

# 送電系統向け類似事例検索手法の提案と評価

西川 記史<sup>†</sup> 高田 実佳<sup>††</sup> 山崎 潤<sup>††</sup>

<sup>†</sup>(株)日立製作所 〒244-0817 神奈川県横浜市戸塚区吉田町 292 番地

<sup>††</sup>Hitachi America, Ltd. 3315 Scott Blvd. 4th Floor, Santa Clara CA 95054-3103

E-mail: <sup>†</sup>norifumi.nishikawa.mn@hitachi.com, <sup>††</sup>{mika.takata,jun.yamazaki}@hal.hitachi.com

あらまし 電力データの蓄積・管理のために大規模なデータベースが用いられている。送電系統で事故が発生した場合に、データベースから過去の類似した事例を検索し、それを参考に事故の対策を検討することはシステムの安定化に有用である。しかし、検索を高速化するためには特徴量を事前に抽出し索引付を行う必要があり、データベース運用の複雑化や検索時の柔軟性の低下を招いていた。本研究では、類似事例の検索処理の一部を DBMS 内部に push-down することにより、事前の特徴量抽出を不要とした対話的な類似事例検索手法を提案する。電力システムシミュレータにより生成したデータを用いて提案手法を評価した結果、過去 2 年の PMU データから類似事例を 1 分以内で検索できる見込みを得た。これにより、送電系統運用において、データベース運用を単純化でき、多様な特徴量による過去類似事例検索を可能とする。さらに、新しい知見をアプリケーションにして組込み易くすることが可能となる。

キーワード 送電系統, 類似事例検索, 類似波形検索, プローニー解析

## 1. はじめに

諸外国では大規模な停電が深刻な問題となっている [1]。2003 年 8 月に米国及びカナダで発生した大停電以降、大規模停電の回避は送電システムの安定化における最重要課題の一つである。また、大規模停電は、経済的にも大きな損失をもたらす。Electricity Consumers Resource Council により報告された 2004 年に発生した大規模停電のレポートでは、経済的損失は 100 億ドルにも達したことが報告されている [2]。

大規模停電が発生する理由の一つに、送電系統の状況理解・把握 (Situational Awareness) が不十分であったことが挙げられている [3]。このため、送電系統の制御システムでは、系統を精緻に監視しオペレータの Situational Awareness を改善する WAMS (Wide Area Measurement System) の導入や、近年では監視に加え系統の保護・制御をも行う WAMPAC (Wide Area Monitoring, Protection and Control) の導入計画が進んでいる [4,5]。

WAMS や WAMPAC では、非常に細粒度の情報の収集が可能な位相計測装置 (Phasor Measurement Unit; PMU) [6,7] と呼ばれる新たなセンサーが広く活用されている [8-10]。PMU は、電圧、電流、周波数の大きさと位相差を毎秒 30 回から 120 回の周期で計測するセンサーである [11]。PMU は、従来から使われていた SCADA (Supervisory Control And Data Acquisitions) と比較して数百倍詳細な情報を収集することが可能であり、電力システムにおける時系列データ分析の可能性を広げるものであると考えられる。

大規模停電を防ぐためには、事故発生から数分間の初期段階での的確な事故対策が重要である。送電系統で発生した事故に的確に対応するために、ワーキンググループでの過去の事故事例の共有が活発に行われており [25]、今後事故事例を活用したオペレーション支援が重要になると考えられる。このため我々は、事故が発生した場合に類似した過去事例を検索し、それに

関連付けられた操作履歴を提示することで事故発生時のオペレータの対策を支援するシステムの研究を進めている [12-16]。

大規模な PMU データベースから対話的に過去の類似事例を検索するには、検索を高速化するための特徴量を事前に抽出し蓄積する手法が考えられる。このためにはデータを取り込む際に同時に特徴量を抽出する、あるいは蓄積された PMU データを対象にオフラインで特徴量を計算し、PMU データとは別に蓄積する必要がある。しかし、これらはデータベース運用の複雑化や特徴量を事前に定義しておくことによる類似事例検索の柔軟性の低さを引き起こす。

