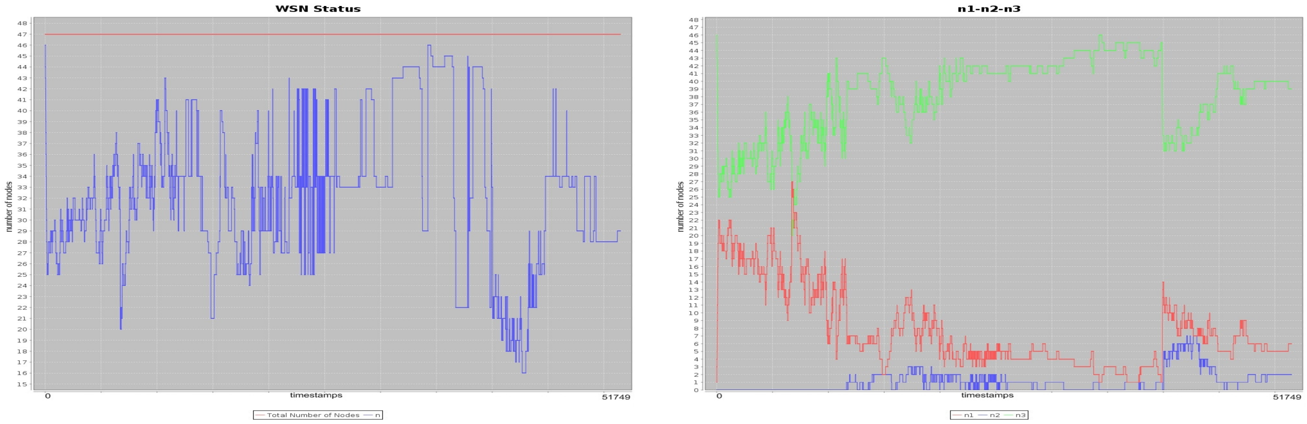
2. CLASSIFICATION OF SENSOR TRACES

The efficiency of most of the algorithms depends mostly on the correlation among data measured by the sensor nodes and on the size and the frequency of the topology changes. This correlation of measurements can be *spatial*, *temporal* and *semantic* [6]. In this section we propose a mapping of all these parameters of each trace in a feature space, an abstract space where each sensor trace is represented as a point in n-dimensional space.

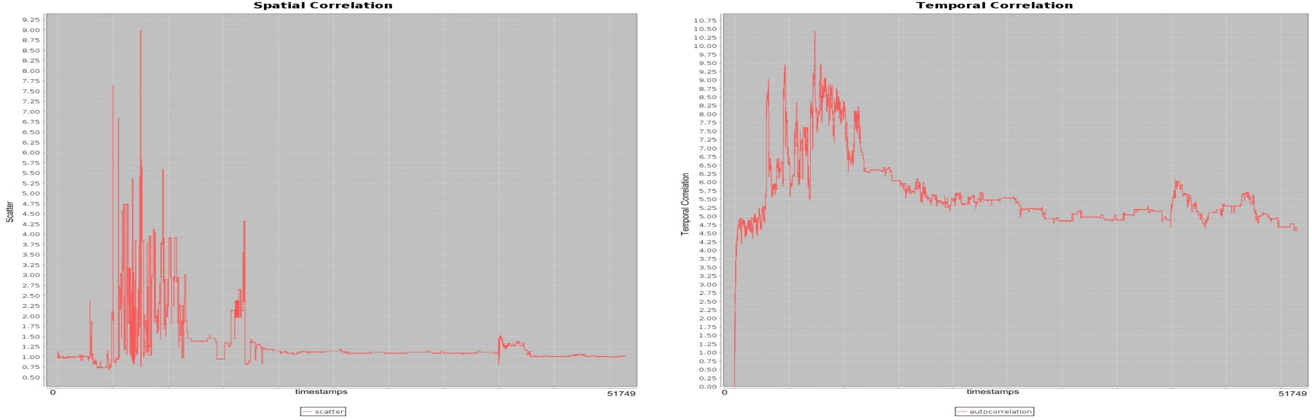
So we assume that each trace starts from time-stamp t_0 to t_T, where the first one is the first recorded time while the latter the last one. N expresses the number of installed nodes. For each of of these time instants, we calculate gradually a set of measures/properties trying to depict the status of the WSN at that moment.

