Polling strategies for predictions in wireless sensor networks

LE BORGNE Yann-Aël

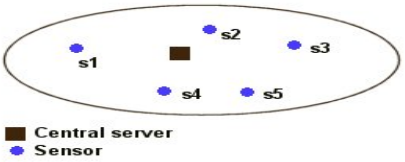
ULB Machine Learning Group
Computer Science Department
Université Libre de Bruxelles
1050 Brussels – Belgium

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



S sensors s_i, 1<=i<=t

Readings are collected from all sensors at regular time intervals

Observation database obtained over time →

t	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	$s_1(2)$	$s_2(2)$	$s_3(2)$	$s_4(2)$	$s_5(2)$
3	$s_1(3)$	$s_2(3)$	$s_3(3)$	$s_4(3)$	$s_5(3)$
…	…	…	…	…	…

2

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

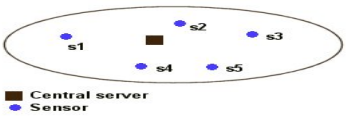
3

Prediction models

Observation database D_T, collected over a period T

t	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	$s_1(T)$	$s_2(T)$	$s_3(T)$	$s_4(T)$	$s_5(T)$

h_4 h_5



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

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

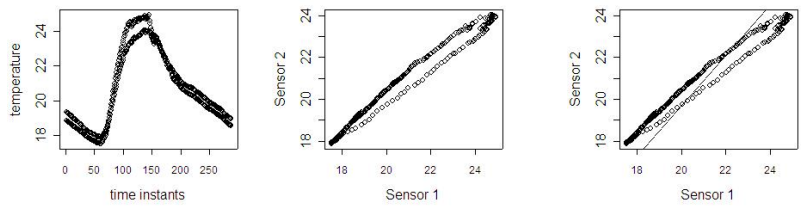
4

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

5

Illustration



•Fitting: $s_2(t)=h_2(t)+r_2(t)$

with $h_2(t)=1.2*s_1(t)-4.3$ (parameters obtained by the least square method)

•Choice for $F(r_2(t))$:

- $P(|r_2(t)|>\varepsilon)$ (ε-approximation error)

  Examples: $P(|r_2(t)|>0.5)=0.14$, $P(|r_2(t)|>1)=0$

- $\Sigma r_2(t)^2$ (quadratic error)
- …

6

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.

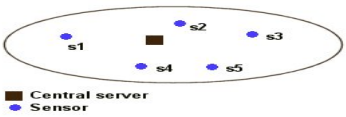
7

Prediction models

Observation database D_T, collected over a period T

t	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	$s_1(T)$	$s_2(T)$	$s_3(T)$	$s_4(T)$	$s_5(T)$

h_4 h_5



$s_4(t)=h_4(s_1, s_2, s_3)+r_4(t)$

$s_5(t)=h_5(s_1, s_2, s_3)+r_5(t)$

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

8

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.

t	s_1	s_2	s_3	s_4	s_5
T+1	$s_1(T+1)$	$s_2(T+1)$	$s_3(T+1)$	$h_4(T+1)$	$h_5(T+1)$
T+2	$s_1(T+2)$	$s_2(T+2)$	$s_3(T+2)$	$h_4(T+2)$	$h_5(T+2)$
T+3	$s_1(T+3)$	$s_2(T+3)$	$s_3(T+3)$	$h_4(T+3)$	$h_5(T+3)$
T+…	…	…	…	…	…

- But two problems:
  - Unequal energy consumption
  - If dependencies change between {s4,s5} and {s1,s2,s3}, prediciton models are not valid anymore

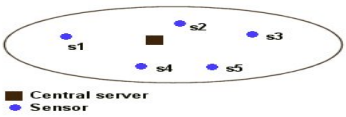
9

Prediction models

Observation database D_T, collected over a period T

t	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	$s_1(T)$	$s_2(T)$	$s_3(T)$	$s_4(T)$	$s_5(T)$

h_4 h_5



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

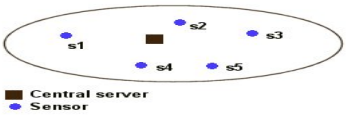
10

Prediction models

Observation database D_T, collected over a period T

t	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	$s_1(T)$	$s_2(T)$	$s_3(T)$	$s_4(T)$	$s_5(T)$

h_1 h_2 h_3



Other predictions possible?

Ex: $s_1(t)=h_1(s_4, 5_4)+r_1(t)$

$s_2(t)=h_2(s_4,s_5)+r_2(t)$

$s_3(t)=h_3(s_4,s_5)+r_4(t)$

11

Cyclic activity schedule

t	s_1	s_2	s_3	s_4	s_5
T+1	$s_1(T+1)$	$s_2(T+1)$	$s_3(T+1)$	$h_4(T+1)$	$h_5(T+1)$
T+2	$h_1(T+2)$	$h_2(T+2)$	$h_3(T+2)$	$s_4(T+2)$	$s_5(T+2)$
T+3	$s_1(T+3)$	$s_2(T+3)$	$s_3(T+3)$	$h_4(T+3)$	$h_5(T+3)$
T+4	$s_1(T+4)$	$s_2(T+4)$	$s_3(T+4)$	$h_4(T+4)$	$h_5(T+4)$
T+5	…	…	…	…	…

Cycle length=3

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

s1	***Send***	***Idle***	***Send***
s2	***Send***	***Idle***	***Send***
s3	***Idle***	***Send***	***Idle***
s4	***Idle***	***Send***	***Idle***
s5	***Idle***	***Send***	***Idle***

12

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:

	T=P(|r|>ε)	
	0.9	0.95
ε =0.1	*52*	*52*
0.5	*33*	*37*
1	*24*	*27*
1.5	*17*	*21*
2	*11*	*16*
2.5	*7*	*11*

•See [*Y. Le Borgne, G. Bontempi, « Round Robin Cycle for Predictions in Wireless Sensor Networks », submitted to ISSNIP 2005].*

13

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

14

Thank you!

References

- Y. Le Borgne, G. Bontempi. « Round robin cycle for predictions in wireless sensor networks ». Submitted to ISSNIP 2005..
- A. Desphande, C. Guestrin, S. Madden, J. Hellerstein, W. Hong. « Model driven data acquisition in sensor networks ». Proc. of the 30th VLDB conference, 2004.
- M. Birattari, G. Bontempi, and H. Bersini. « Lazy learning meets the recursive least square algorithm ». *NIPS*,1999.
- T. Mitchell. « Machine Learning », 1997.