Accurate Estimation of Battery SOH and RUL Based on a Progressive LSTM with a Time Compensated Entropy Index

Taejun Bak[1] and Sukhan Lee[2]

[1,2]*Sungkyunkwan University, Suwon, Gyeonggi-do, 16419, Republic of Korea*

ptjvv@skku.edu
lsh1@skku.edu

ABSTRACT

The accurate estimation of the State of Health (SOH) and Remaining Useful Life (RUL) has been a subject of keen interest due to its impact on safety and condition-based maintenance services. A number of approaches have been proposed to tackle this problem based either on a model-driven or on a data-driven framework. Due to the electro-chemical complexity involved in battery aging, they are yet to achieve the accuracy required, especially, for real-world applications. This is because of the difficulty either in identifying the time-varying nature of model parameters and in collecting the real-world training dataset from widely varying modes of battery usage.

In this paper, we propose a method of estimating SOH and RUL simultaneously in such a way as to contribute to its real-world applicability. First, noticing that battery aging causes the time sequence of charging and discharging voltage and current in a cycle to be shortened and dispersed, we define an aging index, referred to here as the time compensated entropy, for SOH and RUL. Second, for LSTM-based RUL prediction, we optimize the number of SOH input and the RUL prediction sequences for the minimum prediction error associated with a sequence of cycles. Third, we adopt a progressive framework of LSTMs such that whatever learned from the prior predictions are transferred to the subsequent prediction, starting with learned SOH. For experimental verification, we train the proposed progressive LSTM network based on CALCE datasets and apply to various cases of charging and discharging cycles. With SOH estimated online, we achieve less than 10 cycles of accuracy in RUL prediction, moving closer to real-world applicability.

1. INTRODUCTION

The advent of lithium-ion batteries has changed our lives. The small, wireless gadgets like the smartphones and wireless devices could have been developed with the advance of lithium-ion battery technology. This technology provides us a breakthrough to overcome spatial and temporal limits and is the key to advance in Electric Vehicles (EV), Robotics and Aerospace industries.

As the importance of batteries grows, the battery management system (BMS) technology becomes more significant. The biggest disadvantage of batteries is the aging. A primary cause of battery aging is a change in the electrode and electrolyte in batteries. Solid Electrolyte Interphase (SEI) is formed on the electrode surface due to chemical reactions during the operation of batteries (Arorat, White, Carolina, & Doyle, 1998). The SEI is becoming more unstable and increasing in quantity as the battery charges and discharges repeatedly. The growth of SEI results in capacity and power fade in batteries (Vetter, Nov, Wagner, & Veit, 2005). Because of the aging effect which induces degradation of batteries and devices, accurate prediction or estimation on the degree of aging in batteries is necessary.

In order to manage the battery efficiently, it is necessary to accurately estimate and manage the battery State of Health (SOH) and Remaining Useful life (RUL). SOH represents the maximum discharge capacity to the rated capacity of batteries (Murnane & Ghazel, 2017). Since the maximum discharge capacity is characteristic of battery aging, SOH is used as an indicator of degree of ageing. SOH of new battery with no degradation is processed in is 100%. When SOH value reaches 80%, that is the time for the replacement and can be regarded as reaching at an End of Life (EOL). RUL represents remaining life until the EOL.

It is major research subject in battery health management to estimate degree of aging in battery. There are many approaches to achieve that goal. These methods could be categorized into three methods, i.e. direct approach, model-based approach and data-driven approach.

In the direct approach, there are few methods using measured values directly. The coulomb counting method is a method of estimating SOH through the calculated discharge capacity by integrating the discharge current (Ng, Moo, Chen, & Hsieh, 2009). The open circuit voltage method is estimating SOH using the relationship between SOH and OCV obtained by prior experiment. Weng, Sun, and Peng (2014) used OCV model to monitor SOC and SOH. The direct approach is easy to implement but having disadvantages of on-line estimation.

The model-based approach utilizes filtering such as kalman filters or particle filters. Liu, Yin, Song, and Peng (2018) estimate SOH using an Unscented Kalman Filter (UKF) with the health indicator as the operating time corresponding to the discharge voltage range. The filter methods have been used to predict RUL. SVR-UPF, which combines a Support Vector Regression (SVR) and an Unscented Particle Filter (UPF), has been proposed by (Peng, Zhang, Yu, & Zhou, 2016). The model-based approach has the advantage of real time estimation, but high computational capacity is needed, and multiple tasks can't be performed on one model.

The data-driven approach uses learning models with pre-performed battery measurement data. Several SOH estimation and RUL prediction studies based on NN structure have been conducted due to its excellent ability to learn features from the data. Especially Long Short-Term Memory (LSTM) is widely used in aging estimation model. You, Park, and Oh (2017) studied various methods of estimating SOH with LSTM structure. Zhang, Xiong, He, and Pecht (2018) proposed the model with LSTM and Monte Carlo simulation for RUL prediction and its uncertainty. This approach is based on the data collected by simulation in laboratory, i.e. voltage, current and temperature etc. It can lead to inaccurate results in real environment applications because these parameters vary depending on the battery operating environment.

We proposed a *time compensated entropy index* from battery voltage distribution which has low variation and a *Progressive LSTM networks* for SOH estimation and RUL prediction. The advantage of our approach is that it is capable of estimating SOH as well as predicting RUL. Many studies considered these two separately. A number of studies to predict the timing of battery replacement have been carried out under the assumption that SOH is known as label data. However, in a real application, BMS needs to estimate SOH and predicting RUL would be made through accurate SOH estimation. In order to perform different tasks with one model, we apply the progressive neural networks (Rusu, Rabinowitz, Desjardins, Soyer, Kirkpatrick, Kavukcuoglu, Pascanu, & Hadsell, 2016) and combine them with LSTM for extracting features from time-series data. Knowledge for SOH estimation accumulated in a network is transferred to another network for predicting sequence of SOH accurately.