We would like to capture for time instant t_i the status of the WSN. In the ideal case, we would like beside the number of total sensor nodes (N), active sensor nodes (n) and various other network invariants like: the number of active links (e), the maximum distance from the basestation (d), the average distance from the basestation (d_{avg}), the average node degree ($degree_{avg}$) etc. If the network topology is constantly recorded all these invariants can be calculated directly. In the opposite case, we have to infer some of them:
Number of active nodes: The number of active nodes n_i in t_i. By using a parameter w_1 as a threshold, we assume that a node is active if it has a recorded transmission in $[t_{i-w_1}, t_i]$ time period. The w_1 may the a number of timestamps or a time period. This threshold can also be static, for example $T_{w_1} = 100$ timestamps or $T_{w_1} = 1$ hour, or can change according to the average message record period T_p that is observed, for example $T_{w_1} = 10 \cdot T_p$.
WSN cohesion: With observation until timestamp t_i, if we cluster the nodes according to the their time of last recorded transmission, their average transmission period P in $[t_0,$

PE-WASUN'10, October 17–18, 2010, Bodrum, Turkey.



(a) Active Nodes (b) WSN Cohesion



(c) Spatial Correlation (d) Temporal Correlation

Figure 1: Trace #1 (LOFAR-agro deployment) for a subset of 51749 measurements

$t_i]$ and their average transmission period P_{w_1} in window $[t_{i-w_1}, t_i]$, we can make some observations:
→ If the application periodically measures values and sends them to the basestation, then each cluster is going to have similar periods P and P_{w_1}. In this case, if we observe a few clusters and their total number of nodes n is very close to N, we can assume that all the WSN is coherent. In the opposite way, that we find clusters with very large P and P_{w_1} compared to other clusters or that the total n is quite smaller than N, we can assume than there are some parts of the WSN that can not communicate with the base station.
→ In the opposite way when P_{w_1} is quite smaller than P the application may work asynchronously, meaning that when it sends a message only when it detects an event. If we observe that few clusters and their total n is very close to N, we might say that most of the nodes detect the event while if there is variance among the clusters a local event has occurred.

In order to express the status of the WSN at time-stamp t_i we use the number of active nodes n as defined below:

$$n = n_1 + n_2 + \begin{cases} 0, \text{if} \quad n_1 \leq \frac{1}{2} n_2 \\ \frac{2}{3} n_3, \text{if} \quad n_1 \leq n_2 \\ \frac{n_1}{n_1+n_2} n_3, \text{if} \quad n_1 \geq n_2 \end{cases}$$

where n_1 is the number of nodes with $P_{w_1} < P$, n_2 is the number of nodes with $P_{w_1} \simeq P$ and n_3 the number of silent nodes at t_i $(n_1 + n_2 + n_3 = N)$.
Spatial Correlation. In time instant t_i, we would like to find if the sensed values of a node have a correlation with adjacent nodes. If the network topology is stored explicitly then something like that is simple to handle. In the other case, we have to guess the neighbors of a node. For the time window $[t_{i-w_1}, t_i]$, we cluster all active nodes according to their sensed data assuming that neighbouring nodes will measure the same values. We may cluster the nodes not only by their last measured values but with the sequences of their sensed data in this time window [4]. So we construct a few clusters (e.g., 3-5) by using K-means, to capture potential variations in the measurements. Then we use as as a metric index the average scattering of the clusters defined as

$$Scat(c) = \frac{1}{c} \sum_{i=1}^{c} \frac{\|\sigma(C_i)\|}{\|\sigma(X)\|}$$

where $\sigma(C_i)$ is the variance of the cluster C_i and $\sigma(X)$ the variance of the trace which is a vector with d dimensions, defined as

$$\sigma(X)[p] = \frac{1}{n} \sum_{i=1}^{n} ((x_i[p] - \overline{x}[p])^2$$

$$\overline{x}[p] = \frac{1}{n} \sum_{i=1}^{n} x_i[p],$$

where $p \leq d$ is the dimension of measurement vectors x_i that form trace X. With the similar formulas we calculate the variance of cluster C_i. In the case where in the WSN we have a few events in some subareas, some of the clusters are going to be concentrated and others more scattered. The variance of the scatter of the clusters can express such non uniformity.

