

PHƯƠNG PHÁP CHẨN ĐOÁN LỖI CHO HỆ THỐNG PHI TUYẾN ỨNG DỤNG THUẬT TOÁN KPCA VÀ MẠNG NƠ-RON NHÂN TẠO

NON-LINEAR SYSTEM FAULT DIAGNOSIS METHOD BASED ON KPCA AND ANN CLASSIFIER

Huỳnh Duy An¹, Võ Ngọc Yến Phương²

¹Khoa Điện – Điện tử, Trường Cao Đẳng Lý Tự Trọng Thành phố Hồ Chí Minh

²Khoa Công nghệ Cơ khí, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh

TÓM TẮT

Bài báo đề xuất phương pháp phát hiện và chẩn đoán lỗi quá trình công nghệ đa biến phức tạp. Ở phương pháp này, các lỗi hệ thống được phát hiện dựa theo những sai lệch giữa đáp ứng từ hệ thống thực tế và dữ liệu tiêu chuẩn. Phép phân tích thành phần chính sử dụng hạt nhân (Kernel Principal analysis hay KPCA) kết hợp với bộ phân lớp được xây dựng dựa trên mô hình mạng nơ-ron nhân tạo (ANN) giúp nhanh chóng phát hiện và phân loại lỗi trong quá trình chẩn đoán hệ thống.

Trong nghiên cứu này, hệ thống bồn khuấy liên tục có gia nhiệt được khảo sát để kiểm chứng cho phương pháp đề xuất. Hệ thống được mô hình hóa và mô phỏng, kết quả mô phỏng chứng minh phương pháp chẩn đoán lỗi có khả năng phát hiện và phân loại lỗi với độ chính xác trên 85%.

Từ khóa: Chẩn đoán lỗi hệ thống; Mạng nơ-ron nhân tạo; Phân tích thành phần chính dữ liệu (PCA, KPCA); Hàm Softmax.

ABSTRACT

This study proposes a method of fault detecting and diagnosing for the technological process based on real-time system responds. In this method, the system faults are detected based on the errors between standard data and real-time process responses. The Kernel Principal Component Analysis (KPCA) combining with ANN model is utilized to detect and classify immediately errors during fault diagnostic process. In this research, the method is verified on Stirred-Tank Heating Process. The system is modeled and simulated, all results prove that the proposed method could detect and classify faults with more than 85% in accuracy.

Keywords: Fault diagnostics; PCA; KPCA; ANN; Data-driven; Softmax function.



1. TỔNG QUAN

Những năm trở lại đây, nhu cầu cho các quy trình công nghệ với hiệu năng cao đang gia tăng một cách nhanh chóng như hệ quả của quá trình công nghiệp hóa. Trong đó, việc điều khiển và giám sát các quy trình hóa học nói riêng và các hệ thống phi tuyến nói chung luôn thu hút sự chú ý của các nhà nghiên cứu bởi tầm quan trọng và tính chất phức tạp của chúng. Một số đối tượng trong các quá trình này có thể kể đến như: lưu lượng, áp suất, nồng độ, nhiệt độ,... Bên cạnh phát triển các giải thuật điều khiển mới nhằm cải thiện độ ổn định và hiệu năng của quá trình, việc giám sát hệ thống cũng hết sức quan trọng nhằm đảm bảo hệ thống luôn được duy trì ở trạng thái tốt nhất cũng như hạn chế tối đa các trường hợp ngoài ý muốn xảy ra.

Ngày nay, có rất nhiều sự lựa chọn các thuật toán được xây dựng hướng đến việc xử lý dữ liệu hiệu quả nhất có thể. Tuy nhiên, mỗi phương pháp đều có những ưu và nhược điểm riêng. Nhìn chung, chúng ta có thể phân chia các phương pháp này thành hai nhóm: phương pháp định tính và phương pháp định lượng. Các phương pháp định tính đòi hỏi các nhà phát triển phải hiểu biết thấu đáo về hệ thống và việc áp dụng vào các quy trình thực tế là rất khó khăn. Vì vậy, trong phạm vi bài nghiên cứu này, tác giả chỉ xét đến phương pháp xử lý dữ liệu dựa trên phương pháp định lượng. Các phương pháp định lượng số học có thể được kể đến như: phân tích thành phần chính (PCA hay KPCA), khoảng cách Euclidean, khoảng cách Mahalanobis,... Tuy nhiên, PCA nói chung và KPCA nói riêng có thể được xem như hai trong số các công cụ hiệu quả nhất cho các hệ thống đa biến phức tạp. Khả năng giảm chiều dữ liệu và hàm phi tuyến cho phép phương pháp này trích xuất các đặc trưng cần thiết cũng như loại