In this paper, we show a new health indicator based on information entropy for estimating SOH. Using battery datasets from CALCE in the Maryland university, we demonstrate our progressive networks which use these indices having ability to infer the degree of aging.

Section 2 covers brief introduction to progressive neural networks. Section 3 describes our method including the time compensated entropy index and the progressive LSTM networks. In section 4, we do experiments for evaluating our methods are good at accomplishing our goal. Section 5 comes up with a conclusion.

2. BACKGROUNDS

Progressive neural networks are structures that have been proposed to mimic humans performing new tasks based on past experiences. Existing learning transfer models, such as transfer learning, have the disadvantage that it is impossible to transfer such knowledge that they can do various tasks. To solve these shortcomings, they use the lateral connection that leads to the existing learned network to learn the network to perform the new task (Rusu et al., 2016). It transfers the prior knowledge existing in the existing network to other networks via this connection. Through such a structure, it is possible to extract useful features for a new task.
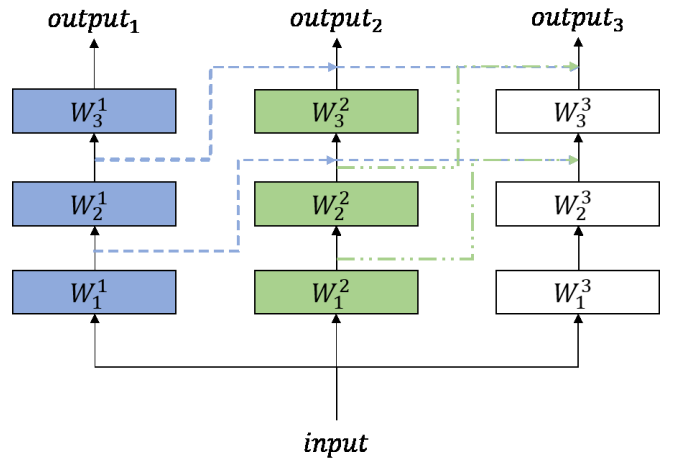

Figure 1. Progressive Neural Networks. Dashed lines are lateral connections for transferring knowledge. Each column learns to deal different tasks.

The figure 1 shows the progressive neural networks structure. The ability to integrate prior knowledge through the structure where the output is transferred to other networks on each floor of each network, hence, reduce oblivion of existing knowledge. The lateral connection in P-NN makes the network capable of doing various tasks. The ability to transfer knowledge and deal with multi-task is the reason why we use P-NN in this paper. From the second column network in P-NN, the output of each hidden layer is written as follows:

$$h_i^k = f\left(W_i^k h_{i-1}^k + \sum_{k<j} U_i^{k:j} h_{i-1}^j \right) \qquad (1)$$

where h_i^k is the hidden output of layer i in a column k, W_i^k is the weight matrix of layer i in column k and $U_i^{k:j}$ is weight matrix of lateral connection from a column k to a column j. $f(x)$ is an activation function such as ReLU function. The learning process in P-NN is sequential. When learning the column j, we freeze weights of the columns before column j and update weights of the column j and weights of all lateral connection from column k to j in all k less than j.

The progressive neural network has been applied in various fields, as it has the advantage of reusing previous results and learning new things. The performance of progressive neural networks is superior to the Transfer Learning or Deep Neural Networks as evident from the experiments on emotion recognition on learned knowledge in order to distinguish speakers and gender in (Gideon, Khorram, Aldeneh, Dimitriadis, & Provost, 2017).

3. PROPOSED METHOD

Aging phenomenon of batteries is caused by physical-chemical changes inside the battery and makes the internal resistance of the battery and the maximum charge capacity be changed (Williard, He, Osterman, & Pecht, 2013). Old batteries are characterized by a wider discharging voltage distribution than that non-aging battery's. It means changes in the distribution can be used as an indirect indicator of battery aging (Lorenzo & Labed, 2005). In this section, the time compensated entropy index based on the battery voltage distribution and the progressive LSTM networks to infer degree of aging is proposed.

3.1. Time Compensated Entropy Index

Entropy in information theory is the concept that Shannon (1948) proposed to represent the amount of information that data has. Information entropy means the number of bits required to represent information. The amount of information increases as the information is chaotic and irregular.

$$H = -\sum_{i=1}^{N} p_i \log_b p_i \qquad (2)$$

Eq. (2), is a basic equation for obtaining information entropy. p_i is the probability that an event i occurs. Information entropy can be used to indicate a disorder, an uncertainty of the system. As the battery ages, the system is disordered due to the transition in the internal characteristics. Information entropy is used to represent the changes emerged as voltage distribution. A voltage in an operating cycle is represented by a histogram. Eq. (2) can be represented as Eq. (3), with the histogram which has M bins.

$$H_k = -\sum_{i=1}^{M} p(x_i^k) \log_{10} p(x_i^k) \qquad (3)$$

where H_k is information entropy in the kth battery cycle and $p(x_i^k)$ is the probability for the bin x_i^k in a histogram of voltage distribution at kth battery cycle. The operating time of the battery varies depending on the operating environment and variation of the information entropy may occur. Information entropy tends to decrease as operating duration is shortened. To minimize this effect, the entropy is normalized by multiplying the normalization term including the duration.

$$E_k = -\frac{\alpha}{D_k} \sum_{i=1}^{M} p(x_i^k) \log_{10} p(x_i^k) \qquad (4)$$

$$E_k = -\frac{\alpha}{C_k \cdot D_k} \sum_{i=1}^{M} p(x_i^k) \log_{10} p(x_i^k) \qquad (5)$$

where E_k is the time compensated entropy index. D_k is the discharging (or charging) duration of the kth cycle and we call this term a time compensated term. α is a length coefficient and can be defined as an operating time of the shortest cycle in the data. In Eq. (5), we can compensate entropy index in Eq. (4) by the discharging current additional term C_k under the condition that discharging current changed. Discharging Time Compensated Entropy Indices and Charging Time Compensated Entropy Indices come from the voltage distribution during discharging and charging.