Temporal Correlation. Concerning temporal correlation, for each node we can observe in window $[t_{i-w_1}, t_i]$ type the autocorrelation of the time lag for each measurement.

$$cor_l = \frac{1}{w_1 - l} \sum_{i=1+l}^{w_1} (m_i \cdot m_{i-l} - \overline{y_i} \cdot \overline{y_{i-l}})/VAR(m)$$
$$\overline{y_i} = \frac{1}{w_1 - l} \sum_{i=1+l}^{w_1} m_i, and\ \ \overline{y_{i-l}} = \frac{1}{N-l} \sum_{i=1}^{N-l} m_i$$

where m are the measurements, and l the time lag. We find for each measurement the maximum time lag l. For all alive nodes in the window we take the variance of all these correlations for each measurement type.

3. EXPERIMENTAL EVALUATION

We here evaluate the proposed four metrics. We wish to assess their suitability for the classification of traces. For performance evaluation we used five traces that were recorded during various field environmental experiments. Due to space reasons we present the results of two of them.

Benchmarking Algorithms: For this paper, we chose three data aggregation algorithms to evaluate the proposed metrics. The first algorithm is *Probabilistic Counting with Stochastic Averaging* (PCSA) [3], which uses only *logarithmic* memory space. The second algorithm is the *LogLog Counting algorithm*, presented and analyzed in [2]. The main contribution of the LogLog Counting algorithm is that it reduces the size of the accumulation synopsis from $\log n$ to $\log \log n$. However, the standard error is increased from $0.78/\sqrt{k}$ to $1.30/\sqrt{k}$, where k is the number of bitmaps. This means that the LogLog Counting trades accuracy for improvement in space complexity (which is improved logarithmically). A different, more complex approach to network data summarization was presented in [1] introducing a new sketch that maintains duplicate insensitivity, asynchronous arrivals and time decay of the processed data simultaneously.

Trace 1: LOFAR-agro experiment. This trace was collected by researchers from Delft University of Technology, the Netherlands, in an outdoor experiment targeted at precision agriculture (more details and results are presented in [5]). A static network was deployed in a potato field and monitored the environmental characteristics of the field over a period of a several months. All the nodes reported measured parameters (temperature and relative humidity) back to the gateway every ten minutes. Additionally, information such as statistics on neighbors, communication links quality, battery power, routing topology, etc. was embedded in the trace. The outside deployment makes this trace extremely interesting due to the slow changes in topology (and overall constant degradation of performance) influenced by the day/night cycles and environmental characteristics. The trace includes information on failures: comm. failures (both missing data and duplicates are represented), undesired reboots of the hardware platforms and time sync. failures.

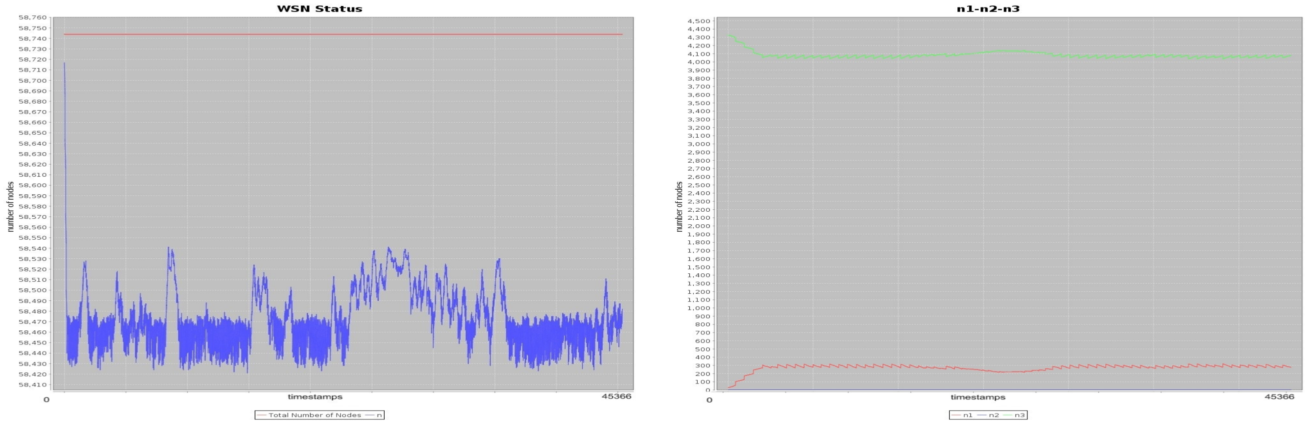
Figure 1 depicts the four metrics for a subset of 51749 timestamps of the trace (the total is over 1 million). Figure 1(a) shows the estimated number of active nodes (WSN status). Figure 1(b) shows WSN cohesion: n_1 is the red curve (number of active nodes with $P_{w_1} < P$), n_2 is the blue curve (number of active nodes with $P_{w_1} \simeq P$) and n_3 is the green curve (number of silent nodes). Recall that $n_1 + n_2 + n_3 = N$. Both metrics seem to capture the sleep/awake cycles of the nodes and the variations in the number of active nodes in the network.

