A Novel Data Collection Strategy Based on Fourier Transform in Power System WSN

Li Wu[1,a], Jiangwen Xiao[1], Jing Wu[2], Bin Liu[1], Jun Li[1], Yiwen Xiao[1]

[1]Huangzhong University of Science and Technology in Wuhan, Hubei Prov, China

[2]Jiangnan University in Wuxi, Jiangsu Prov, China

[a]winnyli@whu.edu.cn

Keywords:data collection; Fourier transform; wireless sensors network.

Abstract: Temperature detection is now used for risk early warning in power system. Adopting wireless sensor network to collect temperature information is one of the most crucial applications. In order to accomplish temperature information gathering, hundreds of sensor nodes are placed in power system equipment, such as transformer, relay protection device, mutual inductor, to name a few. Studies have shown that, for nodes, message transmission consuming more energy than local data computing. It's unrealistic to configure batteries for sensor nodes or replace them on a large scale in power system. For the purpose of lengthening the lifetime of wireless sensor network and reducing energy consumption, the amount of message transmission should be minimized. In this paper, a new data collection strategy in power system based on Fourier transform is proposed. Because of periodic character of temperature detection data, Fourier transform is adopted to linearly fit these data in this paper, and least square strategy is used to optimize parameters. The sink nodes transmit model parameters to the base station instead of thousands of data, when the error is in the range of confidence level. Experiments show that the proposed strategy obtains greater fitness and less error between the measured data and predicted data.

Introduction

Power system troubleshooting is an important task for power grid companies. But artificial maintenance is inefficient and inaccurate, and normally, more expensive. When the fire hazards exist in the power system, it's often accompanied by radiation and local temperature rising. Therefore by detecting temperature of a large number of key points for fire warning is of great significance. This is the places where wireless sensor network(WSN)[1,2] plays a role. Hundreds of temperature sensor nodes are placed at the key points in power system equipment to detect the potential risk.

WSN will have profound impact on people's live. Multipoint temperature detection based on WSN is widely used in industrial and military, like forest fire warning, medical care, coal mine protection. WSN is a new kind of sensor network, which is large in scale, self-organization, dynamic, and data-centric. A typical WSN consists of several parts, sensor nodes, sink nodes, gate way nodes and task management nodes[3]. Sensor nodes are distributed in the network, and is responsible for temperature data perception and gathering. Data are transmitted by sensor nodes to sink nodes, where data is preliminary processed.

For sensor network, message transmission consumes much more energy than computing. Obviously, replacing battery artificially costs more money. Power supply, which is the life of sensors, is an urgent restriction for them. So it is necessary to reduce message transmission. For temperature

detection in power, long sampling period will lose foreseeability, which is what we want to obtain from WSN on the contrast. As a consequence, all temperature data are transmitted to the user is not realistic, no matter from the perspective of energy restriction or data redundancy.

Many strategies fitting data to reduce the dimension of data have been discussed before[4,5,6]. In this paper, a new strategy to fit temperature data is proposed, thus several parameters, instead of a large amount of data, will be sent to the user. When the error between estimated data and measured data is in the range of confidence degree, the user will not receive measured data, and the state of power system is safe. When the error is beyond of range, measured data will be sent to the user for fire hazards prediction and alert.

Temperature change within a day of presents a periodic characteristics. Different from synthesis high order polynomial fitting in the past, we use Fourier transform to fit data, which is proved to have better effect on periodic data. In our simulation, this kind of method shows better fitness than high order polynomial fitting.

The rest of this paper is organized as follows. In section 2, Fourier strategy and algorithm are presented to state our objective. Simulation results are presented in section 3. Finally, a brief summary is made and future work is discussed in section 4.

Algorithms on Data Collection Strategy

The monitoring system aims to extract data effectively without losing its general features[7,8]. By using Fourier transform fitting strategy[9], redundant data is reduced, with no negative effect on decision analysis of the system.

