Improving Accuracy of the Shewhart-based Data-Reduction in IoT Nodes using Piggybacking

Anish Shastri  Vivek Jain  Dr. Sachin Chaudhari  Dr. Shailesh Singh Chouhan†  Dr. Stefan Werner*

Signal Processing and Communication Research Center
International Institute of Information Technology-Hyderabad
†EIS Lab, Luleå University of Technology, Sweden
*Norwegian University of Science and Technology (NTNU), Trondheim, Norway

18 April 2019
Internet of Things offer an array of Sensor Network applications that involve constant measurement of real world phenomena.

The sensors have a limited amount of power supply, with more than 50% of energy spent on communication. Both reception and transmission cost energy.

This has given rise to many algorithms to reduce data communication without loss of events.

Least Mean Square based data reduction algorithm is a widely used algorithm to solve this problem.

However, a much simpler algorithm based on Shewhart test outperforms LMS based algorithm, however, MSE is comparable and is a major issue.

The aim is to solve the MSE issue by the proposed Piggybacking and Interpolation based algorithm, implement it on the customized sensor node and quantify the results.
Motivation

- Data reduction algorithms have been simulated on datasets, but very few hardware implementations, especially on commonly available hardware platforms.

- MSE has been a major drawback in most of the proposed algorithms, so an attempt is made to solve this problem.

- Validation of theoretical results with practical implementation results.

- Implementation of the algorithm on low-power sensor nodes would further increase their battery-lifetime when deployed in various application scenarios.
Contributions

- Use of Shewhart test for data-reduction in IoT networks and performance comparison with the widely-used LMS based algorithm.

- Use of Piggybacking and Interpolation based schemes along with Shewhart algorithm. Experimental results have also been presented.

- Current consumption measurements have been carried out on a custom hardware implemented in our previous work.

- The performance of the proposed algorithm is demonstrated by performance parameters such as savings in the number of transmissions, decrease in the MSE, power consumption and extension in the battery lifetime.
LMS based Prediction Algorithm

Algorithm

\[ \hat{x}[n] = w_{n-1}^T x_{n-1} \]
where, \( w_{n-1} = [w[1] \ldots w[N]]^T \) and \( x_{n-1} = [x[n-1] \ldots x[n-N]]^T \)

\[ e[n] = x[n] - \hat{x}[n] \]

If \( e[n] \) is within error threshold,
\[ x_n = [\hat{x}[n] \times [n-1] \ldots x[n-(N-1)]]^T \]
else,
\[ x_n = [x[n] \times [n-1] \ldots x[n-(N-1)]]^T \]

\[ w_n = w_{n-1} + \mu \cdot e[n] \cdot x_n \]

\[ 0 \leq \mu \leq \frac{1}{E_x} \]
where,
\[ E_x = \frac{1}{K} \sum_{n=1}^{K} |x[n]|^2 \]

Parameters

- \( \mu \) - Step size of the prediction filter.
- \( N \) - Filter tap length.
- \( E_x \) - Mean power of the signal.
- \( K \) - Number of temperature samples.
Shewhart based Data Reduction Algorithm

**Algorithm**

- $x[n]$ is the sensed temperature data.
- $\hat{x}[n]$ is the predicted data at the sink node.
- Prediction at the sink node is done according to the sensed data as follows:
  \[
  \hat{x}[n] = \begin{cases} 
  x[n], & \text{if data received.} \\
  \hat{x}[n-1], & \text{otherwise.}
  \end{cases}
  \]
- $x[n]$ is transmitted to the sink only if $|e[n]| = |x[n] - \hat{x}[n]| > \epsilon_t$

**Observations**

- Significantly lower computational complexity compared to the LMS based algorithm.
- Ensures same performance as LMS in terms of bounded error.
- Unlike LMS, it does not have any tunable parameter like step size which will affect its convergence.
Proposed Piggybacking-based Shewhart test

Algorithm

- Window of sensor values 
  \( \left( x[n_{t-1}], x[n_{t-1} + 1], \ldots, x[n_t - 1], x[n_t] \right) \) of length \( W_t = n_t - n_{t-1} + 1 \). The two end-points \( x[n_{t-1}] \) and \( x[n_t] \) are transmitted from the node to the sink.

- \( p \) approximately equispaced values 
  \( x \left[ n_{t-1} + \left\lfloor \frac{m(W_t-1)}{p+1} \right\rfloor \right] \) for \( m = 1, \ldots, p \) are piggybacked with the transmission \( x[n_t] \).

