Polling strategies for predictions in wireless sensor networks

LE BORGNE Yann-Aël
ULB Machine Learning Group
Computer Science Department
Université Libre de Bruxelles
1050 Brussels – Belgium
yleborgn@ulb.ac.be

http://www.ulb.ac.be/di/mlg

Work supported by the COMP2SYS project, sponsored by the Human Resources and Mobility program of the European community (MEST-CT-2004-505079)

Continuous monitoring

Readings are collected from all sensors at regular time intervals

<table>
<thead>
<tr>
<th></th>
<th>$s_1$</th>
<th>$s_2$</th>
<th>$s_3$</th>
<th>$s_4$</th>
<th>$s_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$s_1(1)$</td>
<td>$s_2(1)$</td>
<td>$s_3(1)$</td>
<td>$s_4(1)$</td>
<td>$s_5(1)$</td>
</tr>
<tr>
<td>2</td>
<td>$s_1(2)$</td>
<td>$s_2(2)$</td>
<td>$s_3(2)$</td>
<td>$s_4(2)$</td>
<td>$s_5(2)$</td>
</tr>
<tr>
<td>3</td>
<td>$s_1(3)$</td>
<td>$s_2(3)$</td>
<td>$s_3(3)$</td>
<td>$s_4(3)$</td>
<td>$s_5(3)$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Energy concerns

- The monitoring task is often required to run for a long period of time
- Sensing nodes can switch between sleeping/active modes (duty cycle)
  - Energy consumption in sleeping mode orders of magnitude lower than in active mode
- Prediction models and mote scheduling to increase time spent in the sleeping mode

Prediction models

Observation database
$D_{1:T}$, collected over a period $T$

1 $s_1$ $s_2$ $s_3$ $s_4$ $s_5$
1 $s_1(1)$ $s_2(1)$ $s_3(1)$ $s_4(1)$ $s_5(1)$
2 $s_1(2)$ $s_2(2)$ $s_3(2)$ $s_4(2)$ $s_5(2)$
... ... ... ... ...
T $s_1(T)$ $s_2(T)$ $s_3(T)$ $s_4(T)$ $s_5(T)$

Question: Can I find prediction models $h_i$ for some sensor $s_i$ given a set of other sensors?

Ex: $s_4(t) = h_4(s_1, s_2, s_3) + r_4(t)$
$s_5(t) = h_5(s_1, s_2, s_3) + r_5(t)$
Predictability

- Choice of a prediction model $h_i$ for $s_i$ (linear regression, K-nearest neighbours, neural networks)
- Learning procedure -> identify $h_i$ and estimate residual error $r_i$
- If $f(r_i) < T$, with $f$ and $T$ user defined, $s_i$ is said to be predictable

Illustration

- Fitting: $s_2(t) = h_2(t) + r_2(t)$
  with $h_2(t) = 1.2s_1(t) - 4.3$ (parameters obtained by the least square method)
- Choice for $F(r_2(t))$:
  - $P(|r_2(t)| > \epsilon)$ ($\epsilon$-approximation error)
    - Examples: $P(|r_2(t)| > 0.5) = 0.14$, $P(|r_2(t)| > 1) = 0$
  - $\Sigma r_2(t)^2$ (quadratic error)
  - ...
How to drive the search?

- Ranking criterion C: e.g. energy
- Run a ‘backward search’:
  - Two subsets \{Sq\}<-\{S\} and \{Sp\}<-\{}
  - Remove sensors \(s_i\) (sorted by C) from \{Sq\} and add them to \{Sp\}, if a prediction model \(h_i\) with \(f(r_i)<T\) can be found.

Prediction models

Observation database \(D_T\), collected over a period T

\[
1 \quad s_1 \quad s_2 \quad s_3 \quad s_4 \quad s_5 \\
1 \quad s_1(1) \quad s_2(1) \quad s_3(1) \quad s_4(1) \quad s_5(1) \\
2 \quad s_1(2) \quad s_2(2) \quad s_3(2) \quad s_4(2) \quad s_5(2) \\
\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \\
T \quad s_1(T) \quad s_2(T) \quad s_3(T) \quad s_4(T) \quad s_5(T)
\]

\[
s_4(t) = h_4(s_1, s_2, s_3) + r_4(t) \\
s_5(t) = h_5(s_1, s_2, s_3) + r_5(t)
\]

Suppose we have \(s_4\) and \(s_5\) with lowest remaining energy, and \(f(r_4)<T\), \(f(r_5)<T\)
Energy savings

- From instant $T$ onwards, prediction models $h_4$ and $h_5$ can be used, leaving sensors $s_4$ and $s_5$ in their sleeping modes.

