Advanced binary search pattern for impedance spectra classification for determining the state of charge of a lithium iron phosphate cell using a support vector machine

Patrick Jansen¹,³, Michael Vollnhals¹, Daniel Renner¹,³, David Vergossen¹, Werner John²,³, and Jürgen Götze³

¹Audi Electronics Venture GmbH, Gaimersheim, Germany
²SiL GmbH – Paderborn/TU Dortmund, Dortmund, Germany
³Technische Universität Dortmund (AG DAT), Dortmund, Germany

Correspondence to: Patrick Jansen (patrick.jansen@audi.de)

Received: 16 December 2015 – Accepted: 19 May 2016 – Published: 28 September 2016

Abstract. Further improvements on the novel method for state of charge (SOC) determination of lithium iron phosphate (LFP) cells based on the impedance spectra classification are presented. A Support Vector Machine (SVM) is applied to impedance spectra of a LFP cell, with each impedance spectrum representing a distinct SOC for a predefined temperature. As a SVM is a binary classifier, only the distinction between two SOC can be computed in one iteration of the algorithm. Therefore a search pattern is necessary. A balanced tree search was implemented with good results. In order to further improvements of the SVM method, this paper discusses two new search pattern, namely a linear search and an imbalanced tree search, the later one based on an initial educated guess start value for the SOC. All three search pattern are tested with a LFP cell by analyzing each impedance spectrum for every 10°C temperature step from −30 to +40°C and every 5% SOC step from 0 to 100%. While being trained with the measured impedance spectra the SVM is tested with an added disturbance to training data.

As the SVM is a binary classifier only two impedance spectra can be computed at one time. As a 5% SOC resolution leads to 21 impedance spectra efficient search algorithms are needed. As described in Jansen et al. (2015) balanced tree search bades on an divide and conquer algorithm

1 Introduction

As shown in Jansen et al. (2015) a support vector machine (SVM) is a powerful possibility to determine the state of charge (SOC) as defined in (Sauer et al., 1999) of a lithium iron phosphate cell. Based on frequency domain data for a specific state of Health and defined temperature states. As an alternative to common SOC estimation methods such as the state space observer or the kalman filter the SVM resigns an electric circuit diagram (ECD) with components such as the open circuit voltage curve and the element parameters of the electric network of said ECD (Piller et al., 2001, Plett, 2004, Codeca et al., 2008). The mentioned SVM and their characteristics have been discussed and it has been demonstrated that the method can correctly determine the SOC for a broad temperature range with respect to the signal to noise raito.

In order of further improvements on the SVM method this paper discusses two new search pattern, namely a linear search and an imbalanced tree search pattern, the later one based on an educated guess. All three search pattern were compared under various aspects like accuracy, efficiency, tolerance of disturbances and temperature dependancy. The imbalanced search tree shows to be the most efficient search pattern if the initial guess is within less than ±5% SOC of the original SOC in both directions and exhibits the best tolerance for high disturbances. Linear search improves the rate of exact classifications for almost every temperature. It also improves the robustness against high disturbances and can even detect a certain number of false classifications which makes this search pattern unique. The downside is a much lower efficiency as all impedance spectra have to be evaluated while the tree search pattern only evaluate those on the tree path.
including was used. For optimization on SVM for SOC determination this paper discusses two new search algorithms, namely a linear search and an imbalanced tree search pattern are shown. These suggestions can improve several aspects of the SVM method like low successful classification rates at extreme temperatures.

2 Principle and methodology

The SVM is a binary classifier from the field of machine learning theory. The SVM is a support vector learning algorithm for pattern recognition, with the aim of classifying quantities with certain attributes and grade unknown samples to one of two classes. A binary support vector classifier such as the SVM is based on a class of linear hyperplanes, to separate a number of elements into two specific classes, based on class specifying attributes using a hyperplane, see Fig. 1.

The SVM is also applicable to non linear separable data, by using the so-called “Kernel Trick” to transform the data into a high-dimensional feature space where the data is linear separable. The kernel depends on several usable functionalities – for instance a polynomial or a radial basis function – to evaluate the hyperplane that separates the data in the feature space. A suitable kernel function has to be chosen specifically for the training data.

2.1 Impedance spectra classification

Impedance spectra classification is a valid technique to obtain information about electrochemical cells. Parameters like state of charge (Li et al., 2010) and state of health (Waag et al., 2010) can be derived from an impedance spectrum among others. However, these methods exhibit certain shortcomings concerning accuracy, reproducibility and sturdiness against disturbances and cell aging. Determining the SOC via impedance spectra classification using SVM is an alternative method to achieve this aim.