本論文では、事前の特徴量抽出を不要とし、データベース運用の単純化と検索時の柔軟性の向上を目指した類似事例検索手法を提案する。提案手法は、類似事例検索処理のうち、特に時間がかかるデータベースからの事故データの抽出処理を DBMS 内部に push-down するとともに、抽出した事故データの特徴量抽出及び類似度計算を並列に行う。

電力システムシミュレータ [17] により生成したデータを用いて提案手法を評価した結果、2 年分の PMU データを蓄積したデータベースから、十数秒で類似した事故データを検索できる見込みを得た。これにより、送電系統運用において、データベース運用を単純化すると共に、多様な特徴量による過去類似事例検索を可能とする。

## 2. 関連研究

これまで、大量のセンサデータを収集・蓄積し、それらの分析を支援するための IT プラットフォームが多数提案されている [18,19]。文献 [18] では、大量のセンサデータを短時間で蓄積するため時系列データを圧縮して蓄積する機構、及び SQL インタフェースを通して時系列データベースにアクセスするための機構を提案している。また、文献 [19] では、時系列データに対する抽象的なアクセス方法を定義し、データを時間ごとに区

切り tree 形式で記憶装置に格納するとともにデータを圧縮することにより、時間指定での時系列データの取出し時間を短縮している。これらの方式は、センサ ID と時刻をキーとしてデータを検索する場合は高い性能を発揮するが、現在発生している事故の時系列データと類似した時系列データを検索することには不向きである。

また、大規模な時系列データベースから類似事例を検索するための手法も多数提案されている [20–24]。文献 [20, 21] は時系列データの次元を削減した索引を作成することで類似時系列データの検索を高速化している。文献 [22] は Dynamic Time Warping [23] の高速化手法を提案している。文献 [24] は、ビットマップ索引を用いて PMU データファイルの索引付を行い、時刻や電圧値による PMU データファイルの検索を高速化している。これらの手法はいずれも、時系列データ全体を検索対象としている。

### 3. 類似事故の時系列データ検索方式

データベースから事故時の時系列データと類似した時系列データを求めるためには、事前に特徴量を抽出しておくか、特徴量を事前に抽出しない場合はデータベースから全データを抽出し、その後に現在発生している事故の時系列データと抽出したデータを比較する必要がある。しかし、前者の方式はデータベース運用を複雑にする可能性が高く、後者の方式ではデータを全てデータベースから取り出す処理がボトルネックとなりやすく、非常に時間がかかる。また、過去の事故が多数ある場合にそれらを現在の事故の時系列データと逐次的に比較していたのでは時間がかかり対話的な事例検索に間に合わない可能性がある。

そこで我々は、送電系統の時系列データの特性に着目し、過去の事故の時系列データの抽出を DBMS 内部で実施することによりデータ転送ボトルネックを排除するとともに、過去の事故の時系列データと現在発生している事故の時系列データの比較を並列に行うことにより対話的な事故事例検索を可能とする類似事故の時系列データの検索方式を提案する。

#### 3.1 送電系統データの特性

まず、送電系統データの特性について述べる。送電系統の監視や制御のために使用される PMU は、電圧や電流とそれらの位相、周波数など、様々な情報を収集する。このうち、電圧と周波数は定格値が決まっているため、事故が発生した時のみ大きく変動し、それ以外はほぼ一定の値を取る。逆に、電流値はその時々々の負荷によって値が大きく変化する。また、位相もゆっくりと変化するため様々な値を取る。

図 1 に送電系統シミュレータ [17] で作成した電圧値の例を示す。図 1 は 2 分間、8 PMU 分のデータであり、先頭から 60 秒後に系統事故が発生している。図から分かるように、事故が起きるまで電圧値は定格値の近辺で一定値であり、事故時のみ大きく変動することが分かる。

#### 3.2 時系列データ検索方式

次に、我々が提案する類似事故の時系列データの検索方式について説明する (図 2)。

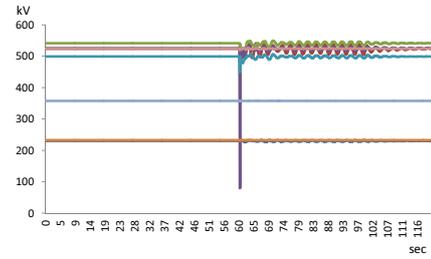


図 1 Voltage Magnitude of PMU data.