- Let \( x_t = \left[ x[n_{t-1}], x \left[ n_{t-1} + \left\lfloor \frac{W_t-1}{p+1} \right\rfloor \right], \ldots, x \left[ n_{t-1} + \left\lfloor \frac{p(W_t-1)}{p+1} \right\rfloor \right], x[n_t] \right]^T \), then the predicted values at the sink node for the sample-indices \( n_{t-1} < n < n_t \) are given as

  \[ \hat{x}[n] = \begin{cases} x[n], & \text{if data received.} \\ f(x_t), & \text{otherwise.} \end{cases} \]

where \( f(x_t) \) is an interpolation function.

Observations

- \( \left\lfloor . \right\rfloor \) denotes the round-off operation.
- \( W_t \) changes with the value of \( t \) as it depends on the time-series data and the threshold.
- If \( 2 < W_t < p \), all intermediate values are sent as piggybacks.
- Only simple and exact interpolation schemes such as linear, cubic and previous (also called as naive) are considered.
- Note that piggybacking only increases the payload of each packet and not the number of packets transmitted.
Sensor-node Implementation

(b) Schematic of the sensor node

(c) Hardware design of the sensor node
Network Deployment

(d) Star topology

(e) Network deployment floor plan
Measured temperature data in the indoor and outdoor scenarios along with transmission instances while using LMS and Shewhart-based data reduction schemes.
### Table – Comparison of the Shewhart and LMS methods for data-transmission reduction for the two data-sets. A total of 11000 sample points are taken for comparison

<table>
<thead>
<tr>
<th>Scenario</th>
<th>No. of Transmissions</th>
<th>% savings</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Indoor Scenario</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With LMS</td>
<td>408</td>
<td>96.29</td>
<td>0.0906</td>
</tr>
<tr>
<td>With Shewhart</td>
<td>58</td>
<td>99.47</td>
<td>0.0489</td>
</tr>
<tr>
<td><strong>Outdoor Scenario</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With LMS</td>
<td>682</td>
<td>93.80</td>
<td>0.0645</td>
</tr>
<tr>
<td>With Shewhart</td>
<td>201</td>
<td>98.17</td>
<td>0.0442</td>
</tr>
</tbody>
</table>

### Parameters for LMS based Algorithm
- $\mu = 2.768 \times 10^{-6}$ (Indoor) and $\mu = 2.77 \times 10^{-6}$ (Outdoor)
- Filter tap length $N = 9$
- $e_t = 0.5 \, ^\circ$C.
- Sensing interval $T = 30s$

### Remarks
- 11000 sample points taken for comparison.
- Shewhart based method outperforms LMS in terms of both MSE and Number of transmissions.
- This result is important as Shewhart based algorithm has less complexity compared to the LMS based algorithm.
(g) Measured current profile for 1 transmission instant with a payload of 4 bytes \((p = 1)\) and payload of 10 bytes \((p = 4)\). The blue and red curves are not completely overlapped for easy visualization.
(h) MSE for the indoor temperature data at the sink as a function of number of piggybacks using various interpolation schemes for the proposed piggyback-based Shewhart test.

(i) Relative decrease in the MSE of the indoor temperature data at the sink using the proposed piggyback-based Shewhart test with respect to no-piggybacking based Shewhart test.
Current Consumption Analysis

Table – Summary of the operation of the sensor node in different operating states

<table>
<thead>
<tr>
<th>State</th>
<th>Operation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I_{P_1}$</td>
<td>Sense temperature and computation</td>
</tr>
<tr>
<td>$I_R$</td>
<td>Turn ON XBee if data needs to be sent to the sink</td>
</tr>
<tr>
<td>$I_{Tx}$</td>
<td>Data transmission instant</td>
</tr>
<tr>
<td>$I_{P_2}$</td>
<td>Increment for further iterations</td>
</tr>
<tr>
<td>$I_S$</td>
<td>Sensor node sleeps for fixed time duration</td>
</tr>
</tbody>
</table>

Table – Average current consumption for different states for payload (2 bytes for Shewhart case, 4 bytes for 1 Piggyback and 10 bytes for 4 Piggybacks).