3.2. Progressive LSTM model

A progressive LSTM model (P-LSTM) with input vectors including time compensated entropy index was designed to predict battery SOH and RUL. The goal of our model is multiple performance capability to accurately estimate the SOH of the battery in time, predict the future SOH sequence through the predicted SOH and obtain it to the RUL. The progressive neural networks (P-NN) is a suitable structure for these objectives. The structure can effectively communicate knowledge between networks to perform multiple tasks. We have designed a P-LSTM based on LSTM that is suitable for extracting features of time series data according to the battery operation.

Figure 2 describes our models. Our model includes two P-LSTM networks: 1) A network that estimates SOH at the current time t and predict SOH sequences based on predicted SOH sequences. 2) A network that predicts future SOH sequences until life of batteries reaches EOL through the predicted SOH sequence at a first network. Both networks are based on P-LSTM architecture. The hidden state in the column is combined with the hidden state of the previous columns through the lateral connection. In P-LSTM, h_i^k is the hidden state of the layer i in the kth column:
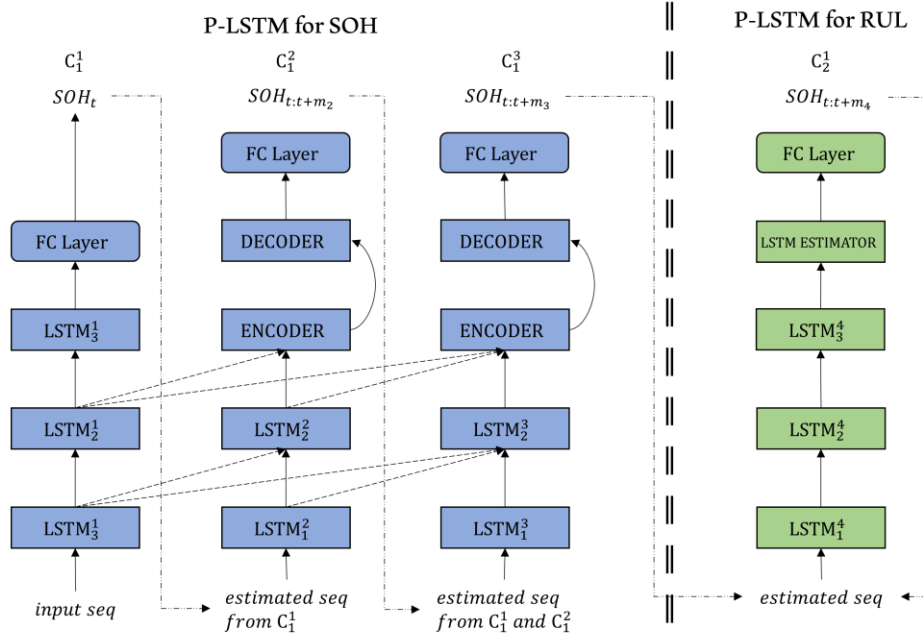
Figure 2. P-LSTM for SOH and RUL

$$h_i^k = f\left(L_i^k(h_{i-1}^k) + \sum_{k<j} U_i^{k:j} h_{i-1}^j\right) \qquad (6)$$

where U is weighted matrix in lateral connection from column k to j and L_i^k is LSTM layer i in the kth column. we use ReLU function as $f(x)$.

3.2.1. P-LSTM for SOH

The first P-LSTM network consists of a P-NN structure with three network columns performing different roles.

The first column (C_1^1) estimates SOH at the current point in time. Inputs of C_1^1 is the vector u_t including the time compensated entropy indices. When voltage data at the cycle t is given, $E_t = [(DE_t, CE_t)]$ is the time compensated entropy indices. where DE_t is the discharging time compensated entropy index at the cycle t and CE_t is the charging time compensated entropy index at the cycle t. To specify battery's operation, input vector includes a discharge voltage distribution. For a given data set until the cycle t, input vector u_t^1 with sequence length k_1 is written as:

$$u_t^1 = \left[(E_{t-k_1-1}, P_{t-k_1-1}), \cdots, (E_t, P_t)\right]^T \qquad (7)$$

where $P_t \in \mathbb{R}^n$ is a discharge voltage distribution with $P_t = (p(x_1), p(x_2), \cdots, p(x_n))^T$. The input vector passes through the LSTM layer on each floor and is changed to knowledge for the next columns. After three LSTM layers and a fully connected layer, SOH at the cycle t S_t as a final output is obtained. Accumulated results used as the input vector of C_1^2.

The second column (C_1^2) predicts a SOH sequence as much as the prediction cycle m_2 from the current cycle. C_1^2 consists of two LSTM layers and one LSTM Enc-Dec. The Enc-Dec is responsible for predicting the SOH sequence of variable lengths. An input vector of C_1^2 made from C_1^1 is estimated SOH sequence. u_t^2 is input vector of the column 2 at the cycle t and can be written as:

$$u_t^2 = \left[S_{t-k_2-1}, \cdots, S_t\right]^T \qquad (8)$$

where k_2 is an input sequence length of C_1^2. The output h_2^l of the LSTM layer l in the second column is added to the output of the LSTM layer l in C_1^1 multiplied with the lateral connection weight. This result is used as an input to the next layer via ReLU function. Through this process, knowledge of C_1^1, which estimates S_t from the battery measurement data including the compensated index, is used to predict the future SOH sequence in C_1^2. The output o_t^2 is:

$$o_t^2 = \left[\bar{S}_t, \cdots, \bar{S}_{t+m_2}\right]^T \qquad (9)$$

where $\bar{S}_t$ is predicted SOH at cycle t and m_2 is a length of prediction in C_1^2.