bỏ nhiều cho dữ liệu thô đầu vào. Phương pháp phân tích thành phần chính (PCA) đơn giản [1] đã được áp dụng thành công để giám sát và chẩn đoán các bất thường trong hệ thống. Những năm sau đó, các mô hình PCA được tinh chỉnh [2], [3], [4], [5] để cải thiện chất lượng xử lý dữ liệu và áp dụng vào các hệ thống phi tuyến đa biến phức tạp hơn. Tuy nhiên, các mô hình PCA kể trên dựa theo kỹ thuật tuyến tính hóa trong khi các quá trình thực tế có độ phi tuyến cao và hết sức phức tạp. Vì vậy, dựa trên mô hình PCA cổ điển, việc phân tích thành phần chính dựa trên mặt nạ phi tuyến (Kernel Principal Component Analysis hay KPCA) được xét đến như là một sự lựa chọn tốt nhất cho các hệ thống này. Mặt nạ phi tuyến (Kernel) thực chất là phép biến đổi dữ liệu nhằm phân phối dữ liệu trên một siêu phẳng để dữ liệu trở nên khả phân tuyến tính. Các nghiên cứu [6], [7] đề xuất phương pháp chẩn đoán lỗi cho hệ thống có yếu tố hóa học sử dụng mô hình KPCA dựa trên hai tiêu chuẩn đánh giá là T^2 và Q . Gần đây, [8] đã ứng dụng tính chọn lọc dữ liệu của mô hình PCA cổ điển để xây dựng mô hình RKPCA (Reduced Kernel PCA) nhằm chẩn đoán lỗi cho hệ thống sản xuất xi-măng. Hơn thế nữa, mô hình hybrid – KPCA [9] lần đầu tiên được ứng dụng thành công để phát hiện các bất thường trong hộp số bánh răng hành tinh.

Mặc dù các phương pháp kể trên có thể xử lý dữ liệu phi tuyến và phát hiện các bất thường một cách hiệu quả nhưng việc xây dựng các bộ phân loại lỗi cho từng trường hợp cụ thể là hết sức khó khăn. Để giải quyết vấn đề này, một số mô hình mạng nơ-ron nhân tạo (gọi tắt là ANN) được sử dụng như là: AE (Auto Encoder), LSTM (Long Short Term Memory), AE-LSTM,... Mô hình mạng nơ-ron chủ động và mạng nơ-ron tích chập (gọi tắt là CNN) [10], [11] được giới thiệu nhằm chẩn đoán cho quá trình hóa học dựa trên dữ liệu được thu tập từ

các cảm biến trong hệ thống. Việc so sánh các mô hình mạng nơ-ron khác nhau trong việc chẩn đoán lỗi hệ thống [12] cho thấy mỗi mô hình đều chỉ có thể áp dụng cho một số hệ thống có đặc tính và yêu cầu nhất định. Hầu hết các mô hình ANN đòi hỏi nhà phát triển phải nhận diện được tất cả bất thường có trong hệ thống để gán nhãn phù hợp cho từng trường hợp cụ thể. Tuy nhiên, với một hệ thống phức tạp hay mới được xây dựng thì rất khó để có đủ thông tin cho việc phân loại tất cả các lỗi xảy ra, nhưng chúng ta vẫn có thể phát hiện được chúng bằng cách sử dụng phương pháp số học ví dụ như KPCA.

Trong phạm vi nghiên cứu này, tác giả đề xuất phương pháp kết hợp mô hình ANN và phương pháp số (KPCA) để xây dựng mô hình chẩn đoán lỗi cho quá trình công nghiệp. Chẩn đoán ở đây có ý nghĩa phát hiện và phân loại lỗi nếu có.

