ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 2, Issue 3, March 2013

Prediction Based Data Collection in Wireless Sensor Network

V.Mathibala , Mrs. Seiya Susan Thomas.

Abstract**—** A wireless sensor network is a network that is made of hundreds or thousands of sensor nodes which are densely deployed in an unattended environment with the capabilities of sensing, wireless communications and computations which collects and disseminates environmental data. For many applications in wireless sensor networks, users may want to continuously extract data from the networks for analysis later. However, accurate data extraction is difficult and it is often too costly to obtain all sensor readings, as well as not necessary in the sense that the readings themselves only represent samples of the true state of the world.

Energy conservation is crucial to the prolonged lifetime of a sensor network. Energy consumption can be reduced for data collections from sensor nodes using prediction. The prediction based algorithms are based on the observation that the sensor capable of local computation generates the possibility of training and using predictors in a distributed way.

An energy efficient framework for clustering based data collection in wireless sensor networks can be done by adaptively integrating enabling/disabling prediction scheme with sleep/awake. The framework consists of a number of sensor nodes which form clusters. Each cluster has a cluster head and set of sensor nodes attached to it. Cluster head collects the data value from its member nodes .The prediction incorporated in the member nodes imply that sensors need not to transmit the data if it does not differ from a predicted value by a certain threshold. A member need not be awake if no data value has to be transmitted but only has to periodically check the data values and awake only if it differs from the predicted value. If prediction is disabled it simply transmits the data values. The performance of power saving in clustering based prediction is evaluated by creating a network scenario for tracking a moving object in NS-2.33 simulator.
.

Index Terms: Cluster head, Prediction operation, Sleep/Awake Algorithm.

INTRODUCTION

Wireless sensor network (WSN) refers to a group of spatially dispersed and dedicated sensors for monitoring and recording the physical conditions. Wireless Sensor Network comprises of numerous sensors and they are interlinked or connected with each other for performing the same function collectively or cooperatively. It is a network that is made of hundreds or thousands of sensor nodes which are densely deployed in an unattended environment with the capabilities of sensing, wireless communications and computations. These spatially distributed autonomous devices cooperatively monitor physical and environmental conditions, such as temperature, sound, vibration, pressure, motion or pollutants, at different locations.Wireless sensor networks have critical applications in the scientific, medical, commercial and military domains. Examples of these applications include environmental monitoring, smart homes and offices, surveillance, and intelligent transportation systems. It also has significant usages in biomedical field. As social reliance on wireless sensor network technology increases, they can expect the size and complexity of individual networks as well as the number of networks to increase dramatically. Wireless sensor networks are typically used in highly dynamic, and hostile environments with no human existence (unlike conventional data networks), and therefore, they must be tolerant to the failure and loss of connectivity of individual node. The sensor nodes should be intelligent to recover from failures with minimum human involvement. Networks should support process of autonomous formation of connectivity, addressing, and routing structure. Architecture of wireless sensor network is designed taking all these factors into consideration. The architecture consists of

.

a. The sensor nodes that form the sensor network. Their main objectives are making discrete, local measurement about phenomenon surrounding these sensors, forming a wireless network by communicating over a wireless medium, and collect data and route data back to the user via sink (Base Station).

b. The sink (Base Station) communicates with the user via internet or satellite communication. It is located near the sensor field or well-equipped nodes of the sensor network. Collected data from the sensor field routed back to the sink by a multi-hop infrastructure less architecture through the sink.

A. ENERGY CONSIDERATION IN WIRELESS NETWORKS

Energy conservation is a critical issue in the design of sensor networks since the sensor nodes are battery-powered. WSN has been considered as a promising method for reliably monitoring both civil and military environments under hazardous conditions. Due to such condition, the power supply for sensor in the network cannot be usually rechargeable or replaceable. The design of protocols and applications for such networks has to be energy aware in order to prolong the lifetime of the network because it is quite difficult to recharge node batteries. A node in WSN depends on batteries for their energy source. However, since a battery's lifetime is limited, the power resource is at a premium. But wireless signal transmission, reception, retransmission, and beaconing operations all consume battery power. Energy efficiency in mobile nodes can be achieved through improvement in various levels, including the communication, and application layers [4]. For example, the power management feature in 802.11 cards allows two modes of operation, the active mode and power save mode. During the active mode, the wireless card is always ready to transmit or receive frames in accordance with the specifications of the 802.11 medium access control protocols. In the power save mode, nodes are temporarily put to sleep and are awake only in scheduled time intervals for short durations.