The task of determining the SOC of a lithium iron phosphate cell can be achieved with an optimal classifier such as the SVM by grading measured impedances to a certain class respectively a classified impedance spectrum. The classes for the SVM are represented by all the impedance spectra for different SOC levels, corresponding to defined temperature and aging states, generated by an impedance spectroscopy, are the foundation of this classification method (see Fig. 2). The data of those spectra represents the training data of the SVM classification function – based on a polynomial kernel function – that is used to calculate the hyperplanes that separate all impedance spectra to each other. As noted above, the SVM is only capable of a binary decision, therefore with more than two classes a separation of every impedance spectrum to each other has to be realized by a hyperplane via SVM. So $n$ classes will yield to $n - 1$ hyperplanes to be evaluated by the SVM to separate all spectra.

The SVM decision function can only make binary decisions so that all the SVM decisions have to be rated and contextualized. The most efficient way to do so is to create a graph to arrange the hyperplanes. The whole quantity of the impedance spectra elements or the superset, the root of the graph, is separated by the median of the hyperplanes represented by their specific linear combinations,

$$\tilde{w}_{\text{med}} = \sum_k \tilde{v}_{k,\text{med}} \gamma_k \tilde{x}_{k,\text{med}}$$

of lagrange multipliers $\tilde{v}_{k,\text{med}}$, corresponding class labels $\gamma_k$ and the support vectors $\tilde{x}_{k,\text{med}}$ of the represented SOC classes $c_i$,

$$\tilde{x}_{k,\text{med}} \in \bigcup_i \{c_i\}$$
The two new nodes of the graph represent the two roughly equal power sets of the above superset. The separation of the generated power sets can be repeated recursively down to the power set elements representing a single impedance spectrum or SOC class \( c_i \). The resulting graph corresponds to a binary search tree, where the root is the whole superset of all elements from the impedance spectra with the nodes as a power set of its parent superset and the leaves representing the single SOC classes (see Fig. 3).

This binary search tree can easily be parsed by a binary tree search algorithm where the edges of the graph represent the binary decisions of the SVM decision function. So a binary tree search algorithm such as a divide and conquer search algorithm, applied to the aforementioned binary search tree, makes it possible to grade measured impedances \( Z_i \) using the SVM decision function,

\[
f_{\text{SVM}}(Z) = \sgn\left( \sum_{k=1}^{l} \tilde{v}_k, \text{med} \cdot y_k \langle Z_i - \tilde{x}_k, \text{med} \rangle^d + b \right)
\]

where \( d \) indicates the degree of the used polynomial kernel function \( k \).

\[
k(Z_i - \tilde{x}_k, \text{med}) = |Z_i - \tilde{x}_k, \text{med}|^d
\]

2.2 Impedance grading

The SOC of the cell can now be determined by grading at least one measured impedance \( Z \) from the cell under load conditions. The impedance spectra of the relevant defined SOC levels therefore have to be classified, as described above. The measured impedance now has to be graded to a single SOC class – SOC specific impedance spectrum – to correlate the SOC of the cell. The binary decision of the SVM decision function can only prove for two classes to which class, separated by the hyperplane, the measured impedance belongs, see Fig. 4. By using a divide and conquer search
algorithm on the binary search tree with the SVM decision function as a key criterion, the SOC can be determined by multiple binary decisions along the search tree.

This divide and conquer algorithm starts at the median hyperplane of all separated impedance curves, that separates this superset into two roughly equal power sets. The binary decision, of the SVM decision function, whether the measured impedance belongs to the power set on one side of the hyperplane or the other decreases the quantity of relevant SVM decisions by half. The remaining power set after the decision containing the measured impedance will therefore also be divided by its median hyperplane into two subsidiary power sets. By continuing this recursive structure the SVM decisions on the binary search tree ultimately grade the measured impedance to a single dedicated SOC class. This class represents the SOC of the cell for the measured impedance. This implementation of this binary SVM based tree search algorithm makes an optimal execution time or complexity respectively of $O(n \log n)$ for the grading of one measured impedance possible.

### 2.3 Alternative classification methods

Other methods for classification problems such as the nearest neighbor decision (NND) and its derivatives such as the $k_{NN}$-nearest neighbor decision are still under investigation. NND-methods use euclidean metrics to evaluate the distance $d$ between the sample and its classified single-nearest neighbor or $d_{NN}$ for $k_{NN}$-nearest neighbors.

An a priori assumption of the underlying statistics of the training data as for a Bayes classifier is not necessary. Therefore the big advantages of this method are the simplicity and performance. A disadvantage in this regard is the fact that, as sets of training data increase, the classification probability decreases (Cover and Hart, 1967 and Dudani (1976)). Compared to the SVM method, NND-methods have higher costs in terms of memory space to store the entire volume of training data and in terms of runtime because all the training data has to be evaluated to grade a single sample. SVMs on the other hand resign on the storage of the training data. Their aim is to detect a pattern within the training data via its class-specific attributes so that the data can be intersected by a hyperplane based on support vectors and the training data can be separated without errors (Cortes and Vapnik, 1995).