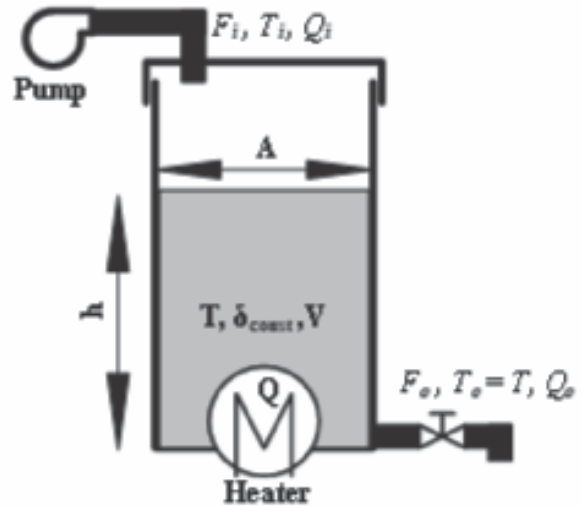
2. CƠ SỞ LÝ THUYẾT

2.1. Đối tượng khảo sát

Trong nghiên cứu này, mô hình bồn nước gia nhiệt được sử dụng để khảo sát phương pháp chẩn đoán lỗi hệ thống. Đây là hệ thống có tính phi tuyến cao thường được xét đến trong các nghiên cứu liên quan. Hình [1] và bảng 1 thể hiện cấu trúc hệ thống và các tham số được xét đến.

Bảng 1. Thông số cơ bản hệ bồn gia nhiệt

Thông số	Giá trị
Tiết diện (A)	1500 cm ²
Hệ số xả (C)	0.6
Tiết diện xả (b)	7 cm ²
Hệ số công suất của bơm (K)	150 cm ³ /Vs
Gia tốc trọng trường (g)	9.81 m/s ²



Hình 1. Mô hình hệ bồn nước gia nhiệt

2.2. Giám sát quá trình dựa trên thuật toán KPCA

Được xét đến như là kỹ thuật biến đổi dữ liệu tuyến tính truyền thống, PCA chuyển dữ liệu nhiều chiều thành dữ liệu đơn giản hơn sao cho giữ lại được tối đa thông tin của dữ liệu ban đầu. Khi thuật toán được xử lý trên miền đặc trưng của dữ liệu, KPCA được áp dụng. Trong ngành khoa học dữ liệu, KPCA được biết đến như phương pháp học máy dựa trên mặt nạ (hay mặt nạ phi tuyến). Trong nghiên cứu này, ba biến số được đưa vào mô hình KPCA bao gồm mức nước, lưu lượng nước và nhiệt độ của nước. Các bước xử lý cơ bản nhằm xử lý dữ liệu dựa theo thuật toán KPCA được mô tả trong [2], [6], [7] và các nghiên cứu liên quan trước đó.

2.3. Mô hình ANN với ngõ ra là hàm Softmax

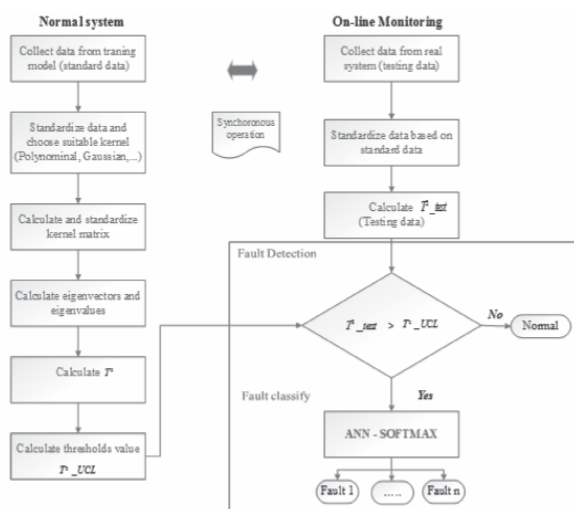
Trong các ứng dụng phân loại đối tượng, hàm Softmax được sử dụng như là hàm kích hoạt cuối cùng cho mô hình mạng nơ-ron nhằm đưa các phần tử trong vector ngõ ra về khoảng [0, 1] để có thể so sánh với ngưỡng phân loại.

giá trị (0,1) đại diện cho kết quả nhận dạng xác suất nhận ra từng đối tượng trong tổng thể.

Trong mô hình thực nghiệm ở phần sau, vector đầu vào cho mô hình ANN bao gồm các tín hiệu điều khiển và đáp ứng hệ thống như là điện áp điều khiển bơm nước, điện áp điều khiển bộ gia nhiệt, mức nước, lưu lượng nước và nhiệt độ nước trong bồn. Vector ngõ ra với mỗi phần tử thể hiện xác suất nhận dạng cho từng lỗi được quy ước trước trong hệ thống.