The third column (C_1^3) predicts a SOH sequence with longer time-length than an output of C_1^2. Estimated SOH sequence from C_1^1 and o_t^2 are concatenated and given as input vector u_t^3 as follows:

$$u_t^3 = \left[S_{t-k_3-m_2}, \cdots, S_{t-1}, \bar{S}_t, \cdots, \bar{S}_{t+m_2-1}\right]^T \qquad (10)$$

where k_3 is an input sequence length of C_1^3. Structure of C_1^3 have same structure with C_1^2. Outputs from the layer l in C_1^1 and C_1^2 are transferred to C_1^3 via the lateral connection. A SOH sequence with longer time can be predicted based on the knowledge to estimate S_t from C_1^1 and to predict future SOH sequences from C_1^2. Predicted SOH sequence with m_3 as length of prediction in C_1^3 can be written as:

$$o_t^3 = \left[\bar{\bar{S}}_t, \cdots, \bar{\bar{S}}_{t+m_3-1}\right]^T \tag{11}$$

where $\bar{\bar{S}}$ is predicted SOH in C_1^3.

3.2.2. P-LSTM for RUL

The second P-LSTM network's goal is to predict the SOH sequence from a certain point in time, which we call the starting point, to the EOL. The second P-LSTM network consists of a network column C_2^1 with LSTM layers and LSTM estimator. C_2^1 predicts a SOH sequence whose length is m_4 from input vector with k_4 of length repeatedly. C_2^1 has the input vector v_t^1 which concatenates o_t^3 and estimated SOH values from C_1^1. After predicted SOH sequences are larger than k_4, v_t^1 comes from results of C_2^1 only.

For example, EOL is 510 cycle when the starting point is 300th cycle, m_4 is 40 and k_4 is 100, the predicting process in C_2^1 from starting point is written as:

$$\begin{cases} \left\{\bar{\bar{S}}_{240}, \cdots, \bar{\bar{S}}_{339}\right\} \rightarrow \left\{\hat{S}_{340}, \cdots, \hat{S}_{379}\right\} \\ \left\{\bar{\bar{S}}_{280}, \cdots, \hat{S}_{379}\right\} \rightarrow \left\{\hat{S}_{380}, \cdots, \hat{S}_{419}\right\} \\ \left\{\bar{\bar{S}}_{320}, \cdots, \hat{S}_{419}\right\} \rightarrow \left\{\hat{S}_{420}, \cdots, \hat{S}_{459}\right\} \\ \qquad\qquad\qquad \vdots \\ \left\{\hat{S}_{400}, \cdots, \hat{S}_{499}\right\} \rightarrow \left\{\hat{S}_{500}, \cdots, \hat{S}_{EOL}, \cdots, \hat{S}_{539}\right\} \end{cases} \tag{12}$$

where $\hat{S}_t$ is predicted SOH from C_2^1. Some experiments are conducted to determine model parameters like sequence lengths and output lengths.

4. EXPERIMENTS AND RESULTS

4.1. Datasets

Experiments to evaluate the performance of proposed model used data from the battery aging experiments provided by the CALCE Battery Research group at the University of Maryland (He, Williard, Osterman, & Pecht, 2011). The dataset was created by repeatedly charging and discharging the battery over several months. The battery named CS2 used in the experiment is a lithium-ion battery with 1.1Ah of rated capacity. We use 6 experimental datasets conducted with CS2, i.e. CS2_33, CS2_34, CS2_35, CS2_36, CS2_37 and CS2_38. These cases are divided into two types based on discharge current. CS2_33 and CS2_34 cycled at constant discharge current of 0.5C while others cycled at constant discharge current of 1C. Figure 3 shows the degeneration curves of test cases.
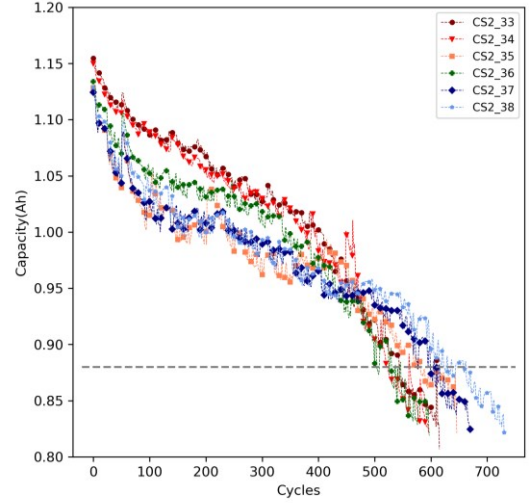

Figure 3. Degeneration curves to the vicinity of EOL from CS2 dataset. EOL is the point where curve reaches at 80% of rated capacity (dashed horizontal line).

The maximum capacity of the battery is lower as the battery operates repeatedly. EOL is where the degeneration curve reaches at a grey dashed line of 80% of rated capacity in figure 3. The raw degeneration curve is a signal with frequent change in short period. However, this variation can interfere with estimation of SOH and RUL for test data. Therefore, a gaussian filter defined in 1-D is applied to raw degeneration curves. Filtered signals are used as label data in training data for RUL prediction.

4.2. Model Configuration

We conducted experiments to evaluate our method in terms of efficient estimation and prediction of the battery SOH. The training dataset includes CS2_33, CS2_34, CS2_35, CS2_37. CS2_38 and CS2_36 are used as the validation dataset and the testing dataset, respectively. Training procedure is as follows:

1. Design the P-LSTM model with the specified parameters. The input vector is composed of the calculated time compensated entropy index, current and voltage distribution. At this time, the bin number of the voltage distribution is smaller than the size used in calculating the entropy index in order to prevent dimension of the input vector from becoming too large.

2. Train P-LSTM for SOH. First C_1^1, which estimates SOH as an input vector, is learned. Second, C_1^2 is trained by inputting the result of C_1^1. The weight of C_1^1 freezes while learning training C_1^2. At last train C_1^3 in freeze state of weights of C_1^1 and C_1^2.