The computing components used in a mobile node, such as processors, memory and input/output devices; usually have low capacity and limited processing power [9]. Therefore, algorithms for communication protocols need to be lightweight in terms of computational and storage requirements. The existing research on energy consumption of sensors is usually based on either theoretical models or computer simulations. One widely cited model of energy consumption has been used extensively as a guide for simulations and the design of low power consumption communication protocols [3] One solution is clustering-based localized prediction[10], where a cluster head also a sensor node maintains a set of history data of each sensor node within a cluster. They expect the use of localized prediction techniques is highly energy efficient due to the reduced length of routing path for transmitting sensor data. On the downside, clustering based local prediction in sensor networks faces a couple of new challenges. First, since the cost of training a predictor is nontrivial, they should carefully investigate the trade-off between communication and computation. To support prediction techniques, energy is consumed on communication(e.g., sending and receiving sensor data) and computation(e.g., processing sensor data and calculating a predicted value). Motivated by this observation, we analytically study how to determine whether a prediction technique is beneficial in this paper, qualitatively derive sufficient conditions for this and reveal that the decision is a function of both the desired error bound and the correlation among the sensor data values. For instance, when the error bound is very tight or the correlation is not significant, a sensor node always has to send its data to the cluster head. The second challenge is due to the characteristics and inherent dynamics of the sensor data. When the data distribution, in particular the data locality, evolves overtime, prediction techniques may not work well for a set of less predictable data. Global re clustering is costly if it is initiated periodically.

B. **Energy Consumption and Predictor**

A sensor network consists of a set of autonomous sensor nodes which spontaneously create communication links, and then, collectively perform tasks without help from any central servers. In sensor networks, accurate data extraction is difficult; it is often too costly to obtain all sensor readings, as well as not necessary in the sense that the readings themselves only represent samples of the true state of the world. As such, one technique so called prediction emerges to exploit the temporal correlation of sensor data. Technology trends in recent years have resulted in sensors' increasing processing power and capacity [5]. Implementing more sophisticated distributed algorithms in a sensor network becomes possible.

Predictor is one of the important techniques, which uses the past input values from the sensors in order to perform the prediction operations. The existence of such prediction capability implies that the sensors do not need to transmit the data values if they differ from a predicted value by less than a certain prespecified threshold, or error bound. A simple approach to developing a predictor in sensor networks is simply to transmit the data from all sensors to the base station (i.e., the sink), which has been realized in many previous studies [6], [10], and [2]. Predictor training and prediction operations are carried out by the base station only, but not the sensor nodes, despite their increasing computing capacity.

This solution while practical has many disadvantages, such as a high energy consumption incurred by transmitting the raw data to the base station, the need for wireless link bandwidth, and potential high latency.

1) Adaptive scheme to enable/disable prediction Operations

Consider a cluster of sensor nodes, which can be awake or sleeping. If the sensor nodes are sleeping, the prediction problem is reduced to estimating data distribution parameters using history data. In this case the estimates are already available. If the sensor nodes are awake, they continuously monitor an attribute X and generate a data value x_t at every time instance t. Without local prediction capability at the cluster head, a sensor node has to send all data values to the cluster head that estimates data distribution accordingly. With local prediction, however, a sensor node can selectively send its data values to the cluster head[10].