3. PHƯƠNG PHÁP THỰC HIỆN

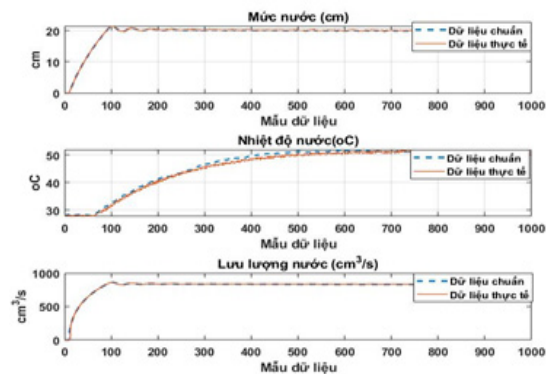
Tác giả đề xuất mô hình phát hiện và phân loại lỗi hoạt động độc lập với nhau. Điều này dựa trên ưu điểm xử lý dữ liệu của phương pháp số (KPCA) và mô hình mạng ANN – Softmax. Hình [2] thể hiện quá trình thu thập và xử lý dữ liệu giữa dữ liệu mẫu tiêu chuẩn và đáp ứng hệ thống thực.



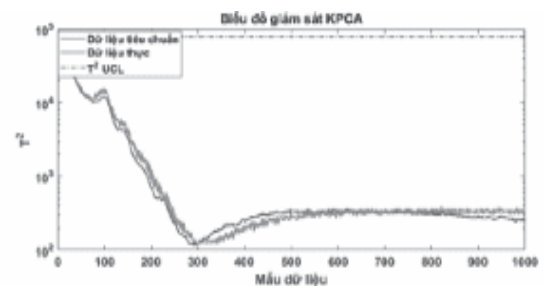
Hình 2. Lưu đồ chẩn đoán hệ thống tổng quát

4. KẾT QUẢ THỰC NGHIỆM

4.1. Khi hệ thống hoạt động bình thường



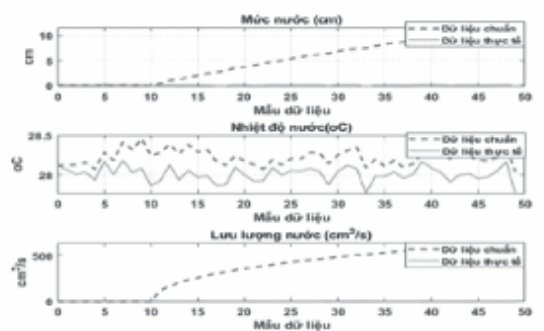
Hình 3. Kết quả đáp ứng khi hệ thống hoạt động bình thường



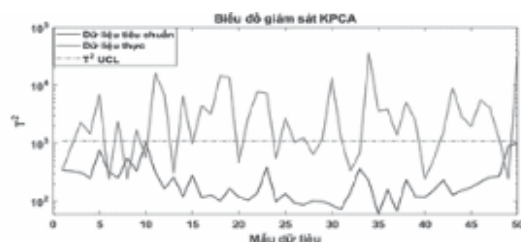
Hình 4. Kết quả biểu đồ giám sát KPCA

Hình 3 và hình 4 thể hiện đáp ứng khi hệ thống hoạt động bình thường không phát sinh lỗi, cả hai đường biểu diễn dữ liệu đều nằm dưới ngưỡng T^2 -UCL do không có sự sai lệch bất thường mà chỉ có biến động ngẫu nhiên trong hệ thống trong ngưỡng cho phép.

4.2. Khi hệ thống xảy ra lỗi



Hình 5. Đáp ứng hệ thống khi xảy ra lỗi



Hình 6. Kết quả biểu đồ giám sát KPCA

Hình 5 và hình 6 thể hiện đáp ứng hệ thống và giám sát dữ liệu trên mô hình KPCA khi có tác nhân gây lỗi (trong trường hợp này là lỗi bơm không hoạt động). Dễ dàng nhận thấy rằng, dữ liệu thực tế có sự chênh lệch lớn so với dữ liệu tiêu chuẩn. Dưới đây là bảng thống kê độ chính xác của quá trình chẩn đoán lỗi khác nhau của hệ thống với 100 mẫu dữ liệu cho mỗi lỗi khác nhau. Kết quả cho thấy mô hình chẩn đoán lỗi có độ chính xác trên 85%, phù hợp với yêu cầu nghiên cứu đề ra cho phương pháp.