Considering the application scenarios of WSN, the sensor nodes storage and processing performance, as well as the cycle of temperature changes outdoor, m latest sensory data having a certain time interval is chosen in this paper. Assume that the time series and measured temperature data form a set of vectors, denoted as $(t_1, y_1), (t_2, y_2), (t_3, y_3), \cdots, (t_m, y_m)$, where $t_i, i = 1, 2, \cdots, m$ is sample time point, and y_i is the measured temperature with error.

Construct a function $Y(t)$ for these m sensory data, which satisfies the approximation error $\delta_i = Y(t_i) - y_i$ is in the range of confidence level. The form of function $Y(t)$ depends on specific issues. Consume the function form as below:

$$Y(t) = \lambda_0 + \sum_{j=1}^{n} \lambda_j B_j(t) \tag{1}$$

Where λ_0 is the average of measured data. $\lambda_0 = \left(\sum_{i=0}^{m} y_m\right)/m$. For the temperature measurement WSN, it is the average temperature of a day. n is the number of term, while B_j is the primary function. It's worth noting that n and the form of B_j depends on the actual problem.

Under normal circumstances, measured temperature is influenced by the outdoor temperature in WSN of power system. And outdoor temperature presents to be periodic within a day in general. So it is reasonable to take the primary function B_j as periodic function. In this paper, we take it as:

$$B_j(t) = \sin\left(\omega_j t + \theta_j\right) \tag{2}$$

where $\theta_j \in (-\frac{\pi}{2}, \frac{\pi}{2})$. Thus $Y(t)$ can be denoted as:

$$Y(t) = \lambda_0 + \sum_{j=1}^{n} \lambda_j \sin\left(\omega_j t + \theta_j\right) \tag{3}$$

If $n \approx m$, we need to compute n parameters to replace m. As a consequence, it isn't up to the aim of reducing information transmission amount and data redundancy. What's more, the accuracy of prediction will be affected. We choose $n \ll m$ in general. Then, by selecting appropriate λ_j, we gain the function Y corresponding to the measured data y.

Here we select three most important component harmonic to form the model. This is to say, $n = 3$. f is a vector which consists of three frequency corresponding to these harmonics. $\omega = (\omega_1, \omega_2, \omega_3)^T$ and $\theta = (\theta_1, \theta_2, \theta_3)^T$ are 3×1 vectors, by solving which we can obtain $B(t)$. When WSN transmit data, the sensor nodes can only send parameters $\lambda = (\lambda_0, \lambda_1, \lambda_2, \lambda_3)^T$, ω, θ to the monitoring center instead of vast measured data. For ω, it is easy to know that:

$$\omega = \frac{2\pi f}{m} \tag{4}$$

From Eq. 2 and Eq. 3. we can rewrite $Y(t)$ as :

$$Y(t) - \lambda_0 = \sum_{j=1}^{3} \lambda_j (\sin \omega_j t \cos \theta_j + \sin \theta_j t \cos \omega_j t) \tag{5}$$

Let $Y_o(t) = Y(t) - \lambda_0$, $Y_o(t)$ is the difference between estimated data and average measured data. Thus we can set the function near X axis. For m sample data, matric $Y_o(t)$ can be written as:

$$Y_o = \begin{bmatrix} Y_o(t_1) \\ Y_o(t_2) \\ Y_o(t_3) \\ \vdots \\ Y_o(t_m) \end{bmatrix} = \begin{bmatrix} \sin\omega_1 t_1 & \cos\omega_1 t_1 & \cdots & \sin\omega_n t_1 & \cos\omega_n t_1 \\ \sin\omega_1 t_2 & \cos\omega_1 t_2 & \cdots & \sin\omega_n t_2 & \cos\omega_n t_2 \\ & & \vdots & & \vdots \\ \sin\omega_1 t_m & \cos\omega_1 t_m & \cdots & \sin\omega_n t_m & \cos\omega_n t_m \end{bmatrix} * \begin{bmatrix} \lambda_1 \cos\theta_1 \\ \lambda_1 \sin\theta_1 \\ \vdots \\ \lambda_n \cos\theta_n \\ \lambda_n \sin\theta_n \end{bmatrix} \tag{6}$$