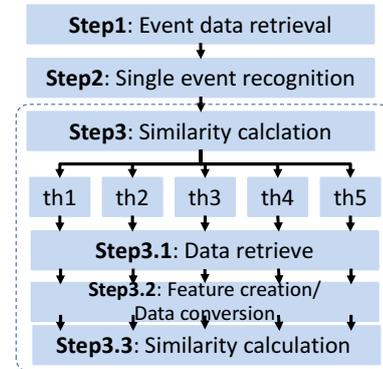


図 2 Outline of Proposed Similarity Search Method.

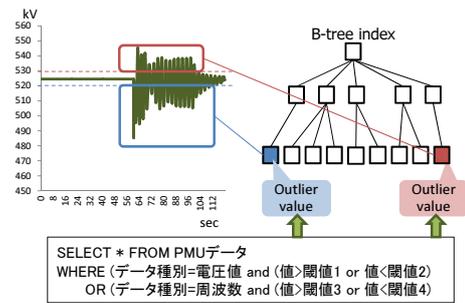


図 3 Event Data Retrieval Method using Outlier Search.

#### 3.2.1 事故時の時系列データ抽出 (Step 1)

前節に示したように、PMU が収集するデータの一部には定格値があること、及び事故が発生した場合のみデータの値が定格値から外れる。この特性を利用することで、DBMS がもともと有している B-tree 索引あるいはビットマップ索引を用いて事故時の時系列データを高速に取得することが可能となる。

図 3 に、B-tree 索引を用いた事故時の時系列データの抽出方式を示す。提案方式では、PMU が計測するデータの値に対して B-tree 索引を定義する。そして、事故時の時系列データ検索時に、PMU が収集するデータのうち、定格値を持つデータである電圧値 (Voltage Magnitude) 及び周波数 (Frequency) について、値が定格値から大きく外れたものを B-tree 索引を用いて検索する。これにより、全データをスキャンすることなく事故時のデータを高速に抽出することが可能となる。なお、我々の実装ではデータベースに關係データベースを用いている。關係データベースのスキーマは文献 [13] に示したものをを用いている。

#### 3.2.2 単一イベント認識 (Step 2)

図 3 から分かるように、単純に B-tree 索引を用いて値が定格

値から大きく外れたデータを抽出した場合、PMU データでは一つの事故に対して、多数の値が抽出されることになる。時系列データから類似事例を検索する場合、これら多数の値を一つの事故としてカウントする必要がある。通常、系統事故は多くても月に数回程度であると考えられるため、時刻印が近いデータを一つの事故としてまとめる。

その後、一つの事故として認識されたデータのうち最も時刻が若いものを、事故の発生時刻とする。これにより比較しなければならない事故の件数を削減し、処理の高速化を図る。

### 3.2.3 類似度計算 (Step 3)

最後に、前節において識別された過去の事故の時系列データと、現在発生している事故の時系列データより特徴量を抽出し、その類似度を計算する。通常、データベース中には複数の事故のデータが格納されている。事故データの特徴量抽出及び類似度計算を逐次的に実施すると、事故の数が増えるに従い類似度の計算に時間を要するようになる。そこで、類似度判定 (Step 3) では、複数のスレッドを用意し、各事故データ毎に並列に特徴量抽出及び比較を行う。これにより類似度計算時間を短縮する。類似度判定 (Step 3) は、大きく3つのステップからなる。

**Step 3.1 データ検索** 各事故の開始時刻を元に、その前後の電圧値、周波数、電力値の時系列データをデータベースより取得する。

**Step 3.2 特徴量抽出** 時系列データの比較に特徴量を用いる場合、Step 3.1 で取得した時系列データの特徴量を on-the-fly で計算する。特徴量の抽出には、送電系統の安定度の判定に広く用いられる Prony 解析手法 [26] を用いた。