One model for selective sending is $\in -$loss approximation. Given an error bound $\in > 0$ a sensor node sends its value x_t to the cluster head if $|x_t - \hat{x}_t| > \in$ the where $\hat{x}_t$ is a predicted representative data value to approximate the true data. The intuition of this choice is that if a value is close to the predicted value there is not much benefit by reporting it. If the value is much different from the predicted value, it is important to consider it for computing the data distribution. To solve this problem, first develop a localized prediction model. Very complex models are not practical in the application due to the limited computational capacity of sensor nodes. Fortunately, simple linear predictors are sufficient to capture the temporal correlation of realistic sensor data as [1, 6]. A history based linear predictor is one of popular approaches to predicting the future based on past n measurement $\hat{x} = p(x_{t-1}, x_{t-2},..., x_{t-n})$. In particular, it has the properties: 1) $k\hat{x}_t = p(kx_{t-1}, kx_{t-2},..., kx_{t-n})$ and 2) $p'_{x_{t-i}}(.) = c,$ where c is constant, and $\sum_{i=1}^{n} p'_{x_{t-i}}(.) = 1$ if the predictor is unbiased $x_t = c + w_t + \sum_{i=1}^{p} \phi_i\, x_{t-i,}$. One of the examples, autoregressive model denoted by $AR(p)$ is written as where $\phi_1, ..., \phi_p$ are the parameters of the model. Here, c is a constant always presumed to be zero and w_t is a white noise process with zero mean and variance σ^2. The process is covariance stationary if $|\phi_i| < 1$. Accordingly, a p -order AR predictor is: $\hat{x}_t = \sum_{i=1}^{p} \phi_i\, x_{t-i}$. The parameters can be calculated by Yule-Walker equations or using the least square method. The parameters are often updated on every measurement and the estimation is carried out by both the cluster header and 2 cluster members to achieve synchronization. If the error tolerant $\in$ and correlation $\rho_x(j)$ satisfy

$$\frac{\in}{\sqrt{1 - \sum_{j=1}^{p} \phi_j\, \rho_x(j)}} > m\sigma_x \Phi^{-1}\left(\frac{k+1}{2k}\right) \quad ....(1)$$

(1) the scheme with local prediction is more energy efficient, where the parameters are given as in table (1.2). Equation (1) gives the error bound condition. This result says, if the correlation coefficient is too small, prediction will not be accurate. As a result, sensor data values are often not within the error bound and will still be transmitted to the cluster head. Meanwhile, if the error bound is small, the condition is not easily satisfied either and transmissions are still required. Together, this corollary tells that algorithm selection should be determined based on both the desired error bound and the predictability of the sensor data and from experimental results the effect on the parameter m is not deterministic. While a large value of 'm' often leads to the condition of (1) hardly satisfied and there by the prediction scheme at the node is prone to being disabled, it can reduce the energy consumption as a long term predictor when the condition is satisfied.

Process at the Cluster Head:
if timeout after $m * \Delta$seconds
for each member i in the Cluster
if Condition (1) holds
Send message to member i to enable prediction
else
Send message to member i to disable prediction
Else
for each member i in this cluster
if receive a data value from member i
update the history data for member i
else
Perform Prediction to update the history data

Algorithm 1.1 Operations at the Cluster Head

ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 2, Issue 3, March 2013

Algorithm 1.1 shows the pseudo code description of the algorithms at the cluster head the cluster head maintains a set (a circular array) of history data for each cluster member. The algorithm shows the cluster head will continuously receive data values from each cluster member to update the set of history data or when no data values are received will use the predicted value instead for update. The cluster head also runs a periodic process, to determine algorithm selection, with or without local prediction. The decision is broadcast to all cluster members.

