Energy Reduction in Wireless Sensor Networks through Measurement Estimation with Second Order Recurrent Neural Networks 9 14 26/12	Giảm Năng Lượng trong Các Mạng Cảm Biến Không Dây nhờ vào Ước Lượng Đo với Các Mạng Nơ Ron Hồi Quy Bậc Hai
Abstract—Wireless Sensor Networks are real time databases to real world phenomena. As wireless Sensor Networks (WSNs) generally rely on batteries for power, the nodes of the network have a limited	Tóm tắt-Các mạng cảm biến không dây là các cơ sở dữ liệu thời gian thực của các hiện tượng trong thế giới thực. Vì nói chung các mạng cảm biến không dây (các WSN) hoạt động nhờ vào pin, các nút mạng có một

operational lifetime. Efficient power consumption is of utmost importance in operation and maintenance of the network. This paper summarizes work in progress in efficient energy consumption during sensor data collecting through time series modeling. A previous approach for energy efficient model in WSNs using a linear time series model for measurement prediction is reviewed and a new model utilizing a non-linear machine learning approach is proposed.	thời gian hoạt động giới hạn. Do đó, tiêu thụ năng lượng có hiệu quả đóng vai trò vô cùng quan trọng trong hoạt động và bảo trì mạng. Bài báo này tóm tắt một công trình đang tiến hành để hỗ trợ tiêu thụ năng lượng hiệu quả trong quá trình thu thập dữ liệu cảm biến thông qua mô hình hóa chuỗi thời gian. Chúng tôi điểm lại các phương pháp trước đây để xây dựng mô hình hiệu quả về năng lượng trong các WSN dùng mô hình chuỗi thời gian tuyến tính để dự đoán phép đo và đề xuất mô hình mới sử dụng phương pháp học máy phi tuyến.
I. INTRODUCTION The rapid advancement in wireless communication technology has led to a significant interest in wireless sensor networks (WSNs). WSNs have emerged as a platform for surveillance, control, and measurement applications. As WSNs are dependent on battery power, energy efficient operation is of paramount importance to the longevity of the network.	[redacted]
In WSNs, sensor data need to be collected and transmitted to a small number of dedicated sinks (base stations), where the end-user can access the data [5] and each time mobile sensors send data to a sink, energy consumption is required. To reduce this energy consumption, the number of communication between sensor nodes and the sink has to be reduced because communication between nodes is the main source of energy consumption [11] as node communication tends to be the most expensive aspect of operation in wireless sensor networks [8].	[redacted]

A new approach to energy efficient sensing through sensor measurement prediction [6], [9], [12], [13] has shown great potential in communication reduction. In these approaches, sink node exploits time series model to predict local readings instead of direct communicating with sensors. Nonetheless, the past approaches have been limited in the sense that only linear autoregressive models have been considered for modeling non-linear phenomena. Naturally, these linear models are inappropriate for the task of estimating noisy, high dimensional, non-linear processes. This paper proposes a new framework incorporating a non-linear time series model for measurement estimation which reduces energy consumption by learning a mapping that fits the long term characteristics of the underlying process, thereby eliminating the need for frequent re-estimation of model parameters. II. TIME SERIES MODELING The underlying theme of time series modeling is to identify a mapping F from Rd to itself which takes the system from z(t) to z(t + 1),via optimization of parameters w = {w\,w2,...,wn} [1]. In time series modeling, the function F(z, w) is usually classified into two main groups: linear, in which F is known a priori, or non-linear, where F is usually unknown (non-parametric) and must be fit to the data. Time series models can be classified	[redacted] [redacted] [redacted] [redacted]

based on the a priori priors of the model, such as whether the functional form of the model is known or not. In this paper, parametric models refer to models whose functional form are pre-specified, and nonparametric models adapt the functional form from the data. For time series processing linear models are, by definition, parametric. A general parametric model has the form $zt = ft(xt, wt) + e(t)$ where f is a known mapping and the error $e(t)$ is normally distributed white noise. It is assumed that the weights of the model wi e w are initially unknown and need to be estimated. Nonparametric models have the form $zt = ft(xt, wt) + e(t)$, where f is an unknown function which must be estimated from the data via optimization of the weights wi e w. The error term $e(t)$ is again assumed to represent white noise.

III. LINEAR SENSOR ESTIMATION IN WSNS

Communication between nodes is the main source of energy consumption [11] in wireless sensor networks [8]. Time series forecasting was proposed as a means to reduce the amount of communication between the wireless sensor and the sink. In previous approaches [2], [12], [13], data from the sensor nodes were processed with a linear autoregressive model contained in both the sink and each sensor (as Figure 1 shows, the sink contains a copy of each sensors model). The linear autoregressive model is

[redacted]

[redacted]

[redacted]

described in the following section.

A. Linear Auto Regressive Models

Auto regressive models analyze autocorrelations in a time series. These models operate in a closed loop and are known as infinite impulse response filters. The pth order AR(p) auto regressive model regresses the most current value z(t) on to its previous x(t — 1),...,x(t — p) values

$$z(t) = -a_1x(t-1) - a_2x(t-2) - \dots - a_px(t-p) + \varepsilon(t) \quad (1)$$

where z(t) is the measurement at time t, a_i is the ith lagged autoregressive coefficient, and e(t) is the noise term which is assumed to be serially uncorrelated and normally distributed.