3. Train P-LSTM for RUL. Capacity data which is label data for training P-LSTM for SOH used as training data in this step.

5

(a) sequence length: 80   (b) sequence length: 100   (c) sequence length: 120

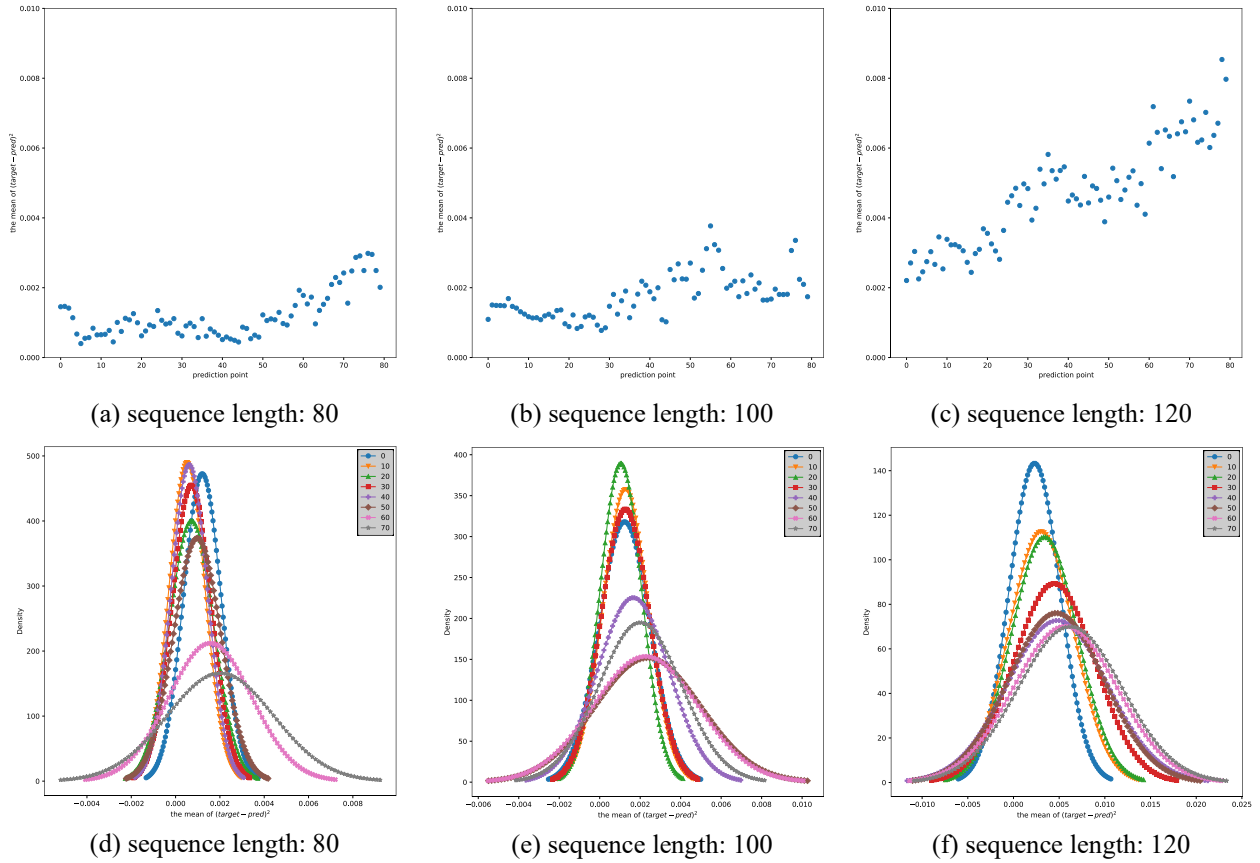(d) sequence length: 80   (e) sequence length: 100   (f) sequence length: 120

Figure 4. The average of errors on each prediction points are described in the first row. The distributions of error (d), (e) and (f) widen according to increasing of prediction length. Optimal sequence length is 80, which shows good performance in terms of error and variance.

The optimizer for model learning uses AdamOptimizer (Kingma & Ba, 2015). The parameters for the P-LSTM configuration is listed in Table 1. Parameters selection, such as sequence lengths, determines the performance of the model. But for most applications the parameters are determined by empirical methods. To determine parameters in P-LSTM, we conduct a simple experiment through one column used in P-LSTM to efficiently determine the seq length and prediction cycle. The average and variance of the error according to the input sequence length and prediction length are applied to choose parameters.

The experiment was performed on the same model as C_2^1 with size of hidden states is 32. Three experiments were performed in which the prediction cycle was fixed at 80 cycles and the sequence length was changed to 80, 100 and 120. The networks learn from training datasets and evaluate the accuracy with prediction at several cycles from 100 to 400 at 50 intervals with the test datasets. The results are listed in the Table 2.

Table 1. Parameters.