Bảng 2. Thống kê kết quả chẩn đoán lỗi hệ thống

Lỗi	Độ chính xác	Lỗi	Độ chính xác
Lỗi bơm	91%	Lỗi bộ gia nhiệt	88%
Hiệu suất bơm giảm	87%	Hiệu suất gia nhiệt giảm	93%
Lỗi cảm biến mức nước	96%	Lỗi cảm biến nhiệt độ	89%

5. KẾT LUẬN

Phương pháp chẩn đoán lỗi bao gồm cơ chế phát hiện và phân loại lỗi hoạt động độc lập với nhau có độ chính xác cao do tận dụng được ưu điểm của từng phương pháp. Việc vận dụng phương pháp xử lý dữ liệu số học và mô hình mạng ANN cho khả năng thích ứng với các hệ thống đa biến phi tuyến phức tạp. Kết quả thực

nghiệm đã chứng minh tính ứng dụng và hiệu quả cao của phương pháp được tác giả đề xuất. ♦

Ngày nhận bài: 26/3/2023

Ngày phản biện: 10/4/2023

Tài liệu tham khảo:

- [1]. J.-M. Lee, S. J. Qin, and I.-B. Lee; "Fault detection and diagnosis based on modified independent component analysis". AICHE Journal, pp. 3501-3514, 2006.
- [2]. VILLEGAS, Thamara; FUENTE, Maria Jesús; RODRÍGUEZ, Miguel; *Principal component analysis for fault detection and diagnosis, experience with a pilot plant. Advances in Computational Intelligence*, Man-Machine Systems and Cybernetics, 2010, 147-152.
- [3]. F. Harrou, M. N. Nounou, H. N. Nounou, and M. Madakyaru; "Statistical fault detection using PCA-based GLR hypothesis testing". Journal of Loss Prevention in the Process Industries, pp. 129-139, 2013.
- [4]. T. Ait-Izem, M.-F. Harkat, M. Djeghaba, and F. Kratz; "On the application of interval PCA to process monitoring: A robust strategy for sensor FDI with new efficient control statistics". Journal of Process Control, vol. 63, pp. 29-46, 2018, doi: 10.1016/j.jprocont.2018.01.006.
- [5]. M. A. Márquez-Vera, O. López-Ortega, L. E. Ramos-Velasco, A. Rodríguez-Romero, J. C. Ramos-Fernández, and J. A. Hernández-Salazar; "Adaptive threshold PCA for fault detection and isolation". Journal of Robotics and Control (JRC), 2021.
- [6]. V. H. Nguyen and J.-C. Golinval; "Fault detection based on Kernel Principal Component Analysis". Engineering Structures, vol. 32, no. 11, pp. 3683-3691, 2010, doi: 10.1016/j.engstruct.2010.08.012.
- [7]. M. Mansouri, M. Nounou, H. Nounou, and N. Karim; "Kernel PCA-based GLRT for nonlinear fault detection of chemical processes". Journal of Loss Prevention in the Process Industries, vol. 40, pp. 334-347, 2016, doi: 10.1016/j.jlp.2016.01.011.
- [8]. F. Bencheikh, M. Harkat, A. Kouadri, and A. Bensmail; "New reduced kernel PCA for fault detection and diagnosis in cement rotary kiln". Chemometrics and Intelligent Laboratory Systems, vol. 204, p. 104091, 2020, doi: 10.1016/j.chemolab.2020.104091.
- [9]. R. Li and Y. Liu; "Fault Diagnosis for the Planetary Gearbox Based on a Hybrid Dimension Reduction Algorithm". 2019 CAA Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS), 2019.
- [10]. P. Jiang, Z. Hu, J. Liu, S. Yu, and F. Wu; "Fault Diagnosis Based on Chemical Sensor Data with an Active Deep Neural Network". Sensors, p. 1695, 2016.
- [11]. H. Wu and J. Zhao; "Deep convolutional neural network model based chemical process fault diagnosis". Computers & Chemical Engineering, pp. 185-197, 2018.
- [12]. G. S. Chadha and A. Schwung; "Comparison of deep neural network architectures for fault detection in Tennessee Eastman process". 2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), 2017.