

METODOLOGÍA BASADA EN EL ANÁLISIS DE DATOS TOPOLÓGICOS PARA LA DETECCIÓN DE OSCILACIONES CON DATOS DE BAJO MUESTREO

Victor Manuel Meza Jimenez, Brian Camilo Perez Romero

Dirección Aseguramiento Operación

XM S.A. E.S.P.

Medellín, Colombia

vmmeza@xm.com.co; bperez@xm.com.co

+57 604 3172929

(Resumen/Abstract) — Oscillation detection in power system is keystone to ensure reliable and secure power delivery. However, most of the existing approaches have been proposed for synchro-phasor measurements, hence unpractical for those countries with non or low deployment of PMUs. This paper introduces a novel topology data analysis-based approach for oscillation detection using low-sampled data from SCADA. This approach operates by extracting some valuable features for oscillation detection classification. The proposed method was tested on real-world data from the Colombian power system, proving its suitability for online applications.

La detección de oscilaciones en sistemas eléctricos de potencia es fundamental para garantizar un suministro de energía confiable y seguro. Sin embargo, la mayoría de los enfoques existentes se han propuesto para mediciones de sincrofasores, lo que los hace poco prácticos para aquellos países con un despliegue nulo o bajo de PMUs. Este artículo presenta un nuevo enfoque basado en el análisis de topología de datos para la detección de oscilaciones utilizando datos muestreados a baja frecuencia desde SCADA. Este enfoque opera extrayendo algunas características valiosas para la clasificación de detección de oscilaciones. El método propuesto fue probado en datos reales del sistema eléctrico colombiano, demostrando su idoneidad para aplicaciones en línea.

Palabras clave: *Análisis topológico de datos, detección de oscilaciones, homología persistente, oscilaciones en sistemas de potencia.*

I. INTRODUCCIÓN

El creciente número de recursos energéticos verdes variables, ya sea basados en inversores o hidroeléctricos a pequeña escala, está cambiando las dinámicas de los sistemas eléctricos e imponiendo nuevos retos en la operación de los sistemas de potencia, especialmente aquellos relacionados a los mecanismos para mantener una red estable y confiable ante diversas condiciones del sistema [1]. En este contexto, la detección de comportamientos oscilatorios en los sistemas eléctricos de potencia proporciona información valiosa para determinar las medidas de corrección adecuadas.

Los tipos de comportamientos oscilatorios en sistemas de potencia son comúnmente clasificados en [2]: (1) Respuesta ambiental debida a pequeños cambios aleatorios, (2) Respuesta transitoria (o ringdown) que ocurre después de una perturbación repentina, (3) Respuesta natural y (4) Respuesta forzada asociada a una entrada externa. Entre estos tipos de respuesta del sistema, las oscilaciones pobre/débilmente amortiguadas, no amortiguadas/sostenidas o crecientes requieren atención y análisis especiales por parte de los operadores del sistema para evitar problemas de inestabilidad.

Lo anterior ha motivado extenso despliegue a nivel nacional de Unidades de Medición de Fasores (PMU, por sus siglas en inglés) en algunos países que buscan mejorar la supervisión de los estados y dinámicas del sistema eléctrico en tiempo real mediante la aplicación de métodos basados en mediciones. Sin embargo, aunque datos sincronizados de PMU y con alta frecuencia de muestreo podrían llevar a un análisis más detallado, la instalación masiva de PMU en toda la red no siempre es posible debido a razones tecnológicas o

ORGANIZA:

APOYA:



ASOCODIS
Asociación Colombiana de
Distribuidores de Energía Eléctrica



MINISTERIO DE MINAS Y
ENERGÍA



MINISTERIO DE CIENCIA,
TECNOLOGÍA E INNOVACIÓN



WORLD
ENERGY
COUNCIL



relacionadas a los costos. En contraste, las aplicaciones de detección de oscilaciones usando los datos disponibles del Sistema de Control y Adquisición de Datos (SCADA, por sus siglas en inglés) deben enfrentar fuertes limitaciones asociadas a la baja frecuencia de muestreo (alrededor de 1 muestra cada 10 segundos).

