

充電データ列のシミュレーションに向けた 電圧値の解析と変異の生成

中里佳樹[†] 関澤俊弦^{††}

本研究は、充電データに変異データを含ませ、変異を含む充電データをシミュレートすることを目的とする。そのために、電圧上昇量を用いて正常値を再現し、クロスエントロピーを利用して変異データを生成する。クロスエントロピーの値を用いることで、変異データの分散の大きさを表現することを可能にする。変異を含むデータ列を生成することで、変異データを検出する分類モデルの発展に繋げる。

Towards charging simulation by analysis voltage value and generation mutation

Yoshiki Nakasato[†] and Toshifusa Sekizawa^{††}

The purpose of this study is to include and simulate mutation data in the charging data. Therefore, the voltage rise is used to reproduce normal values, and cross-entropy is used to generate mutation data. Cross-entropy allows for a representation of the variance of the mutation data. Generating a data sequence containing mutations will lead to the development of a classification model to detect mutation data for future research.

1. はじめに

近年、電気自動車の普及が進み、CO2削減に力を入れている。電気自動車の普及と共に、充電装置の利用が必要である。充電装置の利用に伴い、装置内の情報を書き換える妨害が考えられる。妨害などにより電気自動車の充電が正常に行われなければ、大きな人的被害を及ぼす可能性がある。妨害による被害を抑えるため、妨害を自動的に検知する必要がある。自動的に検知するために、正常データと異常なデータを分類する分類モデルの構築を目指す。分類モデルでは、正常であるデータ列と異常を含むデータ列を認識し、書き換えによる異常な値を自動的に検出することを考える。しかし、分類モデルを構築するには、様々なデータ列で分類を試行し、正常なデータと異常なデータを分類する必要がある。本研究は、分類モデルを構築する上で必要な異常な値を含むデータ列の生成を目標としている。

本研究では、充電装置の出力に妨害が施された状況を想定し、自動的に変異を含むデータ列を生成することを考える。分類モデルの構築には、変異を含むデータ列を用いて試行する必要がある。そのため、変異を含むデータの生成をシミュレートできる環境を作ることで、分類モデルの構築に役立てると考えられるためである。変異の生成は、クロスエントロピーの式を用いてデータ列を生成する方法である。

変異を生成する手法としては、水位の変異データを生成する 1) がある。この文献では、水位をランダムに変化させることにより変異データを生成している。しかし、充電データを完全ランダムに変化させる変異生成は、有効的ではないとする。そのため、本研究では、充電データに対する変異生成のアルゴリズムを構築することを目的とする。また、本研究の初めの段階として、最尤法を用いた正常値の再現を行い、再現が不安定であるという問題点を解決するべく、充電電圧値の解析結果を基に正常値の再現を行う。

以下、第2章では、充電装置の構成と動作について述べる。第3章では、充電装置によって充電されたデータの特徴を述べる。第4章では、充電データ列の電圧上昇量を用いたデータの再現について述べる。第5章では、クロスエントロピーを用いた変異の生成方法について述べる。第6章では、電圧上昇量を用いた再現データにクロスエントロピーの変異生成を取り入れた結果を述べる。第7章では、再現したデータと変異生成の考察を行う。第8章では、本研究のまとめを述べる。

[†] 日本大学大学院工学研究科情報工学専攻
Computer Engineering, Graduate School of Engineering, Nihon University

^{††} 日本大学工学部情報工学科
Department of Computer Science, College of Engineering, Nihon University