### 3 Further investigations on the SVM-Method

To improve the SVM method two modifications on the determination algorithm to contextualize the binary SVM results are investigated. Therefore two new search methods are proposed, a linear search and an imbalanced tree search pattern.

#### 3.1 Linear search

The Linear search pattern computes a result for every hyperplane between two adjacent SOC classes. Fig. 5 illustrates the linear search for different SOC classes and hence hyperplanes between them. While the established balanced tree search algorithm only needs to compute hyperplanes along one tree path with a maximum of $O(\log n)$ steps, the linear search on the other hand computes all the hyperplanes with the need of $O(n - 1)$ steps to generate a result vector with $n - 1$ binary decisions of every SVM corresponding hyperplanes.

$$s = 1 \quad s = 2 \quad \ldots \quad s = n - 2 \quad s = n - 1$$

\[
\begin{array}{cccccc}
\pm 1 & \pm 1 & \ldots & \pm 1 & \pm 1
\end{array}
\]

This Vector then needs to be interpreted to identify the SOC class corresponding to the unknown measured impedance. The Position in the vector where the prefix of the result flips indicates the SOC class corresponding impedance spectrum between the two hyperplanes. Problematic with this method is its error prone against false classifications. Several prefix changes in the result vector are possible, because of said false classifications. A statistic verification is the solution to this problem. With more than one measured impedance to be graded, the corresponding impedance spectrum can be detected, from the most often classified SOC class. This way a certain number of false classifications of the hyperplanes can be tolerated without corrupting the correct result.
This method is especially useful with a high number of hyperplanes, where several errors can be corrected. For the tree search pattern every wrong result of a hyperplane on the search path definitely leads to a wrong SOC estimation.

3.2 Imbalanced tree search

The imbalanced tree search pattern is a different approach based on an imbalanced search tree. While the common pattern from Jansen et al. (2015) is based on a balanced search tree, where the maximum difference in height between two arbitrary nodes is one, an imbalanced search tree on the other hand, has its root of the tree close the estimated SOC class and the height of two arbitrary nodes can vary. An initial educated guess of the SOC – e.g. via voltage measurement – defines the starting point with the trunk of the imbalanced search tree. Figure 6 shows the tree structure. The imbalanced version of the search tree offers a complexity of $O(\log n + 1)$ close to the balanced search tree. But $n$ is a smaller subset of the whole set of SOC classes containing the wanted SOC class. To benefit from this concept in terms of the average compexity, it is important that the wanted SOC has to be in the lower branch. In this case and if the initial educated guess was sufficiently good the imbalanced tree search is an improvement to the balanced tree search in both efficiency and classification rate. The difficulty lies in obtaining a suitable initial guess.

3.3 Comparison

The two new binary search pattern for the SOC determination via SVM are presented and compared to the already known search pattern from Jansen et al. (2015). For the test trial test data is needed. The measured single impedances of the SOC related spectra are charged with an error. According to Orazem and Tribollet (2008) the most common error in an electrochemical impedance spectroscopy is white noise. To simulate this behaviour on the impedance spectra the single impedances are charged with normally distributed random noise. So for the test trials the test data is generated out of training data by charging with an error $\epsilon$, in both directions ($\Im \eta$ & $\Re \eta$), consisting of normally distributed random noise $X_{(\omega)} \sim \mathcal{N}(\mu, \sigma^2)$ charged with different magnitudes $m$.

$$X : \Omega \rightarrow \Re | X_{(\omega)} = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{\omega - \mu}{\sigma}\right)^2}$$

$$\epsilon = X_{(\omega)} \cdot m, \quad \{ m \in \Re | 10^{-7} \leq m \leq 10^{-4} \}$$

where $\omega$ is a random value, the mean value of $\mu = 0$ and the variance of $\sigma^2 = 1$, where added to each impedance of a impedance spectrum. These error charged impedance objects are the specification for the SOC determination algorithm, to clarify that they would be graded correctly to their origin impedance spectrum.

The Linear search pattern principally analyzes all $n-1$ SVM hyperplanes that separates every SOC class to each other to grade an measured impedance in order to determine the SOC.
the actual SOC of a cell. By a statistical verification through grading a hundred impedances the error prone regarding false classification decreases. Therefore the linear search is able to correct a certain number of false classifications as there is no tree path. As shown in Fig. 7 the classification for low temperatures and high errors is quite around 100 %, but the classification rate decreases with the increase of the temperature and increase of the error as well.

The balanced tree search pattern, see Fig. 8, shows a similar behaviour than the linear search pattern. In low temperature ranges the classification rate is very high even with high errors. The classification rate decreases with the increase of the temperature and with this comes a sensitivity against the increasing error.