Figure 1(c) shows the spatial correlation of the sensed data; that is the correlation of the data sensed by neighborhoods of nodes. Figure 1(d) shows the temporal correlation of the sensed data; that is the correlation of the data sensed during a particular time frame. Both metrics initially show high variance (i.e., data is not correlated). This is due to various types of failures present in the particular subset of the trace. In the remainder, both metrics show a low variance (i.e., data are correlated). This is due to the nature of the experiment; neighboring nodes measure similar metrics and nodes measure similar values during the same time of the day. It is clear that the "interesting" parts of the trace are clearly pointed out by the proposed metrics.

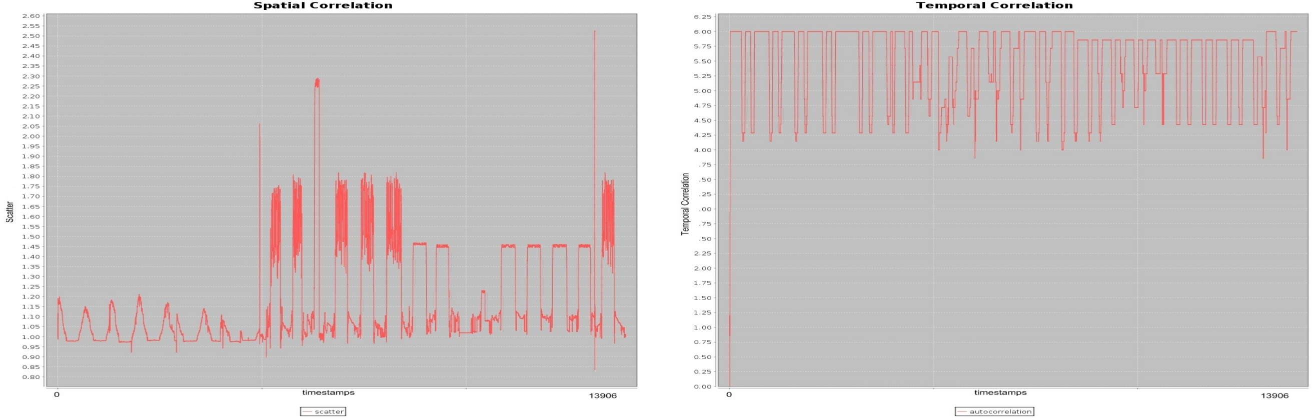
Trace 2: Husbandry deployment. The animal house unit of B.S.R.C. "Alexander Fleming" provided a husbandry of experimental mice (it is used by the biomedical research community). The main goal of the monitoring system was to provide 24 hour monitoring of the mice mothers and newborn mice, and to investigate the impact of various factors such as temperature, humidity, ammonia concentration and man-made noise to newborn mice development. The data collection system was based on a combination of static nodes (to be deployed in particular places in the house) and mobile nodes of very small size (to be attached to mice). The data collection lasted for 2 months and 117454 measurements were collected. Figure 2 depicts the four metrics for the complete trace. In this trace, due to the high mobility of the nodes and the delay-tolerant nature of the experiment, we cannot validate that the estimation of the number of nodes, as shown in Figure 2(a) and Figure 2(b) is correct. However, both figures show very clearly the high dynamics of the network. Also due to the delay-tolerant style of reporting the sensed data, the WSN cohesion metric indicates that the majority of the nodes are in "inactive" state. Also due to the mobility of the nodes, the scatter of the sensed data in terms of spatial correlation is very high (see Figure 2(c)). This is in line with the spirit of the experiment: e.g., different areas of the husbandry have different levels of ammonia. Also as the humans maintain the house, the sensed data are heavily affected. This is reflected by the high scatter in the temporal correlation of the trace (see Figure 2(d)). We believe that based on the proposed metrics we can characterize this trace as "hard". To further emphasize the hardness of this particular trace, we have executed the PCSA, LogLog and TDS algorithms on the sensed data. We used a window size of 1000 measurements. We then compared the aggregated values (e.g., average temperature) to the actual one. Figure 3 depicts the error in the estimate for each algorithm. The PCSA algorithm error is about 85% of the actual value (see Figure 2), the LogLog algorithm error is about 130% (see Figure 3) while the TDS algorithm computes the aggregates with the lowest error of about 20% (see Figure 3). These error levels are clearly beyond the expected ones in simple WSN deployments.