Let $x = [\lambda_1 \cos\theta_1, \lambda_1 \sin\theta_1, \cdots, \lambda_n \cos\theta_n, \lambda_n \sin\theta_n]^T$ \hfill (7)

x is what we need to solve later. Construct a matric M, and write it as:

$$M = \begin{bmatrix} \sin\omega_1 t_1 & \cos\omega_1 t_1 & \cdots & \sin\omega_n t_1 & \cos\omega_n t_1 \\ \sin\omega_1 t_2 & \cos\omega_1 t_2 & \cdots & \sin\omega_n t_2 & \cos\omega_n t_2 \\ & \vdots & & & \vdots \\ \sin\omega_1 t_m & \cos\omega_1 t_m & \cdots & \sin\omega_n t_m & \cos\omega_n t_m \end{bmatrix} \tag{8}$$

Then Eq. 6 can be shorted as:

$$Y_o = Mx \tag{9}$$

In order to get x, least square method[10,11,12] can be used. Approximate error δ can be denoted as :

$$\delta = Mx + \lambda_0 - y. \tag{10}$$

Here vector λ_0 consists of m elements, of which each element equals to the average value λ_0. In order to minimize the approximate error δ, we need to make its vector 2-norm as small as possible, like this:

$$min \left(\|\delta\|^2 = \left(\sum_i^m \delta_i^2 \right)^{1/2} \right) \tag{11}$$

Substitute Eq. 10 into Eq. 11, then a new formula is obtained:

$$min \|\delta\|^2 - min(\|Mx + \lambda_0 - y\|^2) - min\left(\sum_{i=1}^{m} \left(\sum_{j=1}^{n} m_{ij} x_j + \lambda_0 - y_i \right)^2 \right) \tag{12}$$

Now in order to get the minimal δ, we can derivate $\|\delta\|^2$ by each $x_k, k = 1, \dots, n$, thus we can minimize the value of 2-norm. Derivative equation is:

$$\frac{d\|\delta\|^2}{dx_k} = \sum_{i=1}^{m} 2 \left(\sum_{j=1}^{n} m_{ij} x_j + \lambda_0 - y_i \right) m_{ik} = 0, k = 1, 2, \dots, n \tag{13}$$

Referring to Eq. 13, the following matrix equation is deduced:

$$(Mx + \lambda_0 - y)^T M = 0 \tag{14}$$

Eq. 14 can be deformed like this: $M^T(Mx + \lambda_0 - y) = 0$. After further deformation, it can be written as:

$$M^T M x = M^T(y - \lambda_0) \tag{15}$$

From the knowledge of matrix theory, it is sure that Eq. 15 has a solution. M is a column nonsingular matrix due to $\omega_i \neq \omega_j$, where $i \neq j, and\ i, j = 1, ..., n$. As a result, $M^T M$ is positive definite. Further to say, the inverse of it exists. Thus solution x is obtained:

$$x = (M^T M)^{-1} M^T(y - \lambda_0) \tag{16}$$

To simplify Eq. 16, we import matrix A and vector z. Let $A = M^T M$, and $z = M^T(y - \lambda_0)$. Then we get $x = A^{-1} z$, A is the dot product matrix of primary function matrix M, z is the projection of measured temperature data subtracting the average on primary function matrix. Since defined before that $\theta_j \in (-\frac{\pi}{2}, \frac{\pi}{2})$, we know that $\lambda_i, i \in 1, 2, ..., n$ is determined and unique. From $x = [\lambda_1 \cos\theta_1, \lambda_1 \sin\theta_1, \lambda_2 \cos\theta_2, \lambda_2 \sin\theta_2, \lambda_3 \cos\theta_3, \lambda_3 \sin\theta_3]^T$ and λ_0, we can obtain the unknown vector λ and θ.