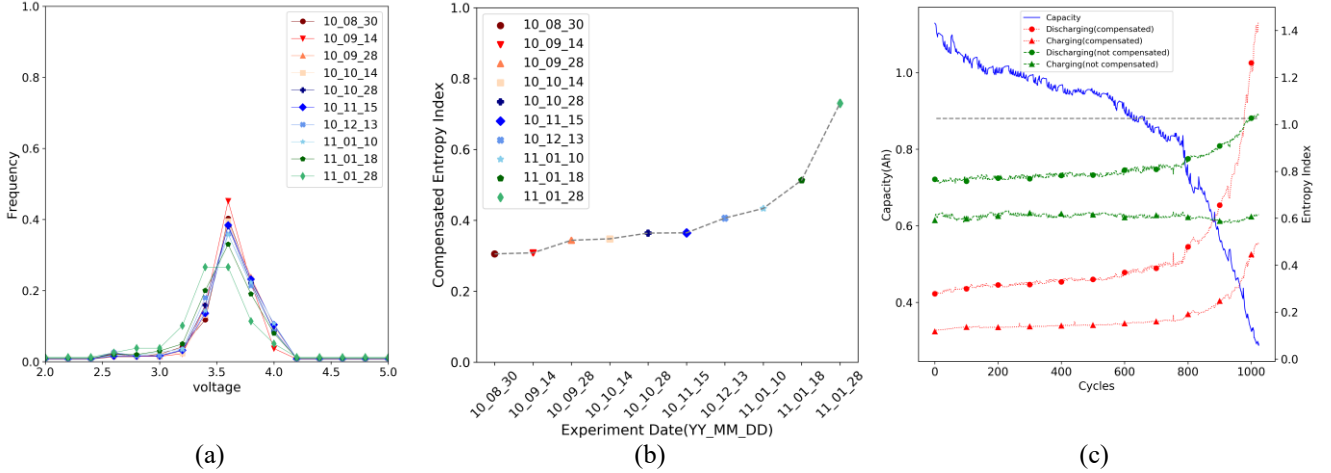| Parameter	Value
k_1, k_2	20
k_3	40
k_4	80
m_2	20
m_3	40
m_4	40
hidden state size at C_1^1	5
hidden state size at C_1^2	10
hidden state size at C_1^3	10
hidden state size at C_2^1	32

Figure 5. (a): Change of discharging voltage distribution of CS2_38. As the battery gets older, the distribution becomes wider. (b): The discharging time compensated entropy indices from (a). (c): The capacity degeneration curves and the entropy curves at discharging (circular mark) and charging (triangular mark) whether compensated or not.

Table 2. The results of error variation according to the sequence lengths and prediction cycles.

Sequence Length	Prediction Cycle	$E_{[0:40]}$	E_{all}
80	80	0.001197	0.000879
100	80	0.001750	0.001315
120	80	0.004682	0.003705

E_{all} is the mean of $(prediction - target)^2$ of all cycles in prediction coverage while $E_{[0:40]}$ is error of 40 cycles. Both errors are lowest if the sequence length is 80. Figure 4 describes results including error per each prediction points and variance per 10 cycles. In all conditions, the error increases as the point far from the starting point is predicted; see the figure 4 (a), (b) and (c). The distribution of errors also gets bigger and bigger; see the figure 4 (d), (e) and (f). In figure 4 (a) and (b), the error increase rapidly after the prediction point 40. It means prediction results in a column is more accurate until 40 cycles. Therefore, we determine the sequence length and prediction cycle in P-LSTM for RUL as 80 and 40, respectively.

4.3. Results

4.3.1. Time Compensated Entropy Index

Figure 5 shows the time compensated entropy index in CS2_38 according to the battery operating cycles. To simplify, one cycle per each experiment that were performed on different days is considered. Time compensated entropy indices of battery are calculated based on voltage distribution in Eq. (4). Figure 5 (a) describes discharging voltage distributions based on histogram which has 16 bins in different experiment dates. The distribution has a uniform distribution as the experiment continues. The wider distribution means, the more uncertain in expected measurement voltage. The uncertainty makes information entropy increase. In figure 5 (b), discharging time compensated entropy index (DE) from distributions in figure 5 (a) change as time goes by. Eq. (4) is used to calculate DE with $\alpha = 50$. DE increases over time as the variance of voltage distribution increases. The change in time compensated entropy index until the maximum charge capacity reaches about 0.2 is shown in figure 5 (c). In figure 5 (c) we consider the variation in entropy indices depending on whether compensated or not about every cycle in CS2_38. When entropy index is obtained from Eq. (4) with a time compensated term, the tendency to increase with aging becomes more pronounced than without compensated term like in Eq. (3). As chargeable capacity decreases, time compensated entropy index in charging and discharging have drastic growth.

4.3.2. Estimation SOH and RUL based on P-LSTM

Through the designed P-LSTM for SOH and RUL, experiment to estimate the SOH per cycle and predict the future SOH sequence is conducted. The results of estimation of SOH in C_1^1 from battery discharging information including entropy are described in the figure 6. CS2_36 dataset is used as test dataset. Predicting SOH based on battery measurement data can be done by SOH estimating column network and time compensated entropy indices. An additional experiment was conducted to estimate SOH using CS2_33 discharged with 0.5C current. In this case, CS2_34, CS2_36, CS2_37, CS2_38 are training datasets and C2_35 is used as a validation dataset. For predicting near future SOH, estimated

SOH values from C_1^1 and C_1^2 are utilized. The results from two experiments with P-LSTM for SOH are in the table 3.
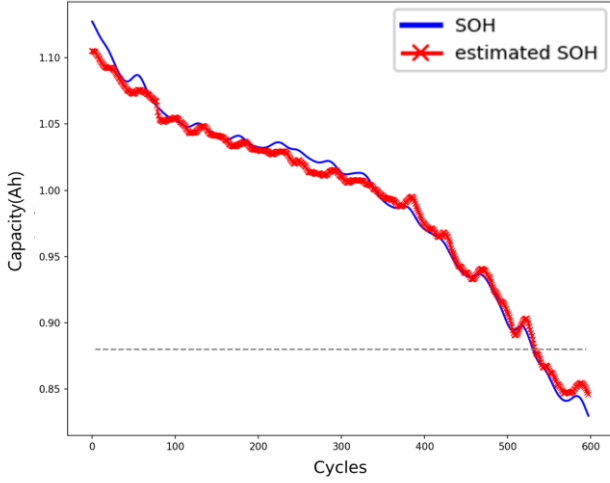

Figure 6. SOH estimation results with CS2_36.

Table 3. The results of SOH estimation from P-LSTM for SOH.

