Model-Aided Data Collecting for Wireless Sensor Networks

Chongqing Zhang, Minglu Li, and Min-You Wu

Department of Computer Science and Engineering,
Shanghai Jiaotong University, Shanghai, China
zhangchongqing@sjtu.edu.cn

Abstract. In this paper, we address the problem of collecting data from sensor nodes using a model-aided approach. In our approach, a model is maintained by a node and a replica of the model is maintained the base station. The base station uses the replica model to estimate the actual measurement data of the sensor node in usual time, and an actual measurement datum is sent to the base station only when the error of the model's corresponding estimation exceeds allowable error bound. In such a way, energy can be saved by reducing the transmission of actual measurement data. Experimental results show the effectiveness of our approach.

Keywords: Wireless Sensor Networks, Model-aided, Data Collecting, Data Fitting.

1 Introduction

Wireless sensor networks can offer us revolutionary new methods of accessing data from real environment [1]. However, because of the limited power of sensor nodes, collecting data is still a challenging work. For example, a Berkeley mote is only powered by two alkaline AA batteries [2]. Furthermore, it is infeasible to replenish the energy of sensor nodes by replacing the batteries in many applications [1]. Therefore, data collecting approaches of high energy-efficiency are strongly needed.

Motivated by the need of extending the network lifetime of energy-constrained wireless sensor networks, there has been considerable research in the area of energy-efficient data collecting in sensor networks and many techniques [3, 4, 5, 6, 7, 8, 9, 10, 14] have been proposed and developed. Among these techniques in-network aggregation and compression are two noticeable techniques. Although the measures they take are different, they are both trying to save energy by reducing the total amount of data transmitted.

Aggregation [3] is an in-network query processing technique for wireless sensor networks. By such a technique, for an aggregation query (e.g., the average rainfall of the monitored area), sensor readings are accumulated into partial results that are combined as messages propagate toward the base station. TinyDB [4] and Cougar [5] are two examples of utilizing aggregation to reduce energy consumption. On the other hand, compression attempts to take advantage of the correlation in the data and exploit coding techniques to reduce the size of data transmitted. For example, in [6],

M. Gerndt and D. Kranzlmüller (Eds.): HPCC 2006, LNCS 4208, pp. 692 – 699, 2006.

Ganesan et al. used wavelet based approach to compress the sensor data; while in [7] Chou et al. used distributed source coding to reduce the redundancy of the data to be transmitted to the sink.

All sensor nodes are still needed to transmit their data both in aggregation and compression. In [10], a model-aided approach was proposed to overcome this problem. However, the models adopted in this approach are fixed and not adaptive to the phenomenal changes. In this paper, adaptive models are adopted to improve the model-aided approach. In our approach, a predictive model M_i is induced by a sensor node N_i using data fitting [12] and an identical model M_i' is sent to the base station. The base station utilizes M_i' to estimate the actual measurement data of node N_i. At the same time, M_i is used by node N_i to judge how the estimations of model M_i agree with the actual measurement data. A measurement datum will be reported to the base station only when the error of corresponding estimative figure exceeds allowable error bound. In such a way, communication cost can be reduced and measurement error can be controlled in an allowable range.

The rest of this paper is organized as follows. In section II, We give the WSN model on which our research are based and present an overview of our approach. In section III, we discuss our approach in detail. In section IV, the implementation issues are discussed. Experimental results are presented in section V to show the effectiveness of our approach. We conclude in section VI.

2 Overview of Approach

We give an overview of our approach using an example of monitoring the blood pressure of patients in a hospital. Fig. 1 gives how the blood pressure of hypertension patients changes in 24 hours [11]. DBP and SBP denote diastolic blood pressure and systolic blood pressure, while SH and EH represent secondary hypertension and essential hypertension. As the figure shows, the blood pressure does not change desultorily. On the contrary, it fluctuates cyclically (with a period of 24 hours). The blood pressure changes continuously and it reaches its highest and lowest points at approximately 8 o'clock AM and 2 o'clock PM. Our approach uses these rules to achieve its energy-efficiency.

