Online Temperature Estimation of Li-ion Battery Pack Using Principal Component Analysis

Taejin Kim[1], Sunuwe Kim[2], and Byeng D. Youn[3]

[1,2,3]*Department of Mechanical and Aerospace Engineering, Seoul National University, Seoul, 08826, Republic of Korea*

godori16@snu.ac.kr
lunashisun@gmail.com
bdyoun@snu.ac.kr

ABSTRACT

Thermal management is one of the important function of the battery management system (BMS). The thermal management system monitors and equalizes the temperature distribution of the battery pack to prevent the different cell degradation rate and to keep the battery on its best performance. In this study, as a part of the thermal management system the in-situ temperature estimation method is developed based on the principal component analysis (PCA) reinforced with the measured temperatures. To begin with, the PCA is used for finding the basis vectors of the battery thermal system, which is the eigenvectors of the covariance matrix of the training data set. Then an arbitrary thermal map can be expressed as the linear combination of these basis vectors and their amplitudes. The amplitude for each basis vectors is estimated from the measured temperatures. The performance of the thermal map reconstruction depends on the accuracy of this amplitude estimation which again is related to the temperature measurement locations. The measured locations are determined considering two aspects: the prediction accuracy and the robustness of the sensor network. To find the sensor location satisfying both criteria, the sensor network optimization problem is accordingly formulated, and solved by the genetic algorithm. The proposed study is validated for various operating conditions including the distributed heat generation condition and different cooling conditions.

1. BATTERY PACK MODELING

In this section, the battery pack thermal model is introduced, which will be later used to generate the training data for extraction of the principal components.

1.1. Battery Pack Overview

A thermal simulation model for a battery pack is developed using the lumped parameter. The battery pack is composed of 50 numbers 18650 cylindrical cells (ICR18650B4, LG Chem.) with 2600mAh capacity, and 3.7V nominal voltage. They are arranged to have 5 by 10 configuration. To construct the finite element model, each cell is represented as a node and it is connected with the surrounding cells by thermal resistance. At each cell, heat is generated by electrochemical reactions. The heat is transferred to the surrounding cells by conduction and to the air by convection. The energy balance equation for the i^{th} cell in the battery pack is given as (Forgez,, Do, Friedrich, Morcrette, & Delacourt, 2010):

$$C_h \frac{dT_i}{dt} = R_t \sum_j (T_i - T_j) + \dot{Q} + h(T_i - T_{amb,i}) \quad (1)$$

where C_h is the heat capacity of the cell, T_i is the temperature at the i^{th} cell, and T_j is the cell around the i^{th} cell, $T_{amb,i}$ is the ambient temperature, t is the time, R_t is the thermal resistance between cells, $\dot{Q}$ is the heat generation rate, and h is the heat transfer coefficient. To simulate the model, the parameters including the heat generation rate are identified by the characteristic tests, described in the following section.

1.2. Heat Generation Model

During charging and discharging, heat is generated inside the cell. There are several heat sources, but in this study, two dominant heat sources, irreversible ohmic heating and reversible entropic heating are considered (Forgez, et al., 2010). The heat generation rate is given as

$$\dot{Q} = I^2 R + IT \frac{dV_{OCV}}{dT} \quad (2)$$

where I is the input current, R is the impedance of the cell, and V_{OCV} is the open circuit voltage. The first term is the ohmic heating and the second term is the entropic heat. Since the impedance R is dependent on the state-of-charge (SOC), the hybrid pulse power characterization (HPPC) test is performed to obtain the SOC and impedance relationship (Onda, Kameyama, Hanamoto, Ito, 2003). Also, dV_{OCV}/dT for entropic heat is obtained as a function of SOC according to Onda et al. (2003). Finally, for calculation of the SOC during operation, the coulomb counting method given in Eq. (3) is used.

$$z_{k+1} = z_k - \left(\frac{\eta_i \Delta t}{C_n}\right) I_k \tag{3}$$

In the above equation, k is the time index, z_k is the SOC at time index k, η_i is the Coulombic efficiency, Δt is the time difference between k+1 and k time indices, C_n is the nominal capacity of the cell, and I_k is the input current. With the impedance, dV_{OCV}/dT, and SOC as a lookup table, the heat generation can be calculated.

The model parameters of a single cell, the heat capacity and the heat transfer coefficient, are calibrated and validated with experimental results. First, for model calibration, the temperature under 1C discharging condition is obtained and used for target vector to optimize the model parameters. The optimized heat capacity and heat transfer coefficient are 69.43 J/K, and 0.1555 W/m^{2}K, respectively. The model is then validated with the urban dynamometer driving schedule (UDDS) current profile, which emulate the real driving schedule. The UDDS profile and 10% SOC discharging is performed alternately until the SOC is near zero, then the charging process follows. The measured and the simulated temperatures for this current profile are shown in Figure 1. The simulated temperature result shows good agreement with the measured temperature during the discharge. The temperature at charging deviates a bit from the measured one, because the impedance used is obtained from discharging pulse. For better results during charging, the impedance needs to be calculated using the charging impulse.