**Step 3.3 類似度計算** Step 3.2 で求めた特徴量と、現在発生している事故の時系列データの特徴量の距離を計算する。

最後に、事例を距離の小さい順にソートし類似検索結果とする。

時系列データの類似度の計算には、以下の式を用いた。

$$\text{similarity} = \sqrt{w_V \Delta V^2 + w_F \Delta F^2 + w_P \Delta P^2} + \text{Pro}/n - 1$$

$$\text{Pro} = w_f \sum_{k=1}^n \Delta f_{\text{mode}.i}^2 + w_d \sum_{k=1}^n \Delta d_{\text{mode}.i}^2$$

$w_V, w_F, w_P$  はそれぞれ電圧、周波数、電力に対する重み、 $\Delta V, \Delta F, \Delta P$  はそれぞれ電圧、周波数及び電力の、現在の事故のデータ値と過去の事故のデータ値の差分である。 $w_f, w_d$  はそれぞれ動揺周波数及び減衰率に対する重み、 $\Delta f, \Delta d$  はそれぞれ動揺周波数及び減衰率の、現在の事故のデータ値と過去の事故のデータ値の差分である。 $n$  は類似検索で使用するモード数である。

## 4. 評価

我々は、提案手法の有効性を確認するために、系統シミュレータ [17] により生成した事故データを用いて、提案手法の評価を行った。

### 4.1 評価方法

評価では、前章で述べた提案手法、及び従来手法として事前に全データの特徴量を抽出しデータベースに挿入するとともに事故時点をタグ付する手法について、特徴量のデータベース容

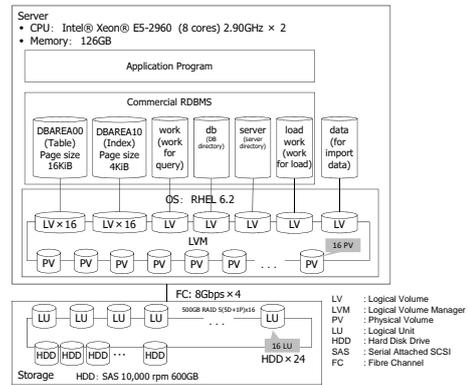


図4 System Configuration.

量、特徴量データの準備時間、及び検索性能を比較した。データ準備時間及び検索時間の評価においては、3回検索性能を測定し、その平均値を求めた。なお、提案手法の Step 3 の並列度数は論理コア数と同じ 32 とした。

### 4.2 評価環境

サーバ及びストレージ、ソフトウェア構成は図4に示すとおりである。ストレージは、HDD 24 台から 16 個の LU を作成した。サーバ側はこれら 16 個の LU に対応したデバイスを用いて LVM を構築し、そこから DB 用の LV を切り出している。DBMS ソフトウェアとして日立製作所の Hitachi Advanced Data Binder<sup>(注1)</sup>を用いた。

データベースには、8 PMU、3ヶ月分のデータを格納した。事故の件数は7件、データベース容量は637.3GBである。

### 4.3 評価結果

評価の結果、提案方式を用いた場合の特徴量のデータベース容量は0バイト、特徴量データの準備時間0秒、検索時間の平均値は約6.3秒であるのに対し、従来方式の特徴量のデータベース容量は約66MB、特徴量データの準備時間約3,770秒、検索時間の平均値は約0.03秒であった。電力会社は平均的に約150台のPMUデータを2年間保有する。従来方式では検索時間は約0.24秒であるが特徴量のデータベース容量が約9.9GB、特徴量データの準備時間が約157時間かかることになり、新たな知見を組み込むのは困難である。

一方、提案方式は実運用で用いられる約2年分、150 PMUデータを蓄積したデータベースを対象とした場合でも、類似事例の発生頻度が同等の場合は、類似事例を約50秒で検索できる見込みを得た。これはシステムの崩壊が始まるまでの時間より十分短く実運用上問題ない検索時間と考える。