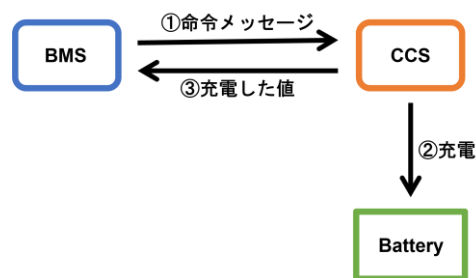


図1 装置の構成

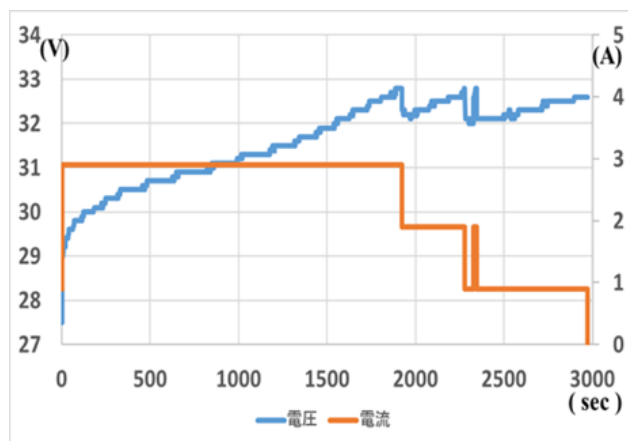


図2 取得した充電データ列

2. 充電装置

充電装置は大きく、Battery Management System（以下BMS）、Charger Control System（以下CCS）とバッテリーの3つで構成される（文献1）。図1に、装置の構成を示す。図中の①,②,③は、動作の順番を表し、次のとおり動作する。動作①では、BMSがバッテリーへ充電する電流値と電圧値を決定し、CCSへ命令メッセージとして送信する。動作②では、CCSがBMSから受信した値を元にバッテリーの充電を行う。動作③は、CCSが実際に充電した電流値と電圧値をBMSへ返す。本論では、図1の動作③でCCSがBMSへ返した値を扱う。解析に利用するデータ列は、テキサス大学サンアントニオ校で収集された充電データセットである。充電データセットは、BMSと

CCS間のメッセージのやり取りが記録されており、1系列のデータ列に0秒から約3000秒までの電圧値、電流値のデータ列が存在する。本研究のデータは、収集されたデータのうち9系列を対象とし、いずれも正常な充電データ列である。

3. 充電データの特徴

充電装置によって行われた充電データ列には、以下の特徴を持つ。充電中の電流値は一定を保ち、充電中の電圧値はBMSによって定められた上限値まで上昇する。電圧値は上昇量が0であるときが多い。電流値が一定であるときは、電圧の変動量は $\pm 0.1V$ であり、少しずつ上昇する特徴を持つ。電圧変動量が0.1であるときと-0.1であるときの割合は以下のとおりである。変動量0.1の割合は、電流値が3A時89%、2A時73%、1A時63%である。また、変動量-0.1の割合は、3A時11%、2A時27%、1A時37%である。電圧値が一定の電圧上限値に到達したとき、電流値が一定量下がる仕組みを取っている。そのため、電流値が下がる時、同時に電圧値も少し下がる。図2は、9系列の充電データ中の1つである。横軸が時間であり、左縦軸が電圧値、右縦軸が電流値を示す。図中の青のデータが電圧データであり、橙のデータが電流データである。

4. 電圧上昇量を用いたデータの再現

本章では、電圧上昇量を用いたデータの再現手法を提案する。第1節では、生データを基に電圧値の上昇量の解析方法について述べる。第2節では、解析した上昇量を基に再現データの出力について述べる。

4.1 データから電圧値の解析

正常な充電データ列9系列の電圧上昇量に基づいて正常値を再現する。時間 t 秒における電圧値を V_t とすると、電圧の上昇量 ΔV_t は次のように与えられる。

$$\Delta V_t = V_{t+1} - V_t \quad (1)$$

充電には、電圧上昇量 ΔV_t が0を継続している時間と微小に変動している箇所が存在する。電圧上昇量 ΔV_t が0を継続している時間を利用し、データの再現を行う。電圧上昇量 ΔV_t が0を継続している時間を CT と表し、正常データ9系列から CT を求める。継続時間 CT は一律ではなく、小さい値と大きい値がはっきりと分かれる。そのため、継続時間 CT は、大小2つに分ける。継続時間 CT の大小を分ける値は $CT=40$ とする。継続時間 CT が40未満であるものを小継続値 CT_{small} とし、継続時間 CT が40以上であるものを大継続値 CT_{large} とする。また、継続時間 CT が40未満である CT_{small} が続く回数を小継続回数 nCT_{small} とし、大継続値 CT_{large} が続く回数を大継続回数 nCT_{large} と表す。継続値 CT_{small} 、 CT_{large} と継続回数