2. SENSOR NETWORK DESIGN

2.1. Principal Component Analysis (PCA) and Signal Reconstruction

The sole use of thermal model for estimating the temperature distribution of the battery pack becomes unreliable as time passes due to the error accumulation, which comes from the uncertainty of the model and the state change of the battery. Hence, the estimated thermal map needs to be corrected by the measured data, and this correction depends on the sensor locations or the sensor network design. The basic idea of the sensor network design is to locate sensors that maximize the total amount of information. In other words, the sensor locations should be selected to assure accurate reconstruction of the temperature distribution. To this end, the temperature reconstruction from the arbitrarily measured signal is first introduced, and the criteria to determine the sensor locations is explained next.

The reconstruction of temperature distribution starts from the decomposition of the signal. The principal component analysis can be used to decompose the signal as (Kammer, 1991)

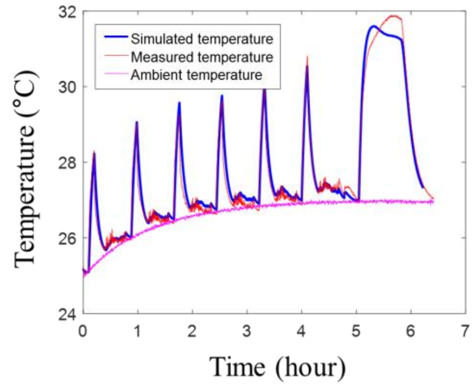


Figure 1. UDDS test results: the measured and the simulated temperatures.

$$\mathbf{y} = \mathbf{\Phi q} \tag{4}$$

where y is an arbitrary signal, Φ is the matrix whose columns are the principal components of the training data set, and q is the corresponding weight. The principal component is a vector that represents the principal axis of the covariance matrix and is obtained as the eigenvector of the covariance matrix. Since the eigenvectors are in common, it is the weight that determines the specific signal. For this reason, now the weight is estimated from the measured signal. According to the method of the best linear unbiased estimator (Kammer, 1991), the weight is estimated as

$$\hat{\mathbf{q}} = \left(\mathbf{\Phi}_s^{\mathrm{T}} \mathbf{\Phi}_s\right)^{-1} \mathbf{\Phi}_s^{\mathrm{T}} \mathbf{y}_s \tag{5}$$

where y_s is a p×1 vector, p is the number of sensors, Φ_s is the p×p principal component matrix, where its row is reduced according to the measured location, and $\hat{q}$ is the estimated weight. Finally, the signal reconstruction is given as (Kammer, 1991)

$$\hat{\mathbf{y}} = \mathbf{\Phi}_r \hat{\mathbf{q}} \tag{6}$$

where Φ_r is n×p principal component matrix, n is the number of samples, and $\hat{y}$ is the reconstructed signal.

The signal reconstruction so far is discussed in terms of the arbitrarily measured signal. Now selecting y_s (p×1) out of y (n×1), or measurement location is discussed. As seen in the above, the reconstruction of the signal depends on the estimation accuracy of the weight, which can be measured as the variance P of the estimation (Kammer, 1991).

$$\mathbf{P} = E[(\mathbf{q}-\hat{\mathbf{q}})(\mathbf{q}-\hat{\mathbf{q}})^T] = \left[\frac{1}{\sigma^2}\mathbf{\Phi}^T\mathbf{\Phi}\right]^{-1} = \mathbf{Q}_0^{-1} \tag{7}$$

From the equation, the variance between the target weight and the estimated weight is expressed in the quadratic form of the principal component matrix divided by the measurement noise variance σ^2. The inverse of this quadratic equation is also known as the Fisher information matrix, $\mathbf{Q}_0$.

Therefore, the $\boldsymbol{y}_s$ out of $\boldsymbol{y}$ should be selected to give the low variance of weight estimation, or equivalently high value of Fisher information matrix. To measure the Fisher information matrix, its determinant is used (Kammer, 1991). In the next section, the sensor network design with the Fisher information matrix as the objective function is discussed.