Por otro lado, algunos métodos que aprovechan tanto la calidad de los datos de PMU como la cantidad de mediciones de SCADA han sido propuestos en la literatura [3]-[5], abordando desafíos tales como las tasas de muestreo variables, señales no lineales, señales no estacionarias y altamente ruidosas o distorsionadas. Entre las soluciones analizadas, los enfoques basados en datos surgen como una alternativa para ofrecer aplicaciones robustas al superar las desventajas de los métodos tradicionales en el dominio del tiempo y en el dominio de la frecuencia, tal como se resume en [6].

Los avances recientes en tecnologías y algoritmos de datos permiten el desarrollo de nuevas metodologías en diversos campos de aplicación, tal es el caso del Análisis Topológico de Datos (TDA, por sus siglas en inglés), un conjunto de métodos que estudian la estructura subyacente en los datos. Las aplicaciones con TDA varían desde análisis de nubes de puntos hasta la clasificación de series de tiempo, ambas abordando la naturaleza oscilatoria y muestreada de los datos de SCADA, lo que representa un gran potencial para desarrollos futuros.

Aunque se han estudiado el análisis de trayectorias oscilatorias con TDA en aplicaciones biológicas [7] y la detección de oscilaciones de baja frecuencia basada en TDA a través de la cuantificación de la respuesta cíclica con datos PMU [8], se necesita más investigación en el campo de las oscilaciones de los sistemas eléctricos.

Bajo este marco teórico, el presente trabajo propone una nueva metodología basada en TDA para clasificar automáticamente señales oscilatorias con datos altamente ruidosos y muestreados a baja frecuencia del sistema SCADA.

II. MARCO TEÓRICO

A. Detección de oscilaciones

El primer paso hacia la mitigación de riesgos asociados a las oscilaciones es la detección. Lo anterior implica identificar cuándo inicia y finaliza el evento oscilatorio observando el comportamiento de las señales de potencia, la tensión o la frecuencia eléctrica, principalmente de forma automática.

Para la mayoría de los métodos de detección, la precisión de la detección depende de las características de la oscilación: Modo de oscilación, forma del modo, amplitud y frecuencia del modo. Esta última es la frecuencia a la que ocurre la oscilación y puede agruparse a su vez en las siguientes categorías [9]: Modos de planta (de 0.01 a 0.15 Hz), modos electromecánicos inter-área (de 0.15 a 1 Hz), modos

electromecánicos locales (de 1 a 5 Hz) y oscilaciones forzadas (ocurren a cualquier frecuencia).

Según el teorema de Nyquist, la frecuencia del modo puede ser reconstruida con precisión si los datos de medición se obtienen a la tasa de muestreo mínima, que es el doble de la frecuencia máxima de las oscilaciones [10]. De lo contrario, podrían aparecer falsas oscilaciones de frecuencia más baja. Esto significa que no sería posible caracterizar las oscilaciones o identificar la fuente de forma precisa, pero aún podría ser posible detectar las oscilaciones, incluso identificar si están poco amortiguadas o no amortiguadas dependiendo de la duración del fenómeno, como se muestra en la Fig. 1. Esta última suposición es válida si el procesamiento de compresión y excepción de los datos SCADA está configurado para capturar variaciones sutiles de voltaje, potencia o frecuencia.

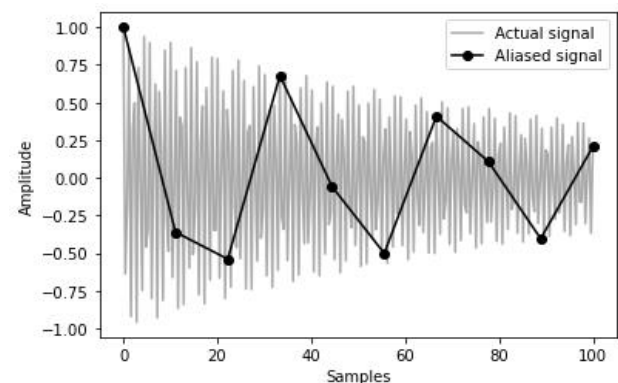


Figura 1. Ejemplo de señal oscilatoria

La presencia de armónicos en señales oscilatorias, un fuerte indicador de oscilaciones forzadas [11], no se observaría con estos datos muestreados a baja frecuencia, por lo que distinguir entre oscilaciones forzadas y naturales sería imposible.