1) Autoregressive Moving Averages: By combining the above model with a moving average, we arrive at the autoregressive moving average (ARMA)

where the residuals of the moving average are represented by v(t), v(t — 1),..., v(t — q) and the coefficients are represented by β_1, β_2,..., β_q at times t,t — 1,...,t — q respectively. The notation ARMA(p, q), refers to the order of the model where p is the autoregressive order and q is the moving average order. The ARMA model has played a central role in energy savings through sensor data prediction in wireless sensor networks as explained in the following section.

B. Linear Sensor Estimation in WSNs

In previous models, each sensor contained is own ARMA model

[redacted]

[redacted]

[redacted]

[redacted]

which was trained from a history of sensor measurements as shown in Figure 1. A copy of each model was then sent to the sink to predict subsequent values into the future, thereby eliminating the need for data to be passed between the sensor and sink at each time step. As the ARMA model is linear and most real world phenomena are nonlinear, divergence of the predictive model from the underlying process is inevitable.	[redacted]
To correct this problem, the prediction error was monitored. If the prediction error became bigger than some pre-specified error tolerance, the sensor node would re-compute the parameters of the linear model, and then send the model parameters to the sink. Predictions would then begin again, and continue, until the error tolerance condition is violated.	[redacted]
The problem with this type of framework lies in the application of a linear model to a fundamentally nonlinear problem. As divergence of the time series model is inevitable, re-parametrization of the model is necessary (which incurs communication with the sink). The frequency of model re-parametrization is dependent on the non-linearity of the data and the acceptable error tolerance. In the following section a non-linear modeling approach is proposed to reduce the need for re-parametrization, which can lead to savings in energy.	[redacted]
IV. A NONLINEAR APPROACH	[redacted]

TO ENERGY SAVINGS IN WSNS

The biggest limitation of the previous time series approaches to energy consumption in WSNs is that only linear (ARMA) models were considered. As it is well known that real world phenomena are usually nonlinear, we argue that nonlinear modeling is better suited to the task.

In the proposed model, the sink carries out the training of the predictive model as opposed to the training being carried out at the sensor as in [12]. Here a series of nonlinear predictors form a series of mappings Fk : DT ^ xN +k, where k G [1,2,.. .,P] and where N e Z+. The mappings relate a measured set of data DT = {XH}(NN+T+I+N_2)P sent from the sensor to a series of points {zi}P=P+N+(p+i+T-2)Pin the future. xi represents the sensor measurement and DT is the set of {xi} data points sent from the sensor to the sink. The mappings Fk yield predictions of P steps into the future which are collected in the sink. The subscript T indexes time where at every PT — 1 time steps the sink contacts the sensor for an additional set of measured data DT+i which is fed back into the predictive model. The model is initialized at time step T = 0 where a pre-specified amount of training data w is sent from the sensor to the sink for function estimation, and predictions begin after T > 1. Here the functional mapping between successive data points in time is learned by the Elman network (see Figure 3) as discussed in the

[redacted]

following section.

V. RECURRENT NEURAL NETWORKS

Recurrent Neural Networks directly address the temporal relationship of their inputs by maintaining an internal state. in this study the Elman [4] network is considered.

Fig. 3. Proposed Model with Nonlinear Predictors.

A. RNN Architecture

The Elman network was originally used for speech processing by Robinson and Fallside [10]. This network has the following governing equations:

(3)

and

(4)

where ut(k) represents the state of the kth element in the hidden state vector, and f (■) being the hidden layer transfer function. The weights w(i,j) connect the ith element from the jth element where H represents the hidden layer and I represents the input layer. The output of the network zk passes the state vector ut+1(k) along with the bias w(k,b) through the output transfer function g(). All transfer functions here are sigmoid and xt(l) is the input vector.

B. Derivative Computation with RTRL

The real time recurrent learning algorithm (RTRL) computes the error gradients in weight space directly from the observation of the network state and the network output error at each time step. The quadratic cost function is defined by ed = \ 2fc=i en{k) = \ J2k=Adt(k)-ztik))2 where dt{k) is the target and

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

zt(k) is the network output. The cost function is minimized, by computing the partial derivative of the error function as follows

C. Second Order Training

Second order Newton based methods are attractive extensions to gradient descent learning because they provide a quadratic rate of convergence and the second order information can be used for model selection. Underlying Newton based methods is the local approximation of the cost function by a quadratic form given by E(wt + dwt) = E(v?t) + SE(vft)Tdwt + TjdvtfV2E(vft)dvft where VE(vtt) and V2E(wt) are the Gradient vector and the Hessian matrix of the cost function, respectively. The Newton step is then obtained by dwt = -[V2E(wt)]_1VE(wt) Since the cost function follows the sum of squares form, the Hessian matrix can be approximated as V2E(wt) = JTJt + St where Jt is the Jacobian matrix (first derivatives) of the cost function and St represents the second order derivative information in V2E(wt). If the term St in the above expression for the Hessian is assumed to be zero, the resulting equation becomes the Gauss Newton method.

From Figure 2, each Fi represents a neural model defined in section V which maps an input xt to an output zi, where the mappings range between one step to P steps into the future which saves P steps of communication between the sink and the sensor.

VI. CONCLUSION

The total amount of energy

[redacted]

[redacted]

[redacted]

consumption of both linear and nonlinear models at given different number of data points, error tolerance and intermediate nodes between sensor and sink will be compared through extensive simulations in further research. This paper reports on work in progress in the area of wireless sensor networks using machine learning for the temporal prediction of data. Previous attempts in forecasting sensor readings using linear models have led to a significant reduction in energy usage of the sensors. The model proposed in this paper integrates cutting edge predictive models from machine learning to enhance data quality and reduce energy savings in wireless sensor networks over the previous linear approaches.	[redacted] [redacted]