## 5. 考察

提案手法は、計測値に B-tree 索引を定義している。そのため、従来手法と比較してデータのロード性能、及びデータベース容量に影響を与える。そこで我々は、B-tree 索引を用いない場合、

(注1)：内閣府の最先端研究開発支援プログラム「超巨大データベース時代に向けた最高速データベースエンジンの開発と当該エンジンを核とする戦略的サービスの実証・評価」(中心研究者：喜連川 東大教授/国立情報学研究所所長)の成果を利用。

及び用いた場合について、データロード性能及びデータベース容量を比較した。

(1) データロード性能 B-tree 索引を用いない場合のデータロード性能は毎秒約 224 万件、B-tree 索引を定義した場合のデータロード性能は毎秒約 167 万件であった。

(2) データベース容量 B-tree 索引の容量は、データベース全体の容量の約 5%であった。

これらの結果から、提案手法によりデータのロード性能は約 25%低下するが、PMU データのロードに求められる性能は十分満たしていると考えられる。また、データベース容量への影響は軽微であると考えられる。

## 6. ま と め

本論文では、大規模な PMU データベースから、時系列データの類似度を用いて、事前の特徴量抽出等のデータ準備処理を不要としつつ類似事件事例を短時間で検索できる手法の提案、及び評価を行った。評価の結果、電力系統シミュレータにより生成したデータを用いて提案手法を評価した結果、実運用で用いられる約 2 年分の 150 PMU データを蓄積したデータベースを対象とした場合でも類似事件事例を約 50 秒で検索できる見込みを得た。これにより、送電系統運用において、データベース運用を単純化でき、多様な特徴量による過去類似事件事例検索を可能とする。さらに、新しい知見をアプリケーションにして組込み易くすることが可能となる。

## 7. 謝 辞

本研究の評価の一部は、国立研究開発法人新エネルギー・産業技術総合開発機構 (NEDO) の委託業務によるものである。

## 文 献

- [1] 電気学会 電気広報特別委員会, 過去の大規模停電事例, <http://www2.iee.or.jp/ver2/honbu/16-committee/epress/data/12-jirei.pdf>, 2011.
- [2] Zhaoyang Dong and Pei Zhang, *Emerging Techniques in Power System Analysis*, Springer, 2009.
- [3] U.S. Department of Energy, *Synchrophasor Technologies and their Deployment in the Recovery Act Smart Grid Programs*, <http://energy.gov/sites/prod/files/2013/08/f2/SynchrophasorRptAug2013.pdf>, 2013.
- [4] Michael Weixelbraun, *Implementation of a Wide Area Monitoring System (WAMS) for Austria's Power Grid*, International Synchrophasor Symposium, 2016.
- [5] Panat Sawatpipat, Witchaya Pimjaipong and Sompol Chumnvanichkul, *Wide Area Monitoring System: The Strategy, Planning and Implementation in Thailand Power Grid*, The 21st Conference of the Electric Power Supply Industry, 2016.
- [6] Arun Phadke, *Synchronized phasor measurements in power system*, IEEE Computer Applications in Power, volume 6, number 2, pp.10–15, 1993.
- [7] Aleksandar Radovanovic, *Using the internet networking of synchronized phasor measurement units*, International Journal of Electrical Power & Energy System, volume 23, number 3, pp.245–250, 2001.
- [8] Jaime De La Ree, Virgilio Centeno and James S. Thorp, *Synchronized Phasor Measurement Applications in Power Systems*, IEEE Transactions on Smart Grid, volume 1, number 1, pp.20–27, 2010.
- [9] Reynaldo Nuqui and Arun Phadke, *Phasor measurement unit placement techniques for complete and incomplete observability*, IEEE