Finalmente, aunque es posible reconstruir una señal de datos muestreada a alta frecuencia aproximada utilizando datos muestreados a baja frecuencia de múltiples fuentes, también se requieren datos basados en modelos previamente [10], lo que dificulta su implementación en aplicaciones prácticas. Por lo tanto, es relevante estudiar distintos enfoques que faciliten la detección de oscilaciones de frecuencia o potencia.

B. Análisis Topológico de Datos (TDA)

El Análisis Topológico de Datos (TDA) enmarca una serie de métodos de análisis de datos combinados como teoría de topología usada para extraer información valiosa de bases de datos bajo la premisa de que los datos tienen forma y esta es importante. El estudio de formas apunta a un mejor entendimiento de las estructuras en sets de datos complejos mediante la transformación de nubes de puntos en características. Las características topológicas pueden ser

utilizadas en aplicaciones tales como clasificación, agrupamiento y detección de patrones.

La representación de baja dimensión obtenida con TDA ofrece [12]:

- Reducción de dimensionalidad al mapear estructuras de datos complejas y de alta dimensión en una forma comprensible
- Resistencia al ruido y a los datos faltantes al retener características significativas de los datos.
- Invariancia ya que solo importa la conectividad. La inclinación, el tamaño o la orientación de los datos no cambian fundamentalmente los mismos.

Las técnicas de TDA también pueden aplicarse para el análisis de series temporales (TSA, por sus siglas en inglés) tanto para dato univariados o multivariados. Debido a la naturaleza unidimensional de las series temporales, son necesarias técnicas de procesamiento para obtener representaciones topológicas de los datos. En [13], se sugiere un método para TSA con TDA que sirve como base para el siguiente procedimiento:

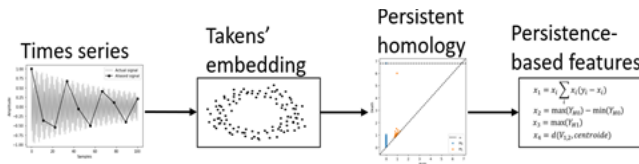


Figura 2. Proceso TSA con TDA

1) El teorema de Takens establece que los datos observados pueden producir un espacio topológicamente invariante de dimensión d utilizando métodos de incrustación de retraso [14]. Eso significa que los datos unidimensionales, como las series de tiempo $\{x(t)\}$, pueden ser convertidas en una nube de puntos finita utilizando una ventana deslizante cuyo tamaño y tiempo de retraso se definen de la siguiente forma:

$$\Phi_{d,\tau} = [x(t_1), x(t_2), \dots, x(t_k)] \mapsto \{x(t), x(t + \tau), \dots, x(t + (d - 1)\tau)\} \quad (1)$$

La dimensión de incrustación y el parámetro de retraso pueden determinarse mediante los métodos revisados en [15].

2) Homología persistente (Persistent homology): El espacio topológico K proporciona una secuencia de grupos de homología $H_k(K_\epsilon)$ que mide las características topológicas que aparecen (nacimiento) y desaparecen

(muerte) a medida que ϵ aumenta [16], tal como se muestra en la Fig 3. (a). Los grupos de homología están representados por el diagrama de persistencia en el que los puntos de pareja nacimiento-muerte se ubican sobre la diagonal, como se observa en la Fig 3. (b).

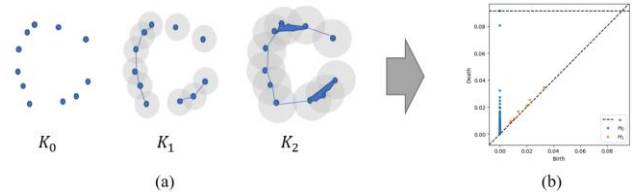


Figura 3. Diagrama de persistencia del espacio topológico K

Las homología persistente de dimensión 0 y dimensión 1 se utilizan para cuantificar la conectividad y los bucles de los conjuntos de datos, respectivamente. Dado que la incrustación de señales oscilatorias da lugar a bucles cerrados y agujeros, la homología persistente con $k \in \{0,1\}$ se utiliza para capturar estas características topológicas relevantes.