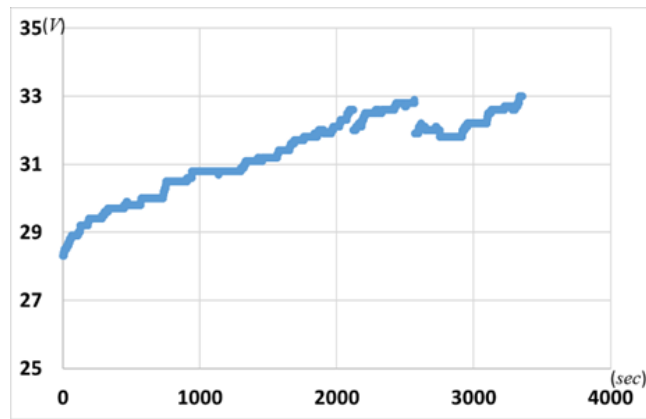


図 3 再現データ列

nCT_small , nCT_large をそれぞれ各生データ 9 系列から算出し、継続値と継続回数を基にそれぞれの標準偏差を求める。各々の標準偏差を基に正規乱数を用いて新たな継続値と継続回数を求める。正規乱数によって求められた小継続値を CT_small_new , 大継続値を CT_large_new , 小継続回数を nCT_small_new , 大継続回数を nCT_large_new とする。

4.2 電圧データの再現

前節で求められた新たな継続値と継続回数を基に電圧値を再現する。算出された継続値と継続回数から電圧上昇量 0 が継続する時間を求める。求められた電圧上昇量を充電開始電圧値に上昇量を加算させ、データを再現する。

算出された小継続回数 nCT_small_new の値の分だけ小継続値 CT_small_new を並べ、次に算出された大継続回数 nCT_large_new の値の分だけ大継続値 CT_large_new を並べ、値が無くなるまで小と大の値を並べを繰り返す。並べられた数列を新たな継続値 CT_new とし、継続値 CT_new を基に新たな電圧上昇量 ΔV_t_new を求める。

継続数 CT_new の値の分、電圧上昇量 0 を並べる。上昇量 0 を並べた後、変動量 ± 0.1 が挿入される。変動量 ± 0.1 が挿入される割合は、生データから統計を取り、挿入する。

上昇量 ΔV_t_new が算出後、充電開始電圧値に上昇量を加算させ、データ列を再現する。充電開始電圧値は、生データ 9 系列を基に、算出される。図 3 は ΔV_t_new を基に生成された正常データの再現である。データが途切れている部分は、電圧値が変更された部分に対応している。

5. クロスエントロピーを利用した変異生成

変異を生成する手法の 1 つとして、クロスエントロピー (2) を用いて偽の確率分布を生成し、変異を生成する。また、クロスエントロピーによって求められた割合を用いて変異を生成する。クロスエントロピーは、真の確率分布と推定する確率分布がどの程度似ているかを表す値であり、値が小さいほど 2 つの確率分布が似ていることを表す。正解データの分布を $p(Vt_i)$, 推定する確率分布を $q(Vt_i)$, エントロピーの値を E としたとき、クロスエントロピー E は次のとおり与えられる。

$$E = - \sum_{i=1}^n p(Vt_i) \log(q(Vt_i)) \quad (2)$$

式 (2) を利用し、真の確率分布に対し大きなエントロピーの値を与えることで、真の確率分布と相違する偽の確率分布を生成する。

现阶段での偽の確率分布は、 $\pm 3\sigma$ の値を 100% としたとき、与えられたエントロピーから算出された割合より外側に確率を持つ分布である。エントロピーから算出される割合を z とすると、式は次で与えられる。

$$z = \pm 100(1 - e^{-E}) \quad (3)$$

割合 z を用いて、変異の生成範囲を求める。正規分布の中央である平均値 μ を 0% としたとき、 3σ の値を 100%, -3σ の値を -100% と表し、 $\pm z\%$ として表す。変異の生成範囲は、割合 -100% から $-z\%$ と $z\%$ から 100% の範囲に一定の確率で変異データを生成する。クロスエントロピーの値 $E=2$ を例としたとき、割合 $z \approx \pm 86.4$ となり、おおよそ $\pm 2.6\sigma$ の位置を指す。これにより、変異の生成範囲は -3σ から -2.6σ , または 2.6σ から 3σ の範囲に生成する。