2.2. Robust Optimal Sensor Network Design

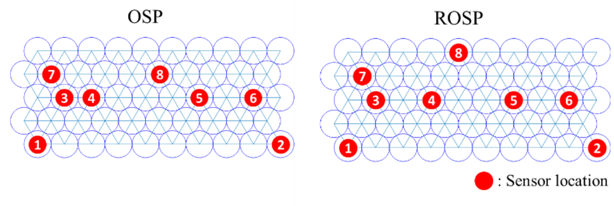
The object of selecting sensor locations is to maximize the determinant of the Fisher information matrix with the ($p\times p$) matrix, Φ_s,which is the matrix that (n-p) rows and columns are eliminated out of the ($n\times n$) principal component matrix, Φ. Therefore, the selections must be made for both rows and columns. The rows are the candidate locations of sensor placement and will be decided by the genetic algorithm later. The columns are the mode shape vectors, and will be selected as the principal components corresponding to the p highest eigenvalues of Φ. The choice of these principal components is adequate to represent the general behavior of the system.

The choice of rows, or sensor locations, is considered in two aspects. First, the determinant of the Fisher information matrix needs to be high enough to reconstruct the whole thermal map out of the measured signal. Second, for the sensor network to be robust, the sensor network needs to keep its functionality under latent sensor failure. Then, the objective function and the constraints are formulated as shown in Eq. (8):

$$\begin{aligned}&\text{Maximize: } f(i_1,...,i_p) = \mathrm{E}[\det(\boldsymbol{\Phi}[\{i_1,...,i_p\}\backslash\{i_j\};1,...,p-1]^{\mathrm{T}}\boldsymbol{\Phi}[\{i_1,...,i_p\}\backslash\{i_j\};1,...,p-1])]\\&\text{Subject to: } \det(\boldsymbol{\Phi}[i_1,...,i_p;1,...,p]^{\mathrm{T}}\boldsymbol{\Phi}[i_1,...,i_p;1,...,p]) \ge \alpha\cdot\det(\boldsymbol{\Phi}[i_1^{opt},...,i_p^{opt};1,...,p]^{\mathrm{T}}\boldsymbol{\Phi}[i_1^{opt},...,i_p^{opt};1,...,p])\end{aligned} \quad (8)$$

where $\Phi[i_1,\ldots,i_p;j_1,\ldots,j_p]$ is the submatrix of Φ formed from the rows $\{i_1,\ldots,i_p\}$and columns $\{j_1,\ldots,j_p\}$ such that $i_l\in\mathbb{N}$, $1\le i_l\le$n, and $i_l\ne i_m$. As mentioned before, the principal components are preselected as the first p eigenvectors of the covariance. The random variable for calculating the expectation in the objective function is i_j, which is an element in $\{i_1,\ldots,i_p\}$. Therefore, the objective function indicates the Fisher information matrix of the system when a sensor fails. The constraint is set to satisfy a certain performances. The constraint criterion is determined based on the sensor placement result without considering the failure. The i_{opt} is the optimal sensor location when failure is not considered. The coefficient α determines the performance criterion.

Since the sensor placement is the combinatorial problem that selects a given number p of sensors out of n candidate locations, the integer valued optimization method is required. Integer valued optimization is a non-convex problem that cannot be solved by gradient-based optimization. Thus, in this study, the genetic algorithm is adopted to find the sensor location, giving the best Fisher information matrix value of the system. The genetic algorithm is a sampling-based optimization algorithm inspired by biological evolution. It generates initial random samples (or population) and pass them to the next generation with modification (Vose, 1999).



(a) (b)

Figure 2. The sensor locations: (a) Optimal sensor placement (OSP), and (b) the robust optimal sensor placement (ROSP).

As the generation goes on, the optimal evolution is found as in natural selection. The sensor locations found by the genetic algorithm are shown with the sensor network design that does not consider the sensor failure as shown in Figure 2.

3. CASE STUDY

The estimation accuracy with and without considering failure of sensors are compared using a test data set. The test case is designed to imitate a battery pack with a cooling system. The battery pack is cooled down by forced convection that comes from the left side of the battery pack by a set different heat transfer coefficient for each cell. The simulation results are shown in Figure 3. As expected, the cells near the air inlet and the circumference of the pack have a lower temperature than the cells at the center.

The summary of the estimation error measured by the root mean square (RMS) error are shown in Table 1. The failed sensor number indicates the sensor indices shown in Figure 2. Both optimal sensor placement (OSP), and the robust-OSP (ROSP), which represent the sensor networks without and with consideration of sensor malfunction gave accurate estimation results in normal conditions; however, the accuracy drops with a sensor malfunction. Also, the accuracy drop for the ROSP results is less than the accuracy drop for OSP.