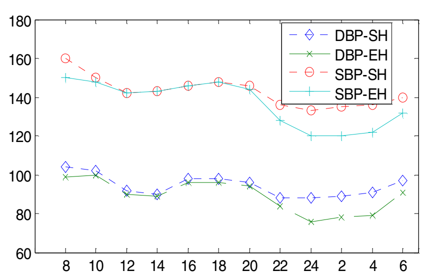

Fig. 1. Blood pressure curve in 24 hours [11]

Fig. 2 gives an overview of our approach. The base station uses commands to request the sensor nodes to sense the patients' blood pressure for a period of time *T* and at a certain frequency *f*, with an error bound ε is allowed. As the figure shows, a pair of models is maintained, with one model M_i distributed on node N_i and the other M_i' on *BS*. M_i and M_i' are always kept in synchronization. Model M_i is induced by a light-weight algorithm running on N_i from the measurement data set. Assume at a time instant *t*, a copy M_i' of M_i is sent to *BS*. Then at next time instant $t + 1$, *BS* can utilize model M_i' to estimate the actual measurement data of the sensor node N_i. At the same time, N_i still measures the blood pressure and compares the estimation E_i^{t+1} of M_i with the actual measurement data X_i^{t+1}. If $|E_i^{t+1} - X_i^{t+1}| \le \varepsilon$ (ε is the allowable error bound), the measurement data X_i^{t+1} is not reported to *BS*, otherwise X_i^{t+1} is reported to *BS*.

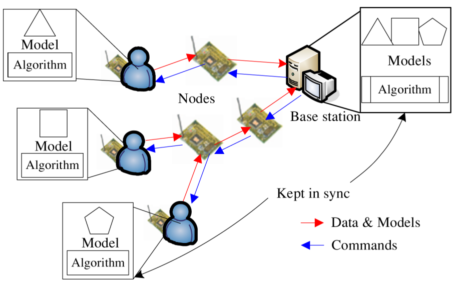

Fig. 2. Overview of Approach

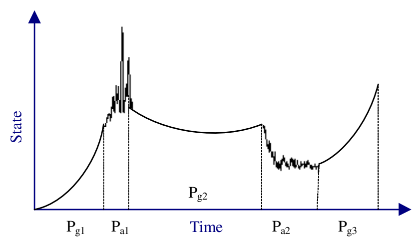

Fig. 3. Gradual changes versus abrupt changes

3 Principle of Approach

3.1 Predictability of Phenomena

The changes of natural phenomena follow some temporal and spatial rules. In this paper we focus on the temporal rules of the changing processes of physical phenomena. One temporal rule is that the changing process of a phenomenon may consist of gradual phases and abrupt phases. An example is the air temperature in a garden. The air temperature may change rapidly and violently in a short time, yet in most of the time it changes slowly and smoothly. Figure 3 shows a changing process that has gradual phases: Pg1, Pg2, Pg3, and abrupt phases: Pa1 and Pa2. During a gradual phase, the state of the phenomenon changes gradually and continuously; while the state of the phenomenon changes rapidly and discontinuously during an abrupt phase.

In many cases, the continuity and gradualness of a gradual phase make it possible to predict the state after time t by the state before time t. For example, we can predict the air temperature in an hour by how the air temperature changes before now. It is this predictability that enables our approach to achieve its energy-efficiency. Examples of such kind of phenomena include: air temperature, air humidity, earth temperature, soil fertility, soil humidity, body temperature, blood pressure, health of machines or buildings, pH value of lake water, concentration of pollutant, diffusion of contaminants, etc.

As for some phenomena, predicting the future state by previous state is very hard, sometimes even impossible. For example, the irregularity of the noise in a workshop makes it difficult to predict its intensity in the future. Our approach is not applicable for monitoring such kind of phenomena.

3.2 Models

Problem Definition. For a sensor node N_i, given the measurement data $\{X_i^0, X_i^1, \ldots, X_i^t\}$ before time instant t, an error bound ε, conclude a predictive model M that minimizes $Num(E_i^{t+a}: |E_i^{t+a} - X_i^{t+a}| \leq \varepsilon)$, where $a \geq 1$.