Simulation results and comparisons made with other algorithms

In this part, we will present our simulation results on the proposed algorithm, and comparisons are made with other collection strategies. We did experiment for 9 times on different data. 23 sets of data are picked in each experiment, as shown in Table 1, which is roughly periodic with highly irregular noise.

Table 1 Experiment Data with Periodic Characteristic

t_i	1	2	3	4	5	6	7	8
y_i	1.2240	0.4626	0.0752	0.0204	0.0006	0.0009	0.0017	0.0039
t_i	9	10	11	12	13	14	15	16
y_i	0.0096	0.3900	0.9394	1.3015	1.4905	1.2535	0.8587	0.4966
t_i	17	18	19	20	21	22	23	
y_i	0.2286	0.1173	0.0694	0.1181	0.2336	0.5034	0.9571	

Take these data for example to explain the construction of regression model and the process to solve parameters. For the sake of simplicity, let $n = 3$. Figure. 1 shows its plane distribution. Fitting line is also drawn it Figure. 1. From the experiment result, we can see that fitting model based on Fourier transform shows the best effect. Appropriate error δ is small enough. As for high order fitting model, δ is much larger. And for difference fitting model, fitting effect turns to be the worst. Prediction effect is also the worst.
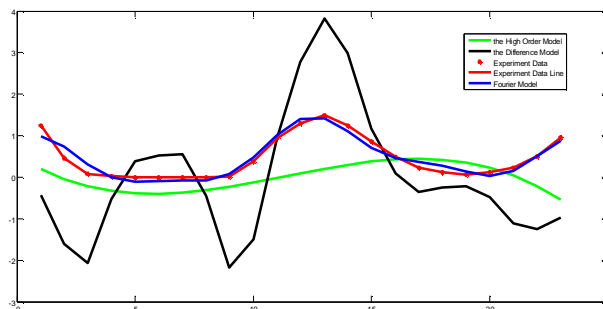

Figure 1 Experiment Result of Data with Periodic Characteristic

In order to measure the fitting effect, we derive a definition of fitness from the cost function given by:

$$\text{fitness} = 100 * \left(1 - \frac{\sum_{k=1}^{m}(y_k - \hat{y}_k)^2}{\sum_{k=1}^{m}(y_k - \bar{y})^2}\right) \tag{17}$$

where $\bar{y}$ is the average value of experiment data, and $\hat{y}_k$ is the estimated data. It is easy to know that a zero error model corresponds to a 100% fit. The less the error, the higher the fitness is. Compute fitness of such three models, in allusion to periodic experiment data, the fitness of Fourier strategy is up to 93.0285. As for high order model, fitness is less than 10%. For difference model, it turns out to be negative.

Another set of experiment data is much less periodic, as shown in Table 2. But the proposed strategy is also show much better performance in data tracking and prediction ability.

Table 2 Experiment Data with Less Periodic Characteristic

t_i	1	2	3	4	5	6	7	8
y_i	0.0946	0.1290	0.1605	0.1832	0.2152	0.4406	0.6417	0.6343
t_i	9	10	11	12	13	14	15	16
y_i	0.6268	0.6665	0.7062	0.6600	0.6012	0.5780	0.5676	0.6807
t_i	17	18	19	20	21	22	23	
y_i	0.8187	0.8625	0.8812	0.7932	0.7053	0.7111	0.7521	