<table>
<thead>
<tr>
<th>State</th>
<th>Avg. Current (in mA)</th>
<th>Duration ($p = 0$) (in ms)</th>
<th>Duration ($p = 1$) (in ms)</th>
<th>Duration ($p = 4$) (in ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I_{P_1}$</td>
<td>5</td>
<td>118</td>
<td>120</td>
<td>124</td>
</tr>
<tr>
<td>$I_R$</td>
<td>34.8</td>
<td>30</td>
<td>31</td>
<td>35</td>
</tr>
<tr>
<td>$I_{Tx}$</td>
<td>47</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>$I_{P_2}$</td>
<td>5</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>$I_S$</td>
<td>0.0061</td>
<td>29837</td>
<td>29834</td>
<td>29826</td>
</tr>
</tbody>
</table>

Instruments

- Agilent 34461A 6½ digit, Digital Multimeter
- Keysight BenchVue software

(j) Measured current consumption profile
Battery Lifetime Estimation

Calculation

\[ I_{\text{avg}} = \frac{(I_{P1} t_{P1} + I_{R} t_{R} + I_{Tx} t_{Tx} + I_{P2} t_{P2} + I_{S} t_{S}) L}{3600} \]

where \( L = \) Total number of transmissions in 1 hr.

\[ \gamma = \frac{Q}{24 I_{\text{avg}}} \]

where \( \gamma \) is the expected lifetime of the battery in days and \( Q \) is the capacity of the battery in mAh.

**Table** — Comparison of average current consumption and expected lifetime of a node in the two scenarios of Shewhart algorithm with and without piggybacking.

<table>
<thead>
<tr>
<th>Method</th>
<th>Avg. Current (mA)</th>
<th>Avg. Lifetime (days)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Indoor Scenario Dataset</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Without Shewhart</td>
<td>0.0735</td>
<td>1513</td>
</tr>
<tr>
<td>Shewhart with ( p = 0 )</td>
<td>0.02599</td>
<td>4488</td>
</tr>
<tr>
<td>Shewhart with ( p = 1 )</td>
<td>0.0263</td>
<td>4430</td>
</tr>
<tr>
<td>Shewhart with ( p = 4 )</td>
<td>0.027026</td>
<td>4316</td>
</tr>
<tr>
<td><strong>Outdoor Scenario Dataset</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Without Shewhart</td>
<td>0.0735</td>
<td>1513</td>
</tr>
<tr>
<td>Shewhart with ( p = 0 )</td>
<td>0.02657</td>
<td>4390</td>
</tr>
<tr>
<td>Shewhart with ( p = 1 )</td>
<td>0.0269</td>
<td>4332</td>
</tr>
<tr>
<td>Shewhart with ( p = 4 )</td>
<td>0.02767</td>
<td>4215</td>
</tr>
</tbody>
</table>
Implementation of Shewhart based data reduction algorithm results in reduction in the data-transmission by 99.47% in Indoor scenario and 98.17% in Outdoor scenario compared to 96.29% in Indoor scenario and 93.8% in Outdoor scenario when employing LMS based algorithm.

A significant decrease in MSE of around 38% and 75%, respectively for one piggyback and four piggyback values with simple linear interpolation scheme is observed.

Lifetime of the customized wireless sensor node increases by 2.97 times and 2.90 times in the Indoor and Outdoor scenarios respectively, when employing Shewhart algorithm without piggybacking.

The decrease in the lifetime of the sensor nodes when employing our proposed algorithm with \( p = 1, 4 \) are much smaller as 1.29% and 3.8%, respectively for indoor scenario, for a decrease in the MSE of around 38% and 75% respectively.
A. Shastri, V. Jain, R. Singh, S. Chaudhari, and S. Chouhan, “On the implementation of
LMS-based algorithm for increasing the lifetime of IoT networks,” in 12th IEEE Int. Conf.
Advanced Telecommun. Syst. and Networks, 2018

W. Shewhart, Economic control of quality of manufactured product. New York : American
Society for Quality Control, 1980.

J. Lunden and S. Werner, “Real-time smart metering with reduced communication and

S. Santini and K. Romer,"An adaptive strategy for quality-based data reduction in wireless
2006, pp. 29-36.

battery life of a wireless sensor node,” Journal of Industrial Information Integration, vol. 4,
pp. 8–14, 2016.

F. Wu, C. Tan, M. Sarvi, C. Rudiger, and M. Yuce, “Design and implementation of a
low-power wireless sensor network platform based on XBee,” in IEEE 85th Vehicular Technol.

F. Aderohunmu, G. Paci, D. Brunelli, J. Deng, L. Benini, and M. Purvis, “An
application-specific forecasting algorithm for ex- tending WSN lifetime,” in IEEE Int. Conf.
Distributed Computing in Sensor Systems (DCOSS), 2013, pp. 374–381.
Thank You!
Any Questions?