3) Características basadas en la persistencia: Una vez se ha construido el diagrama de persistencia (dgm) para $H_0(K_\epsilon)$ y $H_1(K_\epsilon)$, se pueden calcular las características topológicas para resumir las características relevantes de un espacio topológico. Las características topológicas calculadas a partir de los diagramas de persistencia se proponen en [7], algunas de ellas son:

$$feature\ 1 = \max_{i \in dgm} (d_i - b_i) \quad (2)$$

$$feature\ 2 = \sum_{i \in dgm} \frac{(d_i - b_i)}{n} \quad (3)$$

$$feature\ 3 = \sum_{i \in dgm} (d_i - b_i) \quad (4)$$

Donde d_i y b_i indican los i -ésimos valores de muerte y nacimiento de un diagrama dmg para homología persistente de dimensión 0 (0d) o dimensión 1 (1d).

Por otro lado, cuatro métricas de sumas de potencias fueron probadas en datos reales para fines de clasificación en [17], demostrando su idoneidad para asignar cantidades numéricas a códigos de barras persistentes, que contiene la misma información que los diagramas persistentes. Algunas de las características propuestas son:

$$feature\ 4 = \sum_i x_i (y_i - x_i) \quad (5)$$

$$feature\ 5 = \sum_i (y_{max} - y_i)(y_i - x_i) \quad (6)$$

$$feature\ 6 = \sum_i x_i^2 (y_i - x_i)^4 \quad (7)$$

$$feature\ 7 = \sum_i (y_{max} - y_i)^2 (y_i - x_i)^4 \quad (8)$$

Donde x son los puntos de nacimiento y y son puntos de muerte de un diagrama dgm para homología persistente de 0d o 1d.

Finalmente, se proponen dos características más basadas en la distancia al centroide del espacio n -dimensional K calculada con la siguiente expresión:

$$feature\ 8 = \sum_i (x_i - c_x)^2 + \sum_i (y_i - c_y)^2 + \quad (9)$$

$$feature\ 9 = \sum_i (y_{max} - y_i)^2 (y_i - x_i)^4 \quad (10)$$

III. METODOLOGÍA PROPUESTA

Las técnicas de TDA se pueden combinar con técnicas de aprendizaje de máquina (ML, por sus siglas en inglés) para abordar el análisis y la clasificación de series temporales, aprovechando las capacidades de extracción de características de TDA y las habilidades de aprendizaje del ML. Hay 3 categorías en las que los diagramas de persistencia se asignan a espacios auxiliares para utilizar características topológicas para tareas posteriores de aprendizaje automático [18]: Representación basada en kernel, representación basada en funciones y representación basada en vectores. Esta última se considera en el presente trabajo al vectorizar las características mencionadas anteriormente.

La Fig 4. Muestra el proceso de entrenamiento empleado para seleccionar el mejor vector de características en términos del rendimiento de un modelo de regresión logística. Primero, se generan vectores de entrada a partir de varias series temporales etiquetadas como oscilatorias y no oscilatorias, dividiendo el conjunto de datos resultante en conjuntos de datos de entrenamiento y prueba. Luego, se utilizan n características para entrenar el modelo de ML, repitiendo este paso con n tomando valores que van desde 2 hasta 7. Finalmente, se selecciona el conjunto de características con la mejor métrica del rendimiento del modelo. La métrica del rendimiento utilizada en este proceso es el puntaje F1. Una vez que se entrena el modelo, está listo para usar en casos del mundo real.

ORGANIZA:

APOYA:

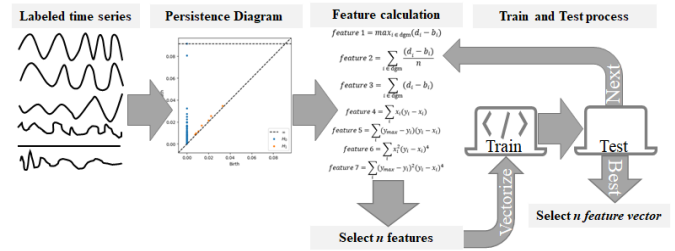


Figura 4. Proceso de entrenamiento de modelo de ML

IV. CASO DE ESTUDIO Y RESULTADOS

La metodología propuesta fue evaluada en las series temporales de potencia activa de salida de generadores, tomadas del sistema SCADA, provenientes de la operación del sistema eléctrico colombiano. Después de evaluar el rendimiento de varios conjuntos de las características mencionadas anteriormente, utilizando la métrica del puntaje F1, la mejor combinación encontrada fue la característica 1, 3, 4 y 5, alcanzando el 85%.