4. CONCLUSIONS

In this paper we investigate and evaluate in practice, by qualitative and quantitative means, data repositories of measurements produced by wireless sensor networks, in order to



(a) Active Nodes (b) WSN Cohesion



(c) Spatial Correlation (d) Temporal Correlation

Figure 2: Trace #3 (Husbandry deployment) trace for the total 117454 measurements

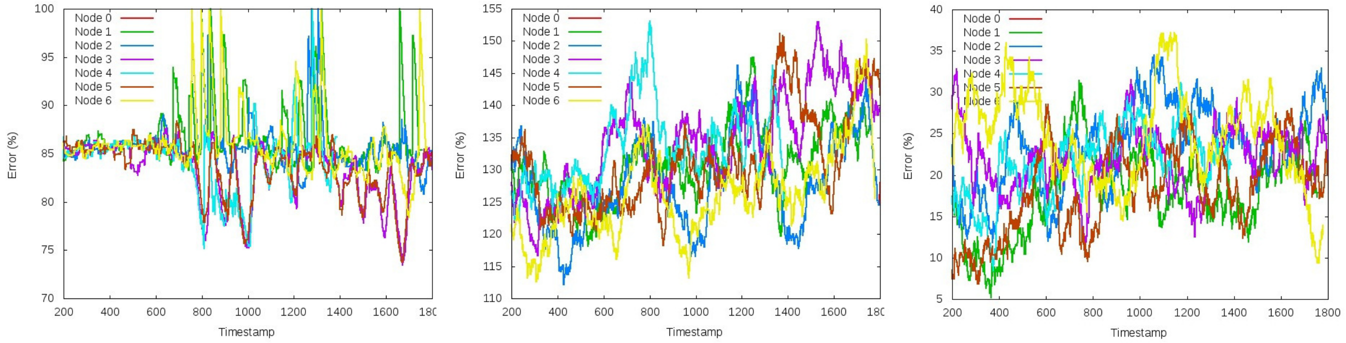


Figure 3: Performance of Data Aggregation under Trace #3 (Husbandry deployment). TDS, PCSA , LogLog.

exploit them as benchmark traces and benchmarked several data aggregation algorithms in wireless sensor networks.

5. ACKNOWLEDGMENTS

This work has been partially supported by the European Commission under contract numbers IST-2008-224460 (WISEBED) and FP7-2007-2-224053 (CONET).

6. REFERENCES

[1] G. Cormode, S. Tirthapura, and B. Xu. Time-decaying sketches for sensor data aggregation. In *26th ACM Symp. on Principles of distr. comp.*, pages 215–224, 2007.

[2] M. Durand and P. Flajolet. Loglog counting of large cardinalities (extended abstract). ESA03, volume 2832 of LNCS (2003), pp. 605-617.:605–617, 2003.

[3] P. Flajolet and G. N. Martin. Probabilistic counting algorithms for data base applications. *J. Comput. Syst. Sci.*, 31(2):182–209, 1985.

[4] D. Goldin, R. Mardales, and G. Nagy. In search of meaning for time series subsequence clustering: matching algorithms based on a new distance measure. In *15th ACM Intl Conf. on Inf. and KM (CIKM '06)*, pages 347–356. ACM, 2006.

[5] K. Langendoen, A. Baggio, and O. Visser. Murphy loves potatoes: Experiences from a pilot sensor network deployment in precision agriculture. In *14th Int. Work. on Par. and Distr. Real-Time Systems* ., Apr. 2006.

[6] S. Nath, P. B. Gibbons, S. Seshan, and Z. Anderson. Synopsis diffusion for robust aggregation in sensor networks. *ACM Trans. Sen. Netw.*, 4(2):1–40, 2008.