Test Case (Discharge)	Column	Prediction Target	RMSE
CS2_36 (1C)	C_1^1	SOH_t	0.0067
	C_1^2	$SOH_{t:20}$	0.0077
	C_1^3	$SOH_{t:40}$	0.0098
CS2_33 (0.5C)	C_1^1	SOH_t	0.0211
	C_1^2	$SOH_{t:20}$	0.0147
	C_1^3	$SOH_{t:40}$	0.0209
*CS2_33 (0.5C)	C_1^1	SOH_t	0.0201
	C_1^2	$SOH_{t:20}$	0.0267
	C_1^3	$SOH_{t:40}$	0.0295

Due to imbalance in datasets which have 4 datasets with 1C and 2 datasets with 0.5C, RMSE results vary depending on the discharge current value. The result shows, for 1C discharging current datasets which are relatively abundant, the accuracy is relatively high while bad performance where 0.5C charging current dataset is used as the test dataset. The additional experiment, *CS2_33 on a table, is predicting SOH of battery discharged by 0.5C with the network trained by datasets only with discharging current of 1C. Our method can be applied to several discharge currents through the current compensated term in Eq. (5). The result of the experiment under limited conditions shows that the proposed time compensated entropy index can be applied in an environment where the discharge current varies at each cycle. A more accurate estimation method and its application in a

changing current environment during one operating cycle will be performed in future studies.

SOH prediction until EOL is generated from C_2^1. Obtained near future SOH from C_1^3 is followed by predicted SOH sequences. After a single prediction with m_4 length is performed, the prediction is utilized as an input to make repeated prediction. Two predictions are conducted in two cases. When CS2_36 is used as a test dataset, actual target EOL (EOL_l) is at 504th cycle and predicted EOL (EOL_p) is at 544th cycle. In the experiments with CS2_33, EOL_l is at 544 and EOL_p is at 530th cycle. The difference between EOL_l and EOL_p at each case is 9 and 14, respectively. In the proposed P-LSTM structure, based on the experience obtained from C_1^1, the model predicts near future SOH from C_1^3 through C_1^2. The RUL prediction is performed without prior learning before the prediction point as well as using not the actual SOH but estimated SOH. In Table 4, we can compare our results with from (Liu, Zhao, Peng, & Hu, 2017) which conduct experiments with same battery dataset used in this paper. Proposed P-LSTM model uses estimated SOH to predict RUL while Liu et al. (2017) assumes label SOH is given as input data. Nevertheless, our model can predict EOL quite well.
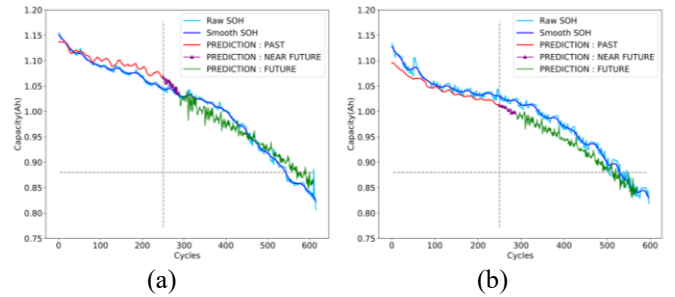

(a)   (b)

Figure 7. The results from P-LSTM for SOH and RUL. (a) and (b) are the results according to test datasets 0.5C and 1C, respectively.

Table 4. RUL prediction results.

Methods	Test Case (Discharge)	Start from	Error (cycle)
Proposed P-LSTM	CS2_33 (0.5C)	250	14
	CS2_36 (1C)	250	9
LSTM from Liu et al. (2017)	CS2_36	322	1
	CS2_36	422	21

Prediction on short-term SOH is critical factor in long-term RUL forecast. The prediction from the start point of the prediction to the EOL is performed independently without the help of the other columns, hence it is inevitably affected by

the result of C_1^3. In order to assure the influence of short-term SOH prediction results on RUL prediction, two prediction was conducted by using different C_1^2 inputs in the same network. The input vector of C_1^2 includes 1) the output from C_1^1 (C_1^1 only) and 2) the outputs from C_1^1 and C_1^3 (C_1^3 with C_1^1). Figure 8 shows the results of two predictions. Figure 8 and Table 5 show predicted near-future SOH from P-LSTM for SOH helps RUL prediction from P-LSTM for RUL get accuracy.
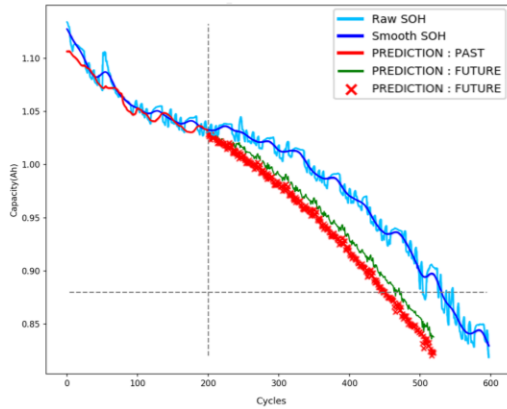

Figure 8. Prediction graph from 200th cycle of P-LSTM for RUL according to whether short-term future prediction results are included in input. **red-cross mark**: prediction with output of C_1^1 only. **green line**: prediction with outputs from C_1^3 with C_1^1.

Table 5. RUL prediction error according to composition of the input vector of C_2^1.

case	EOL$_p$ (*prediction*)	EOL$_l$ (*label*)	Error (cycle)
C_1^3 with C_1^1	485	504	19
C_1^1 only	449	504	55

5. CONCLUSION

Estimating the aging information of lithium-ion batteries is an important research topic because degeneration of lithium ion batteries degrades system performance and lower reliability. However, lithium-ion batteries have difficulty in estimating the aging state due to the operation environment and internal complexity.