For this set of data, fitness of Fourier transform strategy is 89, while it for difference model is 73，and 83 for high order model. This result indicates that for data less periodic, there is obvious progress in performance of difference model and high order model. But Fourier strategy is still better than the two.
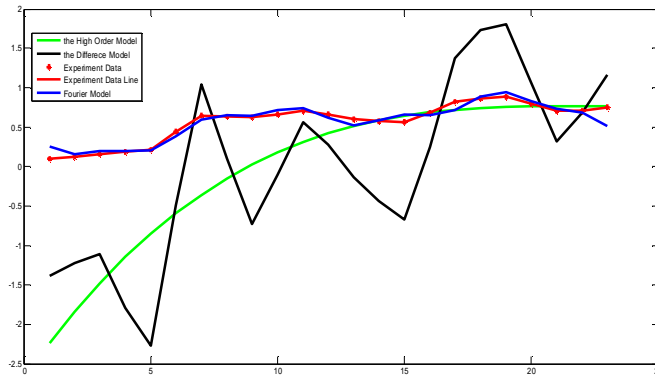

Figure 2 Experiment Result of Data with Less Periodic Characteristic

Conclusion and Future Work

Simulation results show that adopting Fourier model to fit sample data can notably get high fitness and predict ability, and less error. By using Fourier model, what transmitted in WSN is not high-dimensional temperature data, but ten parameters in $\lambda = (\lambda_0, \lambda_1, \lambda_2, \lambda_3)^T$ $(\omega_1, \omega_2, \omega_3)^T$ and $\theta = (\theta_1, \theta_2, \theta_3)^T$. This Fourier model reduce redundancy, and lengthen lifetime of nodes. Though compared to high order model, Fourier model need more computation and parameters, it shows much better performance in data with periodic characteristic.

In the future, we would like to expand our work in other cases. We want to discuss the temperature fitting with weights due to different sampling nodes have different impact on prediction results.

References

[1] D.Cullar, D.Estrin, M.Strvastava, Overview of sensor network. Computer, 2004, 37(8), 41-49

[2] I.Akyliidiz, W.Su, Y.Sanakarasubramnaiam, et al..Wireless sensor networks: A Survey. Computer Networks, 2002,38(4):393-422

[3] Fengyuan Ren, Haining Huang, Chuang Lin, Wireless sensor networks, Journal of Software, vol.14, No.7, pp.1282-1291, 2003

[4] Jianming Zhang, Yingqing Song, Siwang Zhou, Jingcheng Ouyang, Survey on data aggregation techniques in wireless sensor netwoks, Computer Applications, 2007,vol.20,p.1875-1880

[5] J.Gao, L.Guibas, J.Hershberger, Sparse data aggregation in sensor Cambridge, Massachusetts,USA

[6] Lintao Lv, Juanyun Duan, Cui Li, Data gathering model oriented wireless sensor network, Computer Engineering, 2008, vol.34,pp.111-113.

[7] Venuto D D, Rabaey J. RFID transceiver for wireless powering brain implanted microelectrodes and backscattered neural data collection[J]. Microelectronics Journal, 2014, 45(12):1585–1594.

[8] Stapleton S J, Degitz R J. An Innovative Data Collection Method for Investigating Unresolved Pain After ED Discharge: A Pilot Study[J]. Journal of Emergency Nursing Jen Official Publication of the Emergency Department Nurses Association, 2014, 40(6):598–604.

[9] Zhao R, Li X, Sun P. An improved windowed Fourier transform filter algorithm[J]. Optics & Laser Technology, 2015, 74:103–107.

[10]   Fengxia Zhang, Musheng Wei, Ying Li, Jianli Zhao, Special least squares solutions of the quaternion matrix equation $AX = B$ with applications, Applied Mathematics and Computation, Volume 270,1 November 2015,Pages 425-433

[11]   Aràndiga F, Noguera J J. Reconstructions that combine interpolation with least squares fitting[J]. Applied Numerical Mathematics, 2015, 97:30–41.

[12]   Singh S K, Sinha N, Goswami A K, et al. Variable Constraint based Least Mean Square algorithm for power system harmonic parameter estimation[J]. International Journal of Electrical Power & Energy Systems, 2015, 73:218–228.