However, at time instant t, $\{X_i^{t+a}, X_i^{t+a+1} \ldots\}$ are unknown. As a result, these data cannot help us to figure out model M. What we can depend on is the data set $\{X_i^0$, X_i^1, …, X_it}. So what we should do is to derive a proper model M from {Xi0, Xi1, …, Xit} and hope the prediction of M will agree with the actual future measurement data.

The continuity and gradualness of a gradual phase make it can be represented as a unitary function or several unitary functions with time as the independent variables. Based on this, unitary functions with time as the independent variables are adopted as models depicting how the monitored phenomenon changes in a gradual phase. Assume a unitary function f(x) for a phase Pg is derived at time instant t and sent to BS, then BS can use f(x) to estimate the actual state after time instant t.

From above analysis, it can be seen the key problem of our approach is to derive the function f(x) from limited measurement data. This problem can be viewed as a data fitting problem [12]. There are generally three problems to solve: 1) identifying a target function with unknown parameters, 2) identifying a proper data set and 3) determining the unknown parameters of the target function. Problem 2 and 3 are answered in following sections. Here we answer how to solve problem 1.

What target functions should be adopted is strongly application-dependent. As for different applications, the target functions that should be adopted may be quite different. If the change of the monitored phenomenon follows an obvious function type, then we have an obvious choice. Otherwise, if the function *f(x)* is continuous and has $n + 1$continuous derivatives on an open interval (a, b), then according to Taylor Theorem [13], for $x \in (a, b)$, *f(x)* can be represented as the summation of a polynomial of $(x - x_0)$ and a remainder:

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)(x - x_0) + f''(x_0)\frac{(x - x_0)^2}{2!} \\ &+ \cdots + f^n(x_0)\frac{(x - x_0)^n}{n!} + R_n(x) \end{aligned} \tag{1}$$

where $x_0 \in (a, b)$. Based on this, polynomial functions can be adopted to fit the measurement data if there is not an obvious choice.

Note that time also affects the selection of target functions. For example, it can be seen from Fig.1 that the blood pressure curve takes on different shapes in different time phases of a day. As a consequence, using only one function to model the blood pressure curve of a whole day is not appropriate. The proper way is that different functions should be taken to model the blood pressure curve in different time phases. By storing multiple target functions for different time phases on nodes, proper functions can be selected according to the time when the model is building.

Although unitary functions are adopted in this paper to represent how the monitored phenomenon changes, this does not mean unitary functions are the only models that can be used. We think adopting what modeling tool is highly application-dependent.

4 Implementation of Approach

4.1 Building the Models

Identifying the data set concerns how to determine the number and location of the data. At a time instant *t*, what is known to node N_i is the measurement data $\{X_i^0, X_i^1, \ldots, X_i^t\}$ before *t*, and what we need to do is to derive a function that can predict the measurement data after *t*. Generally speaking, data that are adjacent in time are more correlated [15]. Based on this, we use the data measured in an interval before time t to fit the target function. As for the length of the interval, on the premise of that the target models can be constructed successfully; the length should be as small as possible. In the experiments presented in section V, an interval contains only 5 data, yet the result is very satisfactory.

Suppose the chosen data set is (x_i, y_i), $i = 1, 2, \ldots, n$, $x_i \neq x_j$ if $i \neq j$. Let $f(t)$ be the target function and $f(t)$ is expressed as:

$$f(t) = a_1 r_1(t) + a_2 r_2(t) + \ldots + a_m r_m(t) \tag{2}$$

where $\{r_1(t) \ldots r_m(t)\}$ is a group of predefined functions. $\{a_1 \ldots a_m\}$ is the coefficients that need to be determined. Least square fitting [12] is adopted by *MADG* to determine $\{a_1 \ldots a_m\}$. To do this, we need to make expression (2) get the least value.