In this paper, we propose a time compensated entropy index that represents degree of aging by the concept of entropy. The time compensated entropy index confirm that the value increased according to degree of aging and can be used as an indicator to show the degradation phenomenon. We also proposed a P-LSTM structure that can predict battery SOH and RUL together using them. The proposed structure has the advantage of being able to perform two tasks in a single structure, away from the traditional SOH estimation and RUL

prediction approaches. The accurate SOH estimation which is a preemptive condition of RUL prediction is performed with the time compensated entropy index and enable accurate SOH prediction of passing knowledge acquired to the SOH prediction column. The proposed model is able to predict SOH until EOL without any additional training with test data. We used the battery datasets from the CALCE battery to demonstrate that the structure can estimate and predicting SOH.

The current research requires demonstration and improvement for experiments in complex discharge conditions. Therefore, the future issues are the application in more diverse environments and more investigation on optimization the number of SOH input and the RUL prediction sequences for the optimal parameter selections.

REFERENCES

Arora, P., White, R. E., & Doyle, M. (1998). Capacity fade mechanisms and side reactions in lithium-ion batteries. *Journal of the Electrochemical Society*, *145*(10), 3647-3667.

Vetter, J., Novák, P., Wagner, M. R., Veit, C., Möller, K. C., Besenhard, J. O., ... & Hammouche, A. (2005). Ageing mechanisms in lithium-ion batteries. *Journal of power sources*, *147*(1-2), 269-281.

Murnane, M., & Ghazel, A. (2017). A closer look at state of charge (SOC) and state of health (SOH) estimation techniques for batteries. http://www. analog. co m/media/en/technical-documentation/technical-articles/A-Closer-Look-at-State-Of-Charge-and-State-Health-Estimation-Techniques-.... pdf.

Ng, K. S., Moo, C. S., Chen, Y. P., & Hsieh, Y. C. (2009). Enhanced coulomb counting method for estimating state-of-charge and state-of-health of lithium-ion batteries. *Applied energy*, *86*(9), 1506-1511.

Weng, C., Sun, J., & Peng, H. (2014). A unified open-circuit-voltage model of lithium-ion batteries for state-of-charge estimation and state-of-health monitoring. *Journal of power Sources*, *258*, 228-237.

D, Liu., X, Yin., Y, Song., W, Liu., and Y, Peng. (2017). An On-Line State of Health Estimation of Lithium-Ion Battery Using Unscented Particle Filter. *IEEE Access, vol. 6.* pp. 40990-41001. doi: 10.1109/ACCESS.2018.2854224

Peng, X., Zhang, C., Yu, Y., & Zhou, Y. (2016). Battery remaining useful life prediction algorithm based on support vector regression and unscented particle filter. In *2016 IEEE International Conference on Prognostics and Health Management (ICPHM)* (pp. 1-6). IEEE.

You, G. W., Park, S., & Oh, D. (2017). Diagnosis of electric vehicle batteries using recurrent neural networks. *IEEE Transactions on Industrial Electronics*, *64*(6), 4885-4893.

Zhang, Y., Xiong, R., He, H., & Pecht, M. G. (2018). Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries. *IEEE Transactions on Vehicular Technology*, *67*(7), 5695-5705.

Rusu, A. A., Rabinowitz, N. C., Desjardins, G., Soyer, H., Kirkpatrick, J., Kavukcuoglu, K., ... & Hadsell, R. (2016). Progressive neural networks. *arXiv preprint arXiv:1606.04671*.

Gideon, J., Khorram, S., Aldeneh, Z., Dimitriadis, D., & Provost, E. M. (2017). Progressive neural networks for transfer learning in emotion recognition. *arXiv preprint arXiv:1706.03256*.

Williard, N., He, W., Osterman, M., & Pecht, M. (2013). Comparative analysis of features for determining state of health in lithium-ion batteries. *Int. J. Progn. Health Manag*, *4*, 1-7.

Lorenzo, E., & Labed, S. (2005). The battery voltage distribution: a possible tool for surveying the state of health of stand-alone PV systems. *Progress in Photovoltaics: Research and Applications*, *13*(3), 251-260.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell system technical journal*, *27*(3), 379-423.

He, W., Williard, N., Osterman, M., & Pecht, M. (2011). Prognostics of lithium-ion batteries based on Dempster–Shafer theory and the Bayesian Monte Carlo method. *Journal of Power Sources*, *196*(23), 10314-10321.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Liu, Y., Zhao, G., Peng, X., & Hu, C. (2017). Lithium-ion Battery Remaining Useful Life Prediction with Long Short-term Memory Recurrent Neural Network. *Proceedings of the Annual Conference of the PHM Society*, 8(7).

BIOGRAPHIES

Taejun Bak received his B.S degrees in Electrical Engineering from Sungkyunkwan University in 2017. From 2016 to 2017, he was the engineer at ASML Korea. His research interests are in the areas of PHM based on intelligent systems, and time series analysis with machine learning. He is currently a Master's student at Sungkyunkwan University.

Sukhan Lee received the B.S. and M.S. degrees in electrical engineering from Seoul National University, South Korea, in 1972 and 1974, respectively, and the Ph.D. degree in electrical engineering from Purdue University, West Lafayette, IN, USA, in 1982. From 1983 to 1997, he was with the Departments of Electrical Engineering and of Computer Science, University of Southern California. From 1990 to 1997, he was with the Jet Propulsion Laboratory, California Institute of Technology, as a Senior Member of the Technical Staff. From 1998 to 2003, he was the Executive Vice President, and also the Chief Research Officer with the Samsung Advanced Institute of Technology. He has been working as a Professor of information and communication engineering and WCU Professor of interaction science with Sungkyunkwan University, since 2003. He was designated the Dean of the Graduate School of Sungkyunkwan University, in 2011.He is also serving as the Director of the Intelligent Systems Research Institute. His research interests are in the areas of cognitive robotics, intelligent systems, and micro/nano electro-mechanical systems. Dr. Lee is currently a Fellow of the Korean National Academy of Science and Technology