- Transactions on Power Delivery, volume 20, number 4, pp.2381–2388, 2005.
- [10] Gerald Heydt, Chen-Ching Liu, Arun Phadke and Vijay Vittal, *Solution for the crisis in electric power supply*, IEEE Computer Applications in Power, pp.22–30, 2001.
- [11] C37.118.2-2011 - IEEE Standard for Synchrophasor Data Transfer for Power System, IEEE-SA Standard Board, 2011.
- [12] 高田 実佳, 西川 記史, 清水 晃, 茂木 和彦, 藤原 真二, 河原 大一郎, *送電系統安定化のためのデータ分析プラットフォームにおけるアーカイブアクセス技術と評価*, 第 7 回データ工学と情報マネジメントに関するフォーラム (第 13 回日本データベース学会年次大会) E7-1, 2015.
- [13] 西川 記史, 高田 実佳, 清水 晃, 茂木 和彦, 藤原 真二, 河原 大一郎, *送電系統安定化のためのデータ分析プラットフォームにおける蓄積系データ処理技術と評価*, 第 7 回データ工学と情報マネジメントに関するフォーラム (第 13 回日本データベース学会年次大会) E7-2, 2015.
- [14] Mika Takata, Yasushi Miyata and Norifumi Nishikawa, *Proposal of analytics software architecture with data preparation layer for fast event identification in wide-area situational awareness*, 2015 IEEE International Conference on Industrial Technology (ICIT), 2015.
- [15] Mika Takata, Kazutoshi Tsuchiya, Kenta Kirihara, Nao Saito, Yosuke Ishii, Norifumi Nishikawa, Jun Yamazaki, Yutaka Kokai and Brett Amidan, *Decision Support System: real-time event detection and historical event discovery application for operators*, International Synchrophasor Symposium, 2016.
- [16] Norifumi Nishikawa, Jun Yamazaki, Mika Takata and Yutaka Kokai, *A High-Performance Platform for Real-Time Data Processing and Extreme-Scale Heterogeneous Data Management*, i-PCGRID Workshop 2016, 2016.
- [17] PSLF, <http://www.geenergyconsulting.com/practice-area/software-products/pslf>.
- [18] Sheng Huang, Yaoliang Chen, Xiaoyan Chen, Kai Liu, Xiaomin Xu, Chen Wang, Kevin Brown, and Inge Halilovic, *The Next Generation Operational Data Historian for IoT Based on Informix*, In Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data, pp. 169–176, 2014.
- [19] Michael P Andersen and David E. Culler, *BTrDB: Optimizing Storage System Design for Timeseries Processing*, 14th USENIX Conference on File and Storage Technologies (FAST '16), pp.39–52, 2016.
- [20] Eamonn J. Keogh and Michael J. Pazzani, *An Indexing Scheme for Fast Similarity Search in Large Time Series Databases*, Eleventh International Conference on Scientific and Statistical Database Management, 1999.
- [21] Eamonn Keogh, Kaushik Chakrabarti, Michael Pazzani and Sharad Mehrotra, *Dimensionality Reduction for Fast Similarity Search in Large Time Series Databases*, Knowledge and Information Systems Journal (KAIS), 2001.
- [22] Thanawin Rakthanmanon, Bilson Campana, Abdullah Mueen, Gustavo Batista, Brandon Westover, Qiang Zhu, Jesin Zakaria and Eamonn Keogh, *Searching and Mining Trillions of Time Series Subsequences under Dynamic Time Warping*, Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining, pp.262–270, 2012.
- [23] Hui Ding, Goce Trajcevski, Peter Scheuermann, Xiaoyue Wang and Eamonn Keogh, *Querying and mining of time series data: experimental comparison of representations and distance measures*, Proceedings of the VLDB Endowment, Volume 1 Issue 2, pp.1542–1552, 2008.
- [24] Ben McCamish, David Chui, Miles Hestand, Jordan Landford, Robert B. Bass, Rich Meier and Edurado Cotilla-Sanchez, *Managing PMU Data Sets with Bitmap Indexes*, 2014 IEEE Conference on Technologies for Sustainability (SusTech), 2014.
- [25] North American SynchroPhasor Initiative Work Group Meeting, <https://www.naspi.org/meetings>
- [26] 天野 雅彦, 世古口 雅宏, 禰里 勝義, 河田 謙一, *ブローニー解析法による電力系統実測データの動揺モード検出*, 電学論 B, 120 巻 2 号, 2000.