The imbalanced tree search pattern on the other hand uses an initial value – called educated guess – for the creation of an imbalanced search tree. With a good initial guess a shorter search tree than for the balanced search tree can be achieved. Compared to the balanced search tree the imbalanced search tree gives the same results with an good initial guess. With an initial guess of each ±15 % as a worst case estimate the classification rate decreases in best cases from around 100 % to around 68 %, see Fig. 9. The reason for that is the initial guess is so bad, that the search will end up in the wrong subsidiary section tree with no possibility to determine the right SOC class.

The reason that the temperature has such a huge effect on the sturdiness of all three pattern against the decrease of the classification rate regarding the increase of the error is the influence of the temperature on the original impedance spectra, see Fig. 10. At low temperatures the impedance spectra in general are bigger and decrease with the increase of the temperature.

4 Conclusion and further work

These investigations on in situ methods based on the support vector machine have shown their effectiveness on determining the state of charge on a lithium iron phosphat cell. Trials with measured impedance spectra have demonstrated that the new concept for grading impedances using SVM is effective independent of the used search pattern. All three search pattern were compared under various aspects like accuracy, efficiency and tolerance of disturbances.
The linear search pattern improves the rate of exact classifications for every temperature. It also improves the robustness against high disturbances and can even detect a certain number of false classifications which makes this search pattern unique. The downside is a much lower efficiency as all hyperplanes have to be evaluated while the tree search pattern only evaluates those on the tree path. The imbalanced tree search pattern on the other hand does not find the correct SOC in many cases with a bad initial guess, but for a wide range of initial values this search pattern gives very small errors compared to the other search pattern. This is especially useful for temperatures where linear search and balanced tree search fail to give acceptable results, for example +40 °C. Figure 9 shows in comparison to the other search pattern good results for the imbalanced tree search method at +40 °C.

The imbalanced tree search pattern shows to be the best and most efficient way to determine the SOC, if the initial guess is within less than ±5 % SOC of the original SOC in both directions and exhibits the best tolerance for high disturbances with almost the same classification rate then the balanced tree search in low temperature ranges and an enhanced classification rate in high temperature ranges. Although showing a higher number of inexact classifications the magnitude of errors is greatly reduced. The initial value for the SOC was simulated with an error of ±15 %. The imbalanced tree search does not tolerate big errors – higher then 40–70 % –, depending on the temperature.

Table 1 compares basic attributes of the three presented search pattern. It shows that each search pattern has different advantages. The search pattern have thus to be chosen according to the specific requirements. The two new search pattern improve many aspects of the SVM-Method but not all. In a future work a combination of search pattern could be discussed to combine their strengths and create an algorithm with both high accuracy and low error magnitude.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Linear search</th>
<th>Binary search</th>
<th>Educated guess search</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complexity (best case)</td>
<td>( O(n - 1) )</td>
<td>( O([\log n]) )</td>
<td>( O([\log n] + 1) )</td>
</tr>
<tr>
<td>Efficiency</td>
<td>low</td>
<td>high</td>
<td>potentially highest</td>
</tr>
<tr>
<td>Classification rate</td>
<td>very high</td>
<td>high</td>
<td>medium</td>
</tr>
<tr>
<td>Error magnitude</td>
<td>medium</td>
<td>medium</td>
<td>potentially lowest</td>
</tr>
</tbody>
</table>

Further novel methods for the state of charge determination independent of the OCV curve.

The presented novel method shows an alternative way to determine the state of charge of a lithium iron phosphate cell based on SOC related impedance spectra out of frequency domain data by grading single measured impedances. For further improvements the more dynamic behaviour of the impedance spectra in terms of superposed direct current on electrochemical impedance spectroscopy has to be investigated. With the results from this examinations an in operando application based on an online impedance measurement as proposed by Klotz et al. (2011) becomes realistic and is worth further research on this topic.

This investigations on the SOC determination based on frequency domain data shall enable the path to novel hybrid algorithms combining the advantages of both worlds the frequency and the time domain. For instance the high dynamic of the time domain based methods combined with the high accuracy in the middle SOC range of the frequency domain based methods.

Acknowledgements. This contribution was developed within the scope of the project Drive Battery 2015 (Intelligente Steuerungs- und Verschaltungskonzepte für modulare Elektrofahrzeug-Batteriesysteme zur Steigerung der Effizienz und Sicherheit sowie zur Senkung der Systemkosten – AEV-Subproject: Optimierung des Energiemanagements von Fahrzeugen mit Lithium-Ionen Starter- und Bordnetzbatterien) which is funded by the BMWi (Bundesministerium für Wirtschaft und Energie) under the grant number 03 ET6003 I. The responsibility for this publication is held by the authors only.

Edited by: J. Anders
Reviewed by: two anonymous referees

References