6. クロスエントロピーの適用例

データ 9 系列を基に再現されたデータ列に偽の確率分布を基に生成した変異データを混入させ、変異を含むデータ列を生成する。変異を含むデータ列を生成するには、偽の確率分布から生成されたデータを再現した電圧データに含ませる確率を設定する必要がある。本研究では、偽の確率分布によって生成されたデータを再現した電圧データに含む確率を 10% とする。確率を 10% とした理由は、異常データが少量であるとき (数%未満のとき) ノイズとして装置が判定する可能性が高いためである。図 4, 図 5 はクロスエントロピー $E=0.1, 2$ として、10% の確率で変異を含ませたデータ列である。図 4, 図 5 はそれぞれ青のデータが正常値を表し、橙のデータが変異データを表す。図 4 は、 $E=0.1$ を与えたときの変異である。クロスエントロピー $E=0.1$ のときの割合 z は ± 0.95 であり、 $\pm 0.29\sigma$ の位置を指す。図 5 は、 $E=2$ を与えたときの変異デ

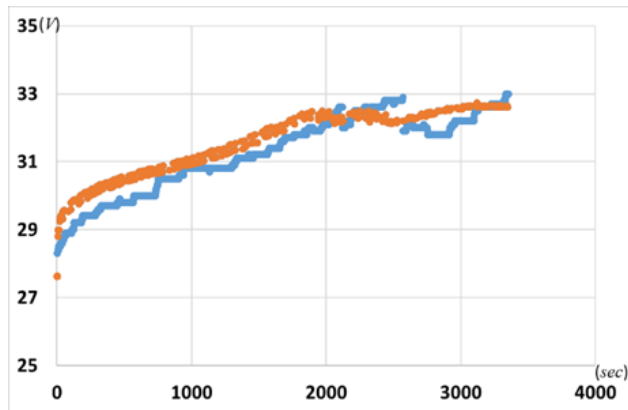


図 4 $E=0.1$ の変異を与えたデータ列

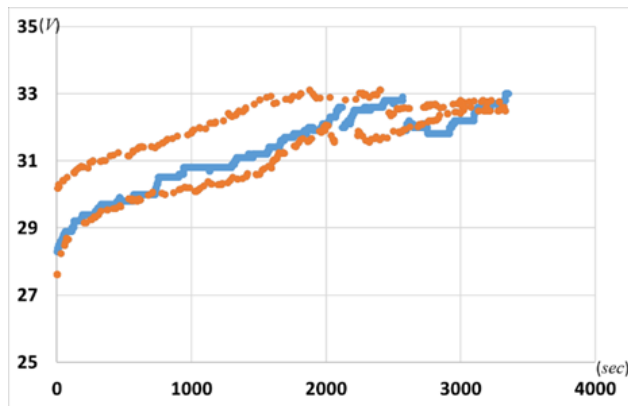


図 5 $E=2$ の変異を与えたデータ列

データ列である。クロスエントロピー $E=2$ のときの割合 z は ± 86.4 であり、 $\pm 2.6\sigma$ の位置を指す。

7. 考察

本研究の正常データの再現は、電圧上昇量 0 の継続時間 CT を用いて正常値の再現を行う方法である。また、変異の生成ではクロスエントロピーを用いて偽の確率分布を生成する。継続時間 CT を取り入れた正常値の再現は、電流値の変動によって生じ

る電圧値の変動の再現を可能にする。しかし、本研究で対象とした装置固有の特性に基づいている可能性があるため、他の装置の特性で対応するかどうかの調査が必要である。

クロスエントロピーを取り入れた変異の生成については、クロスエントロピーの値を変えることにより、変異データの分散の大小を表現することを可能にした。クロスエントロピーの値 $E=0.1$ を与えたときは、変異データの分散が小さく、 $E=2$ を与えたときは、変異データの分散が大きく表れた。クロスエントロピーの値の決定方法や、決定した値が分類モデルに与える影響について更なる研究が必要である。

8. おわりに

本研究は、充電データをシミュレートし、様々なデータで分類を試行し、分類モデルを構築することが目標である。変異を含むデータ列の生成には、電圧上昇量とクロスエントロピーを利用し、変異を含むデータ列の生成を考えた。正常値の再現を他の構成を用いて調査することや、変異の確率を 10% から変えていくこと、クロスエントロピーの与える値の決め方について更なる研究が必要である。

参考文献

- 1) Yuqi Chen, Christopher M. Poskitt, and Jun Sun.: Learning from Mutants: Using Code Mutation to Learn and Monitor Invariants of a Cyber-Physical System, IEEE Symposium on Security and Privacy (2018)
- 2) 栗田多喜夫, 日高章理:統計的パターン認識と判別分析, コロナ社(2019)