$$J(a_1, \ldots, a_m) = \sum_{i=1}^{n} [f(x_i) - y_i]^2 \tag{3}$$

According to the necessary conditions for an extremum: $\partial J / \partial a_k = 0$, where $k = 1, 2, \ldots, m$, following group equation is derived:

$$\begin{cases} \sum_{i=1}^{n} r_1(x_i)[\sum_{k=1}^{m} a_k r_k(x_i) - y_i] = 0 \\ \quad\vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ \sum_{i=1}^{n} r_m(x_i)[\sum_{k=1}^{m} a_k r_k(x_i) - y_i] = 0 \end{cases} \tag{4}$$

Group equation can be expressed as:

$$R^T R A = R^T Y \tag{5}$$

where

$$R=\begin{bmatrix} r_1(x_1) & r_2(x_1)\cdots r_m(x_1) \\ r_1(x_2) & r_2(x_2)\cdots r_m(x_2) \\ \vdots & \vdots \quad \vdots \\ r_1(x_n) & r_2(x_n)\cdots r_m(x_n) \end{bmatrix},\ A=(a_1,\ldots a_m)^T \text{ and } Y=(y_1,\ldots y_n). \tag{6}$$

By solving group equation (5), $\{a_1 \ldots a_m\}$ can be derived. If $r_1(t) \ldots r_m(t)$ are linearly independent, then R^TR is reversible, and equation (5) has sole solution. In our approach, this is guaranteed by always using polynomial functions as the target functions.

4.2 Algorithm

In our approach, most of the work is done on sensor nodes. The base station simply uses the models to compute the expected value. A node N_i performs operations shown in Fig. 4 when a data is measured.

In the algorithm, data fitting is done only when there are enough data in the data set. Lines 13 to 18 are used to guarantee the effectiveness of the model. This can ensure the gain in performance if a correct model is used, and also the performance does not degrade if there is not an accurate model.

```
1: while measures a data X do {
2:    Add measurement X to DataSet;
3:    If X is the first data then {
4:        set y = X as the model  and send model to BS;
5:        continue; }
6:   E = estimate of model;
7:   If |E – X| < ε then continue;
8:   If Num(DataSet) < DataNum then {
9:        set y = X as the model and send model to BS;
10:       continue; }
11:  f(x) = data fitting result;
12:  bool = false;
13:  For each data d in DataSet do {
14:      E = estimate of f(x);
15:      If | E – d | > ε then {
16:          bool = true;
17:          set y = X as the model and send model to BS;
18:          break; } }
19:  If bool then continue;
20:  Set f(x) as model and send model to BS; }
```

Fig. 4. Algorithm running on a node

5 Experimental Results

The performance of our approach (denoted as Model-Aided) was tested against other two approaches. In one approach (denoted as Simple), all data are sent to the base station. In the other approach (denoted as Cache), the latest measurement data X_l of a node is cached by the node and the base station. A measurement data X_c is sent to the base station only when the absolute value of the difference between X_l and X_c is bigger than error bound ε.

MicaZ motes [16] are used to test the performances of all approaches. Four motes are deployed to monitor the air temperature in the garden outside of our laboratory. We monitored the temperature for 5 days. The monitoring results of four motes are quite similar. Fig. 5 shows the air temperature curve that is drawn from the data collected by a node using approach Simple in 40 hours.

The number of transmitted packets is adopted to evaluate the performances of three approaches. For sake of simplicity, a measurement datum or a model is both regarded as a data packet. A node measures the air temperature every 1 minute. The allowable error bound is 0.05 Celsius. Only linear functions are taken to fit the temperature data. A dataset includes 5 data, i.e. data measured in 5 minutes.

Fig. 6 presents the comparative results of three approaches in cost which is evaluated by the number of sent packets. From the figure, it can be seen that even the performance of Cache is much better than Simple. By adopting models, our approach achieves better performance further than Cache.