Process at the cluster members

if prediction is disabled or $\left| \chi_i - \hat{\chi}_i \right| > \in$

send the data value to the cluster head

update the history data using the data value

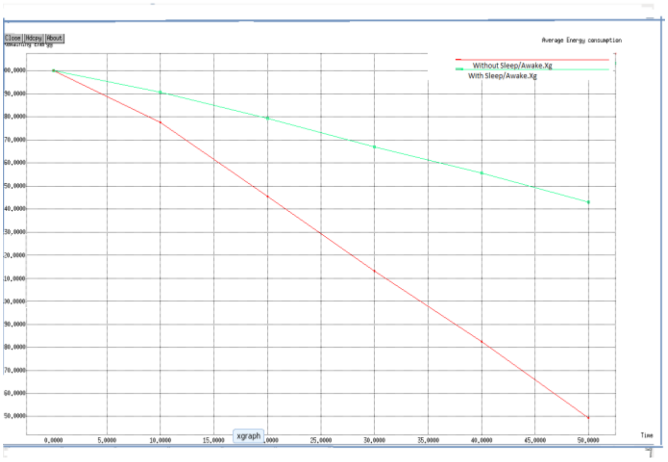
else

Perform prediction to update the history data

Algorithm 1.2 Operations at cluster members

Algorithm 1.2 is the pseudo code description at each cluster node. Each cluster member maintains a set of history data of its own. If "no local prediction", is selected it simply transmits the data values. If local prediction is turned on, the cluster member will perform prediction on each data value. If the data value is not within the error bound, it will be sent to the cluster head too. Meanwhile, the local set of history data should be updated as well. In particular, if local prediction is enabled and the data value is within the error bound, the predicted value not the actual value will be included in the set of history data. The purpose is to maintain the consistency between local and the cluster head.

The result from the simulation has shown below in the line graph with variable nodes. The simulation shows that enable/disable with sleep/awake algorithm consumed less energy when compared to the existing scheme.



Average Energy Consumption

CONCLUSION

A Framework for clustering based data collection with prediction for wireless sensor network has been described. The detailed analysis and description of its two main components, adaptive scheme to enable/disable prediction and integration of sleep/awake algorithms for nodes is presented. An energy-aware predictor is used to find the trade-off between communication and prediction cost. By using sleep/awake Algorithm energy consumption is reduced and network life time is increased.

The usefulness of the framework is demonstrated by accommodating moving object into the area. Via performance evaluation, it has been shown that it achieves energy efficiency when sensor data is spatially and temporally correlated. To summarize, the framework demonstrates that it is viable framework to facilitate data collection in large-scale wireless sensor networks.

REFERENCES

[1] Chu . D, Deshpande . A, Heller stein. J.M. and W. Hong, (2006) "Approximate Data Collection in Sensor Networks Using Probabilistic Models", Proc. IEEE Int'l Conf. Data Eng. (ICDE).

[2] Goel. S and Imielinski. T (2001), "Prediction-Based Monitoring in Sensor Networks: Taking Lessons from Mpeg," Acme Sitcom Computer Comm. Rev., vol. 31, no. 5, pp. 82-98.

[3] Heinemann. W, Chandrakasan .A and Balakrishnan .H, (Jan, 2000) "Energy-Efficient Communication Protocol for Wireless Micro sensor Networks", In Proc. 33rd *HICS*, 4-7.Vol 2, pp. 10.

[4] Jiang Et Al Hongbo, Shudong Jin, Chonggang Wang, (June 2011), Prediction or Not? An Energy-Efficient Framework for Clustering Based Data Collection in Wireless Sensor Networks. IEEE Transactions on Parallel and Distributed Systems, Vol. 22, No. 6,

[5] McConnell .S.M. and Skill corn. D.B. (2005) "A Distributed Approach for Prediction in Sensor Networks," Proc. SIAM Int'l Conf. Data Mining Workshop Sensor Networks.

[6] Meka. A and Singh. A. K (2006) Distributed spatial clustering in sensor networks. In Proceedings of EDBT,

[7] Ossama Younis and Sonia Fahmy. (2004). Distributed Clustering in Ad-hoc Sensor Networks: A Hybrid, Energy Efficient Approach. In Proceedings of IEEE INFOCOM, Hong Kong, an extended version appeared in IEEE Transactions on Mobile Computing, 3(4).

[8] Toh C-K (2007), Ad Hoc Mobile Wireless Networks-Protocols and Systems, Pearson Education, Edition-, ISBN- 978-8-1317-1510-9.

[9] Tulane. D and Madden. S, (2006) "An Energy-Efficient Querying Framework in Sensor Networks for Detecting Node Similarities, "Proc. IEEE/ACM Int'l Conf. Modeling, Analysis, and Simulation of Wireless and Mobile Systems (Swim).

[10] Xu . Y. and W.-C. Lee, (2003) " On Localized Prediction for Power Efficient Object Tracking in Sensor Networks," International Conference of Distributed Computing Systems Workshop, pp. 434-439.