<table>
<thead>
<tr>
<th>$t$</th>
<th>$s_1$</th>
<th>$s_2$</th>
<th>$s_3$</th>
<th>$s_4$</th>
<th>$s_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T+1$</td>
<td>$s_1(T+1)$</td>
<td>$s_2(T+1)$</td>
<td>$s_3(T+1)$</td>
<td>$h_4(T+1)$</td>
<td>$h_5(T+1)$</td>
</tr>
<tr>
<td>$T+2$</td>
<td>$s_1(T+2)$</td>
<td>$s_2(T+2)$</td>
<td>$s_3(T+2)$</td>
<td>$h_4(T+2)$</td>
<td>$h_5(T+2)$</td>
</tr>
<tr>
<td>$T+3$</td>
<td>$s_1(T+3)$</td>
<td>$s_2(T+3)$</td>
<td>$s_3(T+3)$</td>
<td>$h_4(T+3)$</td>
<td>$h_5(T+3)$</td>
</tr>
<tr>
<td>$T+$</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

- But two problems:
  - Unequal energy consumption
  - If dependencies change between $\{s_4, s_5\}$ and $\{s_1, s_2, s_3\}$, prediction models are not valid anymore

Prediction models

Observation database $D_T$, collected over a period $T$

\[ s_4(t) = h_4(s_1, s_2, s_3) + r_4(t) \]
\[ s_5(t) = h_5(s_1, s_2, s_3) + r_5(t) \]
Prediction models

Observation database \( D_T \), collected over a period \( T \)

<table>
<thead>
<tr>
<th></th>
<th>( s_1 )</th>
<th>( s_2 )</th>
<th>( s_3 )</th>
<th>( s_4 )</th>
<th>( s_5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( s_1(1) )</td>
<td>( s_2(1) )</td>
<td>( s_3(1) )</td>
<td>( s_4(1) )</td>
<td>( s_5(1) )</td>
</tr>
<tr>
<td>2</td>
<td>( s_1(2) )</td>
<td>( s_2(2) )</td>
<td>( s_3(2) )</td>
<td>( s_4(2) )</td>
<td>( s_5(2) )</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>( T )</td>
<td>( s_1(T) )</td>
<td>( s_2(T) )</td>
<td>( s_3(T) )</td>
<td>( s_4(T) )</td>
<td>( s_5(T) )</td>
</tr>
</tbody>
</table>

Other predictions possible?

Ex: \( s_1(t) = h_1(s_4, s_5) + r_1(t) \)
\( s_2(t) = h_2(s_4, s_5) + r_2(t) \)
\( s_3(t) = h_3(s_4, s_5) + r_3(t) \)

Cyclic activity schedule

<table>
<thead>
<tr>
<th>( t )</th>
<th>( s_1 )</th>
<th>( s_2 )</th>
<th>( s_3 )</th>
<th>( s_4 )</th>
<th>( s_5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T+1 )</td>
<td>( s_1(T+1) )</td>
<td>( s_2(T+1) )</td>
<td>( s_3(T+1) )</td>
<td>( h_4(T+1) )</td>
<td>( h_5(T+1) )</td>
</tr>
<tr>
<td>( T+2 )</td>
<td>( h_1(T+2) )</td>
<td>( h_2(T+2) )</td>
<td>( h_3(T+2) )</td>
<td>( s_4(T+2) )</td>
<td>( s_5(T+2) )</td>
</tr>
<tr>
<td>( T+3 )</td>
<td>( s_1(T+3) )</td>
<td>( s_2(T+3) )</td>
<td>( s_3(T+3) )</td>
<td>( h_4(T+3) )</td>
<td>( h_5(T+3) )</td>
</tr>
<tr>
<td>( T+4 )</td>
<td>( s_1(T+4) )</td>
<td>( s_2(T+4) )</td>
<td>( s_3(T+4) )</td>
<td>( h_4(T+4) )</td>
<td>( h_5(T+4) )</td>
</tr>
<tr>
<td>( T+5 )</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Cycle length = 3

- An activity schedule is sent to each mote, i.e. in this case: \( s_1 \), \( s_2 \), \( s_3 \)
- This schedule is repeated over time
- Increasing the cycle length allows to solicit more sensors with lower remaining energy

<table>
<thead>
<tr>
<th>( s_1 )</th>
<th>( s_2 )</th>
<th>( s_3 )</th>
<th>( s_4 )</th>
<th>( s_5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( s_1 )</td>
<td>( s_2 )</td>
<td>( s_3 )</td>
<td>( s_4 )</td>
<td>( s_5 )</td>
</tr>
<tr>
<td>( s_1 )</td>
<td>( s_2 )</td>
<td>( s_3 )</td>
<td>( s_4 )</td>
<td>( s_5 )</td>
</tr>
</tbody>
</table>
Results

- 52 sensors, temperature readings from Intel Berkeley lab deployment (2003)
- 2880 readings over 10 days

Average number of sensors used during the tenth day, for different error tolerance parameters:

<table>
<thead>
<tr>
<th>θ</th>
<th>0.9</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>θ =0.1</td>
<td>52</td>
<td>52</td>
</tr>
<tr>
<td>0.5</td>
<td>33</td>
<td>37</td>
</tr>
<tr>
<td>1</td>
<td>24</td>
<td>27</td>
</tr>
<tr>
<td>1.5</td>
<td>17</td>
<td>21</td>
</tr>
<tr>
<td>2</td>
<td>11</td>
<td>16</td>
</tr>
<tr>
<td>2.5</td>
<td>7</td>
<td>11</td>
</tr>
</tbody>
</table>

Current and future work

- Adaptive prediction models
  - Multivariate gaussians
  - Lazy Learning
- Search methods
  - Lasso and PCA
  - Gram schmidt
- Development of a public domain benchmark for these methods
Thank you!

References


- M. Birattari, G. Bontempi, and H. Bersini. « Lazy learning meets the recursive least square algorithm ». NIPS, 1999.