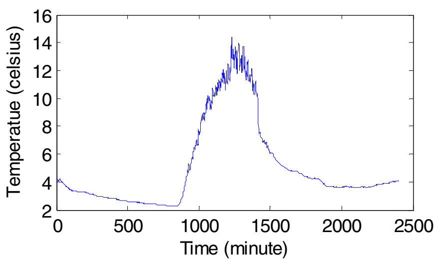

Fig. 5. Air Temperature Curve

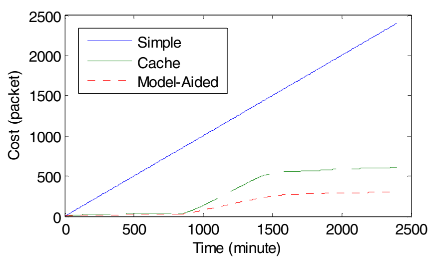

Fig. 6. Packets Sent of Three Approaches

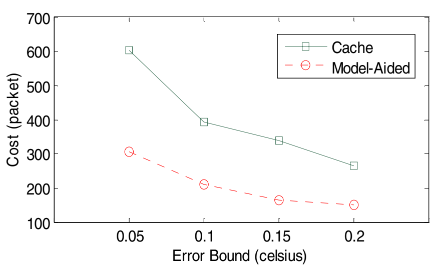

Fig. 7. Error Bounds versus Cost

Fig. 7 reveals the comparative results of two approaches, Cache and Model-Aided, against the error bound. Approach Simple is not compared because the costs of Simple are identical under different error bounds. In both Cache and Model-Aided, the number of sent packets drops as the error bound increases. It can also be observed that Model-Aided outscores Cache under all error bounds.

6 Conclusion

In this paper, temporal rules of the changing processes of natural phenomena are exploited to improve the energy-efficiency of wireless sensor networks. By maintaining replicated models on sensor nodes and the base station, energy can be saved by reducing the data transmitted to the base station. In the next step, we plan to integrate model-aided approach with spatial rules of natural phenomena and make full use of spatio-temporal rules to heighten the energy-efficiency of wireless sensor network.

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless Sensor Networks: A Survey," Computer Networks, Vol. 38, No. 4, pp. 393-422, March 2002.
[2] M. Horton, D. Culler, K. PIster, J. Hill, R. Szewczyk, and A. Woo, "MICA, The Commercialization of Microsensor Motes," Sensors, Vol. 19, No. 4, pp 40-48, April 2002.
[3] B. Krishnamachari, D. Estrin and S. Wicker, "The Impact of Data Aggregation in Wireless Sensor Networks," DEBS, 2002.
[4] S. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, "Tag: A tiny aggregation service for ad hoc sensor networks," OSDI, 2002.
[5] Y. Yao and J. Gehrke, "Query processing in sensor networks," CIDR, 2003.
[6] D. Ganesan, B. Greenstein, D. Perelyubskiy, D. Estrin, and J. Heidemann. An Evaluation of Multi-resolution Search and Storage in Resource-constrained Sensor Networks. ACM SenSys 2003.
[7] J. Chou, D. Petrovic, and K. Ramchandran, "A distributed and adaptive signal processing approach to reducing energy consumption in sensor networks," INFOCOM, 2003.
[8] B. Babcock and C. Olston, "Distributed Top-K Monitoring," ACM SIGMOD 2003.
[9] A. Deligiannakis, Y. Kotidis and N. Roussopoulos, "Hierarchical In-Network Data Aggregation with Quality Guarantees," EDBT 2004.
[10] D. Chu, A. Deshpande, J. M. Hellerstein and W. Hong, "Approximate Data Collection in Sensor Networks using Probabilistic Models," ICDE 2006.
[11] C. Jin, Z. Qian, L. Chen, X. Wang, "Ambulatory blood pressure monitoring in secondary hypertension," Chinese Journal of Cardiology, Vol 27, No 3, pp. 50-53, May 1999.
[12] R. L. Burden, J. D. Faires, "Numerical Analysis," Brooks Cole publishing company, December 2000.
[13] D. Hughes-Hallett, A. M. Gleason, P. F. Lock, D. E. Flath, et al, "Applied Calculus," Wiley, April 2002.
[14] C. Intanagonwiwat, R. Govindan and D. Estrin, "Directed diffusion: A scalable and robust communication paradigm for sensor networks," MOBICOM 2000.
[15] I. F. Akyildiz, M. C. Vuran, O. B. Akan, "On Exploiting Spatial and Temporal Correlation in Wireless Sensor Networks," WiOpt 2004.
[16] Crossbow, Inc. Wireless sensor networks. http://www. xbow.com.