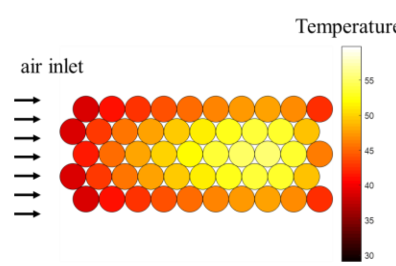


Figure 3. The temperature distribution of a battery pack under forced convection at 83 min.

Table 1 The RMS error of OSP and ROSP methods

	Estimation error (RMS)	
Failed sensor No.	OSP	ROSP
Normal	0.2658	0.3369
1	3.0945	1.3521
2	5.3976	1.5131
3	20.213	8.2920
4	1.9524	1.7199
5	2.3377	2.0404
6	5.5522	2.8814
7	1.8372	1.5646
8	2.1750	1.1891
Mean error for failure	5.32	2.57

From the results, two advantages are observed. First, by combining the model and the measured data, the proposed method captures both the computational efficiency and the estimation accuracy. In practice, such as in the case of a battery management system (BMS) in an electric vehicle, computational efficiency is an important issue. Thus, a complicated model requiring heavy computation cannot be used in spite of its accuracy in estimation. In contrast, the simple model is free from the computational burden; however, it is less accurate. The proposed method requires only a small memory to save the principal component matrix and the capability to calculate the matrix product; inaccuracy is compensated by the measured data and thus, it is suitable for practical use. The use of sensors only to obtain the temperature distribution of the battery pack is not practical because of the cost and the potential for sensor malfunctioning.

Another benefit of the proposed method lies in the sustainability of the sensor network under failure. The sensor network contributes to protect the system as a part of BMS; however, the sensor network itself is vulnerable to failure. Nevertheless, developing another protective system for the sensor network is not a good solution because it requires redundant cost. In this situation, the proposed sensor network design could be a solution because it does not require an extra system, while it does maintain a certain degree of performance under failure. Unlike the PHM, which tries to prevent failure, the proposed design allows the failure and gains time to take care of the failure.

4. CONCLUSION

In this study, we proposed a robust sensor network design for online temperature estimation for the battery pack, based on the principal component analysis. We defined the objective function and the constraint for this purpose. The optimization problem was solved by the genetic algorithm. The selected sensor locations were used to reconstruct the temperature distribution of various cases. The estimation results were compared with the other sensor network design that focuses on accuracy. The results show that the proposed method has estimation capability that is comparable to the existing optimal sensor design and has higher reliability.

In this study, the sensor locations were found through the optimization method, but it requires significant computational resources to find the optimal solution. Hence, to reduce the computational cost, an analytic solution is required. This will be addressed in future research.

Reference

Forgez, C., Do D. V., Friedrich, G., Morcrette, M., & Delacourt, C. (2010). Thermal modeling of a cylindrical $LiFePO_4$/graphite lithium-ion battery. *Journal of Power Sources*, vol. 195, pp. 2961-2968

K. Onda, H. Kameyama, Hanamoto, T., & Ito K. (2003). Experimental study on heat generation behavior of small lithium-ion secondary batteries. *Journal of the Electrochemical Society,* vol. 150, pp. A285-A291.

D. C. Kammer (1991). Sensor placement for on-orbit modal identification and correlation of large space structures. *Journal of Guidance, Control, and Dynamics,* vol. 14, pp. 251-259.

M. D. Vose, (1999). *The simple genetic algorithm: foundations and theory*. MIT press.

BIOGRAPHIES

Taejin Kim received his B.S. degree from Seoul National University, Seoul, Republic of Korea, in 2011. He is a Ph.D. student in Seoul National University. His research topic is the diagnosis and prognosis of the energy devices. He received two awards including the best paper in the IEEE Prognostics and Health Management Conference (2012) and the IEEE PHM Data Challenge Competition Winner (2014).

Sunuwe Kim B.S. degree of mechanical engineering from Korea University in 2014, and is involved in graduate School of Mechanical and Aerospace Engineering at SNU. He is currently doing research on Prognostics and Health Management for lithium ion battery.

Byeng D. Youn received the B.S. degree from Inha University, Incheon, South Korea, in 1996, the M.S. degree from KAIST, Daejeon, Republic of Korea, in 1998, and the Ph.D. degree from the University of Iowa, Iowa City, IA, USA, in 2001. He is an Associate Professor of mechanical and aerospace engineering at Seoul National University (SNU), Seoul, Republic of Korea. Dr. Youn's dedication and efforts in research have garnered substantive peer recognition resulting in notable awards including the ASME IDETC Best Paper Awards (2001 and 2008), the ISSMO/Springer Prize for a Young Scientist (2005), the IEEE PHM Competition Winner (2014), the PHM society Data Challenge Competition Winner (2014, 2015), etc.