A continuación, se observa un caso real en la operación del sistema interconectado nacional colombiano en el que esta metodología propuesta detecta un número creciente de participación en fenómenos oscilatorios en transcurso de 2022 y hasta el 2023. Obsérvese como las unidades 2 y 3 son recurrentes en este comportamiento, volviéndolas candidatas a ajuste de PSS.

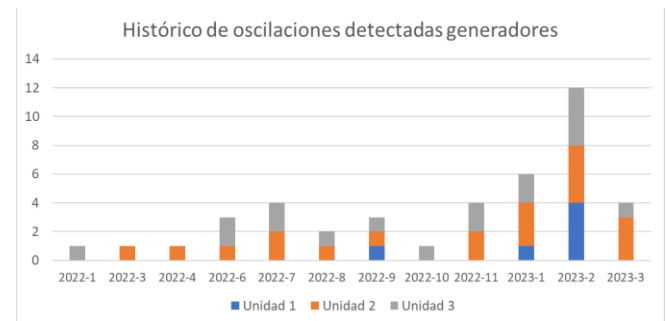


Figura 5. Histórico de oscilaciones detectadas

Al intervenir los ajustes del regulador de velocidad a principios de marzo como contra medida para mitigar el efecto oscilatorio de estas unidades sobre el sistema, se observa la siguiente tendencia en el indicador de oscilaciones débilmente amortiguadas. Ratificando la relación directa entre la cantidad de oscilaciones en las que participaban estas unidades del recurso de generación y la aparición de oscilaciones de este tipo.

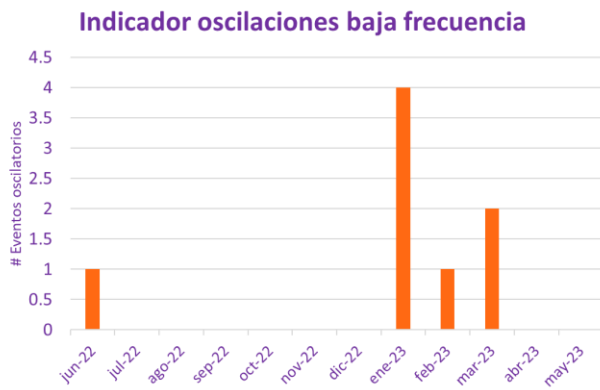


Figura 6. Histórico de indicador de oscilaciones

V. CONCLUSIONES

La metodología basada en datos propuesta para la detección automática de comportamiento oscilatorio de las unidades de generación fue probada demostrando su efectividad para este propósito en casos reales del sistema interconectado nacional colombiano, abriendo la posibilidad a mejorar los análisis de oscilaciones de baja potencia en operadores de sistemas donde el despliegue de medición sincrofasorial es limitado. El potencial de esta metodología para volverse herramienta para detectar cualquier comportamiento oscilatorio de baja frecuencia con señales distorsionadas, con ruido y no estacionarias, no solo de unidades de generación, constituye una oportunidad para el desarrollo de aplicaciones de bajo costo en cualquier sector.

VI. REFERENCIAS

- [1] B. Kroposki et al., "Achieving a 100% Renewable Grid: Operating Electric Power Systems with Extremely High Levels of Variable Renewable Energy," IEEE Power Energy Mag., vol. 15, no. 2, pp. 61–73, Apr. 2017, doi: 10.1109/MPE.2016.2637122.
- [2] D. Trudnowski, J. Pierre, M. Donnelly, M. Venkatasubramanian, and J. Follum, "Bulk Power System Oscillation Terms." WECC, Aug. 17, 2020. [Online]. Available: https://www.wecc.org/_layouts/15/WopiFrame.aspx?sourcedoc=/Administrative/Oscillation%20Terms.pdf&action=default&DefaultItemOpen=1
- [3] B. Mondal, A. K. Choudhury, M. Viswanadh, S. P. Barnwal, and D. K. Jain, "Application of PMU and SCADA data for estimation of source of forced oscillation," in 2019 International Conference on Smart Grid Synchronized Measurements and Analytics (SGSMA), May 2019, pp. 1–7. doi: 10.1109/SGSMA.2019.8784481.
- [4] J. O'Brien, T. Wu, V. Venkatasubramanian, and H. Zhang, "Source Location of Forced Oscillations Using Synchrophasor and SCADA Data," 2017.
- [5] H. Zhang, J. Ning, H. Yuan, and V. Venkatasubramanian, "Implementing Online Oscillation Monitoring and Forced Oscillation Source Locating at Peak Reliability," in 2019 North American Power

Symposium (NAPS), Oct. 2019, pp. 1–6. doi: 10.1109/NAPS46351.2019.9000376.

- [6] S. You, "Ambient PMU Data Based System Oscillation Analysis Using Multivariate Empirical Mode Decomposition," ArXiv, vol. abs/2101.05203, 2021.
- [7] C. M. M. Pereira and R. F. de Mello, "Persistent homology for time series and spatial data clustering," Expert Syst. Appl., vol. 42, no. 15, pp. 6026–6038, Sep. 2015, doi: 10.1016/j.eswa.2015.04.010.
- [8] Y. Chen, H. Chintakunta, L. Xie, Y. M. Baryshnikov, and P. R. Kumar, "Persistent-homology-based detection of power system low-frequency oscillations using PMUs," in 2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP), Dec. 2016, pp. 796–800. doi: 10.1109/GlobalSIP.2016.7905952.
- [9] North American Electric Reliability Corporation, "Reliability Guideline Forced Oscillation Monitoring & Mitigation." Sep. 2017. [Online]. Available: https://www.nerc.com/comm/PC_Reliability_Guidelines_DL/Reliability_Guideline_-_Forced_Oscillations_-_2017-07-31_-_FINAL.pdf
- [10] J.-O. Baek and S. Kim, "Detection and Analysis of Electromechanical Oscillation in Power Systems with Low-Sampled Data Using Modal Analysis Methods," J. Electr. Eng. Technol., vol. 15, no. 5, pp. 1999–2006, Sep. 2020, doi: 10.1007/s42835-020-00471-0.
- [11] Ruichao Xie and D. Trudnowski, "Distinguishing features of natural and forced oscillations," in 2015 IEEE Power & Energy Society General Meeting, Jul. 2015, pp. 1–5. doi: 10.1109/PESGM.2015.7285781.
- [12] N. Saul, "scikit-tda 0.0.4 documentation." 2019. [Online]. Available: <https://scikit-tda.org/theory.html#topological-data-analysis>
- [13] A. Myers, E. Munch, and F. A. Khasawneh, "Persistent Homology of Complex Networks for Dynamic State Detection," Phys. Rev. E, vol. 100 2–1, p. 022314, 2019.
- [14] F. Takens, "Detecting strange attractors in turbulence," in Dynamical Systems and Turbulence, Warwick 1980, Berlin, Heidelberg, 1981, pp. 366–381.
- [15] N. Ravishanker and R. Chen, "Topological Data Analysis (TDA) for Time Series," ArXiv Appl., 2019.
- [16] Edelsbrunner, Letscher, and Zomorodian, "Topological Persistence and Simplification," Discrete Comput. Geom., vol. 28, no. 4, pp. 511–533, Nov. 2002, doi: 10.1007/s00454-002-2885-2.
- [17] A. B. Adcock, E. Carlsson, and G. E. Carlsson, "The Ring of Algebraic Functions on Persistence Bar Codes," ArXiv Rings Algebr., 2013.
- [18] F. Hensel, M. Moor, and B. Rieck, "A Survey of Topological Machine Learning Methods," Front. Artif. Intell., vol. 4, p. 52, 2021, doi: 10.3389/frai.2021.681108.

ORGANIZA:

APOYA: