

Effects of K-value in the K-Nearest Neighbors Algorithm on Performance of Chiller Fault Detection and Diagnosis

Le Ha Dong Quan, Le Minh Nhut*

Ho Chi Minh City University of Technology and Education, Vietnam.

* Corresponding author. Email: nhutlm@hcmute.edu.vn

ARTICLE INFO

Received: 13/11/2022
Revised: 26/11/2022
Accepted: 30/11/2022
Published: 28/04/2023

KEYWORDS

HVAC;
FDD;
KNN;
Energy;
Condenser.

ABSTRACT

Fault detection and diagnosis for water chiller systems help to extend the life of the system, prevent serious damage, and reduce energy consumption. For these reasons, this paper investigates the K value of the KNN algorithm and proposes a fault detection and diagnosis model for the water chiller system based on the K-nearest neighbors algorithm (FDD-KNN). The study results indicated that the FDD-KNN model has an accuracy rate of 99.15% or higher when the value of K is 1. When compared to previous studies for LV3 and LV4 severity faults, the proposed model exhibits high and uniform diagnostic accuracy across all faults and fault levels. In addition, the fault isolation and fluctuating trends of the variables in practice for Normal and fouling of the ConFoul condenser were also checked by the actual database of the water chiller systems, Center Saigon Building, Ho Chi Minh City, Vietnam. The results show that the changing trend of the two variables, TCO and TRC, is consistent with the actual operating conditions of the system. Therefore, the proposed model FDD-KNN with K=1 is completely reliable to apply to fault detection and diagnosis for water chiller systems.

Ảnh Hưởng Giá Trị K Của Thuật Toán KNN Đến Hiệu Suất Chẩn Đoán Lỗi Cho Hệ Thống Điều Hòa Không Khí Trung Tâm

Lê Hà Đông Quân, Lê Minh Nhựt*

Trường Đại học Sư phạm Kỹ Thuật Thành phố Hồ Chí Minh, Việt Nam.

* Tác giả liên hệ. Email: nhutlm@hcmute.edu.vn

THÔNG TIN BÀI BÁO

Ngày nhận bài: 13/11/2022
Ngày hoàn thiện: 26/11/2022
Ngày chấp nhận đăng: 30/11/2022
Ngày đăng: 28/04/2023

TỪ KHÓA

Hệ thống điều hòa không khí;
FDD;
KNN;
Năng lượng;
Bình ngưng.

TÓM TẮT

Phát hiện và chẩn đoán lỗi kịp thời cho hệ thống điều hòa không khí trung tâm (ĐHKKT) giúp tăng tuổi thọ, ngăn ngừa các hư hỏng nghiêm trọng và giảm lãng phí năng lượng của hệ thống. Từ thực tế trên, nghiên cứu này xác định giá trị K của thuật toán KNN, đề xuất mô hình phát hiện và chẩn đoán lỗi cho hệ thống ĐHKKT dựa trên thuật toán K-nearest neighbors (FDD-KNN). Kết quả cho thấy khi giá trị K=1 thì mô hình FDD-KNN cho hiệu suất chẩn đoán chính xác các lỗi đạt tỉ lệ từ 99.15% trở lên. Khi so sánh với các nghiên cứu trước đối với các lỗi nghiêm trọng LV3 và LV4, mô hình đề xuất cho thấy sự chẩn đoán chính xác cao và đồng đều tại các lỗi và mức độ lỗi. Ngoài ra, khả năng cô lập lỗi và xu hướng dao động của các biến số trong thực tế đối với trạng thái bình thường Normal và bám bình ngưng ConFoul được kiểm tra bằng cơ sở dữ liệu thực tế của hệ thống ĐHKKT công trình Saigon Centre, thành phố Hồ Chí Minh, Việt Nam. Kết quả cho thấy sự thay đổi của biến số TCO và TRC phù hợp với điều kiện vận hành thực tế của hệ thống. Do đó, mô hình đề xuất FDD-KNN với K=1 hoàn toàn đáng tin cậy để áp dụng vào chẩn đoán lỗi cho các hệ thống ĐHKKT.

Doi: <https://doi.org/10.54644/jte.76.2023.1309>

Copyright © JTE. This is an open access article distributed under the terms and conditions of the [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/) which permits unrestricted use, distribution, and reproduction in any medium for non-commercial purpose, provided the original work is properly cited

1. Giới thiệu

Hệ thống điều hòa không khí trung tâm được lắp đặt ngày càng nhiều trong các công trình xây dựng nhằm đáp ứng điều kiện tiện nghi và đóng vai trò quan trọng trong một số ngành công nghiệp sản xuất. Các hệ thống này có cấu tạo phức tạp nên trong quá trình hoạt động dễ xảy ra các lỗi sự cố gây hư hỏng. Khi vận hành hệ thống trong tình trạng lỗi dẫn đến suy giảm tuổi thọ thiết bị, gián đoạn quá trình sản xuất và lãng phí năng lượng. Việc phát hiện và chẩn đoán lỗi kịp thời giúp chủ động lên kế hoạch bảo trì, tránh dẫn đến những hư hỏng nghiêm trọng và tiết kiệm được khoảng 15 - 30% năng lượng tiêu thụ. Mặc dù các loại lỗi khác nhau có thể xảy ra trong quá trình vận hành nhưng những lỗi gây suy giảm hiệu suất hệ thống cần được chẩn đoán kịp thời vì chúng có những biểu hiện khó nắm bắt, gây ra lãng phí năng lượng và dễ dẫn đến ngưng hoạt động hệ thống khi lỗi tiến triển nghiêm trọng. Các phương pháp phát hiện và chẩn đoán lỗi (FDD) được chia thành ba nhóm: phương pháp dựa trên định lượng, phương pháp dựa trên định tính và phương pháp dựa trên dữ liệu [1]. Phương pháp dựa trên định lượng cung cấp các mô tả toán học về các quá trình vật lý chi tiết, nhược điểm là rất khó áp dụng đối với những hệ thống phức tạp nên hiếm khi được sử dụng trong FDD. Phương pháp dựa trên định tính dựa trên các mối quan hệ vật lý định tính, các kiến thức cá nhân về hệ thống để xây dựng quy tắc chẩn đoán, nhược điểm là khó áp dụng vào hệ thống lớn và không đảm bảo tính khách quan. Phương pháp dựa trên dữ liệu sử dụng các mô hình toán học để mô tả liên hệ giữa đầu vào và đầu ra trong lịch sử dữ liệu được lưu trữ. Khác với hai phương pháp trên, phương pháp này đảm bảo tính khách quan, chính xác cao đối với những hệ thống phức tạp và dễ dàng phát triển theo sự thay đổi của hệ thống thực.

Trong những năm gần đây, nhiều nghiên cứu đã tập trung vào phát triển FDD cho hệ thống điều hòa không khí trung tâm bằng phương pháp dựa trên dữ liệu. Huang và cộng sự [2] đề xuất phương pháp FDD sử dụng thuật toán AC và 10 biến số vật lý. Kết quả cho thấy mô hình có hiệu suất chẩn đoán chính xác 100% đối với lỗi bám bẩn bình ngưng, tuy nhiên lỗi dư môi chất có hiệu suất chẩn đoán không cao chỉ với 67.4%. Suowei và cộng sự [3] đã phát mô hình SI-PB-BNC có kết quả chẩn đoán chính xác tuyệt đối lỗi khí không ngưng. Tuy nhiên, các mô hình có tỷ lệ chẩn đoán chính xác chưa cao đối với các lỗi rò rỉ chất làm lạnh 69.4% và bám bẩn bình ngưng 73.6%. Mô hình SP-CNN7-16 được đề xuất bởi Gao và cộng sự [4] khắc phục được một phần nhược điểm về chênh lệch mức độ chẩn đoán chính xác giữa các lỗi. Tuy nhiên tỷ lệ chẩn đoán các lỗi dư dầu, rò rỉ môi chất lạnh và dư môi chất lạnh còn chưa nổi bật. Chen và cộng sự [5] sử dụng thuật toán GDW-SVDD và 16 biến số vật lý, kết quả cho thấy mô hình chẩn đoán lỗi giảm lưu lượng nước lạnh qua bình bay hơi và giảm hiệu suất chẩn đoán khi mức độ lỗi giảm. Mô hình CBA được trình bày bởi Liu và cộng sự [6] có tỷ lệ chẩn đoán chính xác lỗi dư môi chất lạnh là 89.74% ở mức độ lỗi nghiêm trọng nhất. Kết quả này cần được cải thiện vì dư môi chất gây các tác động nguy hiểm đến hệ thống. Mô hình EBD-DBN [7] của Wang và cộng sự cũng cho thấy khả năng hạn chế trong chẩn đoán các lỗi môi chất lạnh. Ngoài ra, tỷ lệ chẩn đoán chính xác giảm nhiều khi mức độ nghiêm trọng giảm, điển hình như lỗi bám bẩn bình ngưng giảm độ chính xác từ 98% tại mức độ lỗi cao nhất xuống 74% tại mức độ lỗi thấp nhất.

Các nghiên cứu FDD cho hệ thống điều hòa không khí trung tâm nổi bật về số lượng và đa dạng về thuật toán xây dựng mô hình. Trong số đó, thuật toán K-nearest neighbors (KNN) được xem là một trong những thuật toán hàng đầu, có vai trò quan trọng trong việc giải quyết các vấn đề phân loại và khai thác dữ liệu [8]. Zhang và cộng sự [9] thực hiện kết hợp các kết quả phân loại của các thuật toán KNN và một số thuật toán khác để xác định phân loại cuối cùng cho mẫu dữ liệu. Kết quả cho thấy sự kết hợp này cho tỷ lệ chẩn đoán chính xác cao hơn so với sử dụng từng thuật toán riêng lẻ. Yan và cộng sự [10] so sánh kết quả của các mô hình sử dụng các thuật toán khác để lựa chọn bộ phân loại cho mạng GAN. Mặc dù không đạt hiệu suất chẩn đoán cao nhất nhưng thuật toán KNN thể hiện nhiều tiềm năng cần được khai thác.

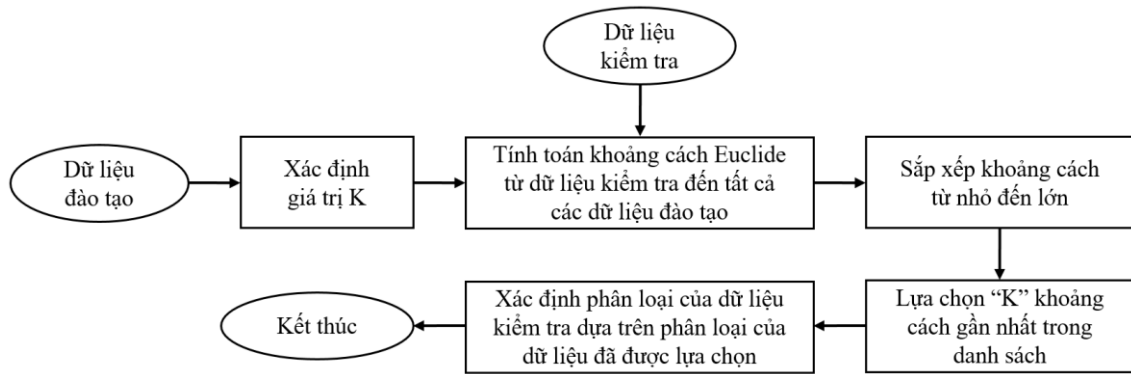
Trong các nghiên cứu vừa nêu trên, giá trị K trong thuật toán KNN chưa được phân tích, đây là tham số quan trọng ảnh hưởng đến hiệu suất chẩn đoán nên cần được quan tâm xác định. Các nghiên cứu trước đây sử dụng nhiều biến số vật lý làm tăng lượng thông tin chồng chéo. Hậu quả là làm giảm hiệu suất chẩn đoán, tăng chi phí cảm biến và thời gian tính toán. Ngoài ra, độ chính xác của các mô hình chẩn đoán lỗi giảm khi mức độ lỗi giảm. Việc duy trì hiệu suất chẩn đoán cho mọi lỗi và mọi mức độ lỗi sẽ nâng cao độ tin cậy và khả năng đáp ứng của mô hình trong các vấn đề thực tế. Trong nghiên cứu này, giá trị K của thuật toán KNN được phân tích và xác định, mô hình phát hiện và chẩn đoán lỗi cho

hệ thống điều hòa trung tâm dựa trên thuật toán K-nearest neighbors (FDD-KNN) được đề xuất để khắc phục những nhược điểm trên. Ngoài ra, mô hình đề xuất cũng được kiểm chứng độ tin cậy dựa trên dữ liệu vận hành thực tế của hệ thống điều hòa không khí trung tâm, giải nhiệt bằng nước của tòa nhà Saigon Centre, thành phố Hồ Chí Minh.

2. Phương pháp nghiên cứu

2.1. Thuật toán K-Nearest Neighbors (KNN) trong mô hình phân loại

Ý tưởng cơ bản của KNN là nếu phần lớn các điểm dữ liệu lân cận thuộc một nhóm thì dữ liệu cần phân loại cũng thuộc nhóm đó. Do đó, theo khoảng cách Euclide, phương pháp KNN tìm kiếm K điểm dữ liệu lân cận đã được phân loại sẵn. Sau đó, các phương pháp biểu quyết và phân loại của các điểm lân cận được sử dụng để xác định phân loại của dữ liệu mới. Trong bài báo này, phương pháp biểu quyết đa số được sử dụng. Phương pháp biểu quyết đa số có kết quả phân loại là nhãn xuất hiện thường xuyên nhất trong K điểm lân cận. Hình 1 thể hiện các bước giải thuật của thuật toán KNN.



Hình 1. Các bước giải thuật của thuật toán KNN

2.2. Quá trình tiền xử lý dữ liệu

Các dữ liệu tại trạng thái không ổn định của hệ thống được loại bỏ bởi kết quả nghiên cứu của Glass và cộng sự [11]. Các điểm dữ liệu dị biệt được loại bỏ bằng việc đặt ngưỡng dưới và ngưỡng trên thông qua các giá trị tứ phân vị của dữ liệu. Khi dữ liệu cao hơn ngưỡng trên và ngưỡng dưới, nó bị loại bỏ.

Ngưỡng dưới được xác định bằng công thức:

$$(Q_1 - 1.5 \times (Q_3 - Q_1)) \quad (1)$$

Ngưỡng trên được xác định bằng công thức:

$$(Q_1 + 1.5 \times (Q_3 - Q_1)) \quad (2)$$

Trong đó, Q_1 và Q_3 lần lượt là tứ phân vị thứ nhất và thứ ba của dữ liệu

Quá trình chuẩn hóa dữ liệu cũng được thực hiện để giảm các tác động bất lợi đến mô hình. Chuẩn hóa Z-score được áp dụng trong bài báo này.

2.3. Các chỉ số đánh giá hiệu suất mô hình phân loại

Ma trận nhầm lẫn: là một ma trận vuông với kích thước mỗi chiều bằng với số lượng các nhóm cần được phân loại. Xét một bài toán phân loại nhị phân, ma trận nhầm lẫn điển hình được thể hiện trong bảng 1.

Bảng 1. Ma trận nhầm lẫn điển hình cho bài toán phân loại nhị phân

	Chẩn đoán: Dương tính	Chẩn đoán: Âm tính
Thực sự: Dương tính	Dương tính thật (TP)	Âm tính giả (FN)
Thực sự: Âm tính	Dương tính giả (FP)	Âm tính thật (TN)

Trong đó dương tính thật (TP) và âm tính thật (TN) là số mẫu dữ liệu được mô hình phân loại chính xác. Dương tính giả (FP) và âm tính giả (FN) là số mẫu dữ liệu được mô hình phân loại sai. Một mô

hình phân loại tốt sẽ có ma trận nhầm lẫn với số lượng lớn mẫu dữ liệu nằm trên đường chéo chính và số lượng nhỏ (lý tưởng là 0) nằm ngoài đường chéo. Đối với những vấn đề phân loại nhiều nhóm như chẩn đoán lỗi hệ thống điều hòa không khí trung tâm, trạng thái đang cần được chẩn đoán có thể xem là nhãn “dương tính” và các trạng thái còn lại có thể xem là nhãn “âm tính”.

Tỷ lệ chẩn đoán chính xác (Hit Rate – HR): cho biết độ chẩn đoán chính xác của mô hình với từng trạng thái. Tỷ lệ HR được tính từ các thành phần trong ma trận nhầm lẫn như sau.

$$HR = TP / (TP + FP) \quad (3)$$

2.4. Mô hình FDD-KNN cho hệ thống điều hòa không khí trung tâm

Cấu trúc mô hình bao gồm ba mô-đun: tiền xử lý dữ liệu, phát triển mô hình, mô hình phát hiện và chẩn đoán.

Tiền xử lý dữ liệu: dữ liệu vận hành được thực hiện quá trình tiền xử lý, các biến số vật lý được phân tích và lựa chọn. Với mỗi phân loại, các mẫu dữ liệu được lấy một cách ngẫu nhiên. Tập hợp các mẫu dữ liệu thu được được phân chia thành các tập dữ liệu độc lập: đào tạo, kiểm định và kiểm tra.

Phát triển mô hình: cấu trúc của mô hình FDD được xác định, giá trị K của thuật toán KNN được phân tích sự ảnh hưởng đến hiệu suất chẩn đoán và xác định giá trị tối ưu thông qua quá trình kiểm định mô hình.

Mô hình phát hiện và chẩn đoán lỗi: gồm giai đoạn 1 là chẩn đoán trạng thái, giai đoạn 2 là chẩn đoán mức độ lỗi. Trong giai đoạn 1, nếu dữ liệu được chẩn đoán đúng trạng thái và trạng thái là bất kỳ lỗi nào thì tiến hành chẩn đoán mức độ lỗi trong giai đoạn 2. Tập dữ liệu kiểm tra được sử dụng để kiểm tra khả năng hoạt động của mô hình.

2.5. Cơ sở dữ liệu ASHRAE RP-1043

Mô hình FDD-KNN được nghiên cứu và xây dựng dựa trên bộ dữ liệu của dự án ASHRAE RP-1043 [12]. Đây là cơ sở dữ liệu được sử dụng rộng rãi trong lĩnh vực FDD cho hệ thống điều hòa không khí trung tâm [2]-[4]. Thiết bị được sử dụng là loại ly tâm có công suất lạnh khoảng 316 kW, môi chất lạnh R134a. Thiết bị ngưng tụ và bay hơi có bộ trao đổi nhiệt dạng ống vỏ, chất tải lạnh là nước chảy trong ống và môi chất lạnh bên ngoài ống. Hệ thống vận hành trong trạng thái bình thường và tiến hành tạo lỗi để ghi nhận dữ liệu. Bảy lỗi độc lập được tiến hành tạo lỗi bao gồm: bám bẩn bình ngưng (ConFoul), dư dầu (ExcsOil), giảm lưu lượng nước bình ngưng (ReduCF), giảm lưu lượng nước bình bay hơi (ReduEF), khí không ngưng (NonCon), thiếu/rò rỉ môi chất lạnh (RefLeak) và dư môi chất lạnh (RefOver). Các thí nghiệm được thực hiện, 64 biến số vật lý của hệ thống được tự động ghi lại sau mỗi 10 giây. Phương pháp tạo lỗi và chi tiết bốn mức độ nghiêm trọng của mỗi lỗi được mô tả trong bảng 2.

Bảng 2. Phương pháp tạo lỗi và các mức độ lỗi

Kí hiệu	Mức độ 1 (LV1)	Mức độ 2 (LV2)	Mức độ 3 (LV3)	Mức độ 4 (LV4)	Phương pháp tạo lỗi
ConFoul	12%	20%	30%	45%	Bịt đầu ống trong bình ngưng
ExcsOil	+ 14%	+ 32%	+ 50%	+ 68%	Nạp thêm dầu
ReduCF	- 10%	- 20%	- 30%	- 40%	Giảm lưu lượng nước
ReduEF	- 10%	- 20%	- 30%	- 40%	Giảm lưu lượng nước
NonCon	+ 1%	+ 1.8%	+ 2.4%	+ 5.6%	Đưa khí nito vào
RefLeak	- 10%	- 20%	- 30%	- 40%	Xả bớt môi chất lạnh
RefOver	+ 10%	+ 20%	+ 30%	+ 40%	Nạp thêm môi chất lạnh

Các biến số vật lý được lựa chọn dựa trên mức độ nhạy cảm đặc trưng với từng lỗi, các tác động chi tiết của lỗi được phân tích như sau:

Lỗi dư dầu: khi lỗi xảy ra lượng dầu trong khoang máy nén nhiều làm tăng tổn thất cơ học nên nhiệt độ dầu TO_{feed} tăng.

Các lỗi về giảm lưu lượng nước: khi lưu lượng nước ngưng tụ giảm, trong khi nhiệt độ nước đi vào và sự thải nhiệt của bình ngưng là không đổi. Kết quả là nhiệt độ nước ra khỏi bình ngưng tăng và độ chênh lệch nhiệt độ đầu ra TCO và đầu vào TCI lớn. Tương tự với lỗi giảm lưu lượng nước bình bay hơi, độ chênh lệch nhiệt độ đầu vào TEI và đầu ra TEO lớn.

Khí không ngưng: khí không ngưng có xu hướng tích tụ trong thiết bị ngưng tụ, nó làm tăng nhiệt độ bão hòa của môi chất lạnh trong bình ngưng TRC.

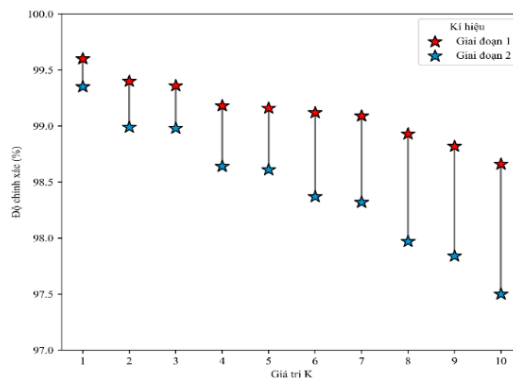
Các lỗi về môi chất lạnh: dư môi chất lạnh làm tăng lượng môi chất trong hệ thống, do đó nhiệt độ bão hòa môi chất trong bình bay hơi TRE và nhiệt độ cuối tâm nén TR_{dis} tăng. Ngược lại khi hệ thống thiếu môi chất lạnh TRE và TR_{dis} giảm.

Bám bẩn bình ngưng: trong hệ thống điều hòa không khí trung tâm giải nhiệt bằng nước, cặn bẩn tích tụ dần dần trong thiết bị ngưng tụ bất kể quy trình xử lý nước được áp dụng. Độ chênh lệch giữa TRC và TCO được mong đợi để có lập lỗi này.

3. Kết quả và thảo luận

3.1. Ảnh hưởng của giá trị K

Trong nghiên cứu này, giá trị K của thuật toán KNN được thay đổi trong phạm vi từ 1 đến 10. Mỗi giá trị K sẽ tham gia vào mô hình FDD-KNN và được ước tính khả năng chẩn đoán lỗi và mức độ lỗi. Kết quả khảo sát được thể hiện trong hình 2.



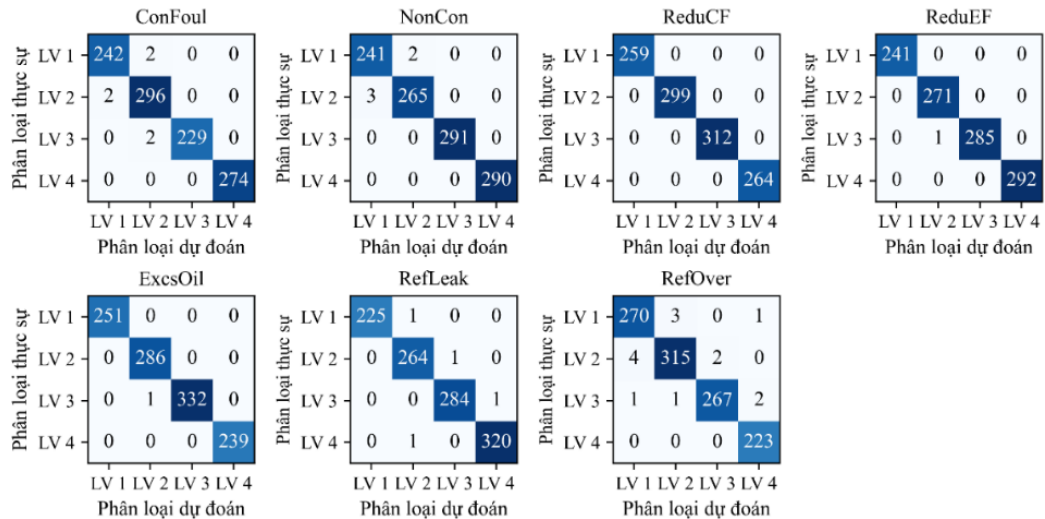
Hình 2. Xác định giá trị K

Hình 2 cho thấy ứng với mỗi giá trị K, hiệu suất mô hình tại giai đoạn 1 cao hơn giai đoạn 2. Nguyên nhân là do giai đoạn 1 phân loại giữa các trạng thái được thực hiện, tuy có một số lỗi có biểu hiện giống nhau nhưng việc lựa chọn các biến số vật lý đảm bảo cho việc phân biệt chúng. Ở giai đoạn 2, mô hình có nhiệm vụ phân biệt các mức độ lỗi. Việc chẩn đoán sẽ gặp khó khăn và dễ nhầm lẫn vì sự thay đổi thông số giữa các mức độ lỗi là rất nhỏ, đặc biệt là những mức độ lỗi nhẹ. Ngoài ra, khi giá trị K tăng hiệu suất mô hình có xu hướng giảm vì khi kích thước vùng lân cận tăng khiến các mẫu dữ liệu thuộc các trạng thái khác nhau có thể cùng tồn tại trong vùng lân cận đó. Do đó, các lỗi có biểu hiện giống nhau được phân biệt bởi khoảng biến thiên của biến số rất dễ bị chẩn đoán nhầm.

Giá trị vùng lân cận là 1 được xác định là giá trị tối ưu của mô hình vì cho hiệu suất chẩn đoán cao nhất. Do đó, mô hình FDD-KNN với K = 1 tiếp tục được thực hiện các đánh giá chi tiết khác để xác định khả năng chẩn đoán.

3.2. Kết quả chẩn đoán

Ma trận nhầm lẫn của mô hình FDD-KNN trong giai đoạn chẩn đoán mức độ lỗi được thể hiện trong hình 3. Trong hình gồm bảy ma trận nhầm lẫn thành phần thể hiện cho từng lỗi.

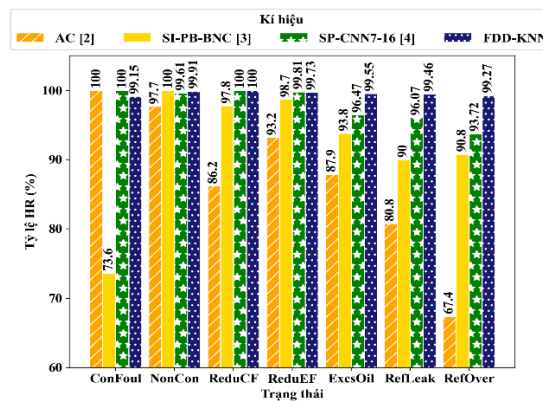


Hình 3. Ma trận nhầm lẫn của mô hình FDD-KNN

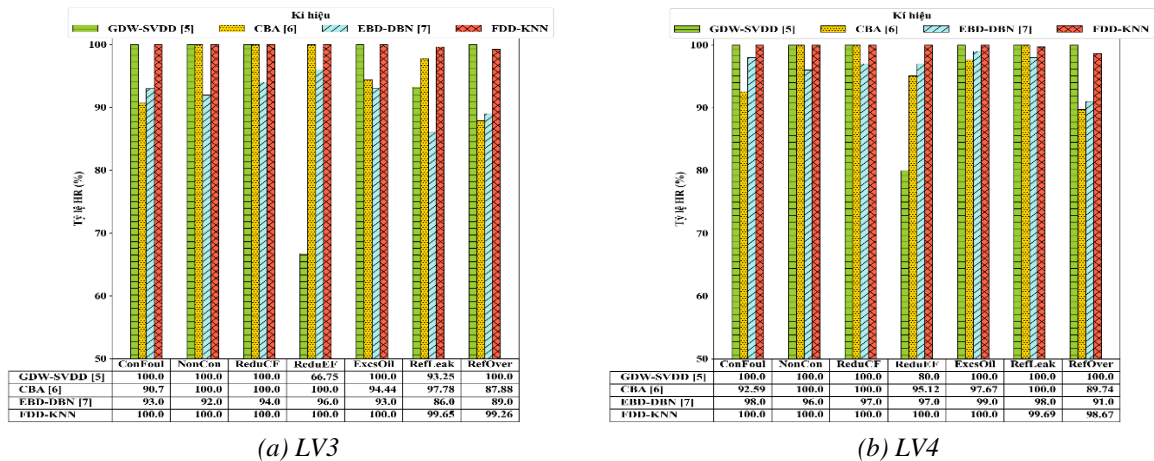
Hình 3 cho thấy các mức độ lỗi gần nhau thường được chẩn đoán nhầm cho nhau, đặc biệt là các mức độ lỗi nhẹ. Không có trường hợp nào của lỗi ReduCF bị chẩn đoán sai về mức độ lỗi. Các lỗi còn lại có hiệu suất chẩn đoán tốt khi số lượng phần tử trên đường chéo chính cao và các phần tử nằm ngoài đường chéo đa số bằng 0 hoặc có giá trị rất nhỏ. Lỗi ConFoul do mất một thời gian dài để tích tụ cặn bẩn và tác động lên các biến số nên có sự chẩn đoán nhầm lẫn giữa nhiều mức độ lỗi như: LV3, LV2 và LV1. Lỗi RefOver có số lượng chẩn đoán sai nhiều nhất, để cải thiện hiệu suất cần bổ sung các biến vật lý nhạy cảm với lỗi này.

3.3. So sánh với các nghiên cứu trước

Hình 4 so sánh tỉ lệ HR giữa mô hình FDD-KNN, mô hình AC [2], mô hình SI-PB-BNC [3] và mô hình SP-CNN7-16 [4]. Mô hình AC có tỉ lệ HR của ConFoul đạt tuyệt đối 100% nhưng các lỗi RefLeak, RefOver chỉ đạt lần lượt 80.8% và 67.4%. Trên thực tế, các lỗi môi chất lạnh dễ xảy ra, đặc biệt là sau khi hệ thống được bảo trì. Một mô hình có tỉ lệ chẩn đoán các lỗi môi chất lạnh cao hơn sẽ đáng tin cậy hơn. Mô hình SI-PB-BNC cho kết quả tỉ lệ HR của lỗi NonCon là 100% nhưng có nhược điểm là lỗi ConFoul có tỉ lệ HR rất thấp chỉ đạt 73.6%. Lỗi ConFoul xuất hiện dần dần trong quá trình vận hành và khó xác định thời điểm lỗi hình thành. Do đó, cần một mô hình FDD có tỷ lệ chẩn đoán của lỗi ConFoul cao hơn. Mô hình SP-CNN7-16 cho kết quả tỉ lệ HR của lỗi ConFoul, NonCon, ReduCF và ReduEF lần lượt là 100%, 99.61%, 100% và 99.81%, tuy nhiên đối với các lỗi ExcsOil, RefLeak và RefOver có tỉ lệ HR thấp hơn và có giá trị lần lượt là 96.47%, 96.07% và 93.72%. Trong khi đó, mô hình FDD-KNN đề xuất có tỉ lệ HR của các lỗi ConFoul, NonCon, ReduCF, ReduEF, ExcsOil, RefLeak, RefOver lần lượt là 99.15%, 99.91%, 100%, 99.73%, 99.55%, 99.46% và 99.27%. Tỉ lệ HR của mô hình FDD-KNN rất cao và đồng đều, điều này cho thấy mô hình này đáng tin cậy để áp dụng cho hệ thống thực tế và khắc phục được nhược điểm của các nghiên cứu trước.



Hình 4. So sánh độ chính xác chẩn đoán lỗi với các nghiên cứu trước



Hình 5. So sánh độ chính xác chẩn đoán mức độ lỗi với các nghiên cứu trước: (a) LV3; (b) LV4

Hình 5 so sánh tỷ lệ HR của mức độ lỗi nghiêm trọng (LV4 và LV3) giữa mô hình FDD-KNN, mô hình GDW-SVDD [5], mô hình CBA [6] và mô hình EBD-DBN [7]. Mô hình GDW-SVDD mặc dù tỷ lệ HR của các lỗi ConFoul, NonCon, ReduCF, ExcsOil, RefOver đạt tỷ lệ chẩn đoán HR rất cao là 100%, tuy nhiên có hiệu suất chẩn đoán lỗi ReduEF thấp chỉ đạt 80% ở LV4 và 66.75% ở LV3. Bên cạnh đó, lỗi RefLeak cũng thể hiện sự suy giảm độ chính xác khi mức độ lỗi giảm, tỷ lệ HR của lỗi này đạt 100% ở LV4 và giảm còn 93.25% ở LV3. Mô hình CBA chẩn đoán chưa tốt đối với các lỗi ConFoul và RefOver khi tỷ lệ HR ở LV3 chỉ đạt lần lượt là 90.7% và 87.88%, đồng thời khả năng chẩn đoán các lỗi ConFoul, ExcsOil, RefLeak và RefOver giảm khi mức độ lỗi giảm. Mô hình EBD-DBN cho thấy tỷ lệ chẩn đoán chính xác giữa các lỗi có sự chênh lệch. Bằng chứng là trong cùng một mức độ lỗi, RefLeak chỉ có tỷ lệ HR là 86% trong khi ReduEF có tỷ lệ HR là 96%. Trong thực tế, các lỗi đều có thể xảy ra nên cần cải thiện tỷ lệ HR và duy trì khả năng chẩn đoán giữa các lỗi để nâng cao độ tin cậy của mô hình. Ngoài ra, việc giảm tỷ lệ HR khi mức độ lỗi giảm thể hiện rõ ở các lỗi RefLeak từ 98% xuống 86% và RefOver từ 91% xuống 89%. Trong khi đó, mô hình FDD-KNN có hiệu suất cao và đồng đều ở cả hai mức độ lỗi LV4 và LV3. Tỷ lệ chẩn đoán các lỗi ConFoul, NonCon, ReduCF, ReduEF và ExcsOil đạt chính xác tuyệt đối ở cả hai mức độ lỗi. Chỉ có lỗi RefOver giảm tỷ lệ HR khi mức độ lỗi giảm nhưng mức độ giảm rất nhỏ (tỷ lệ HR chênh lệch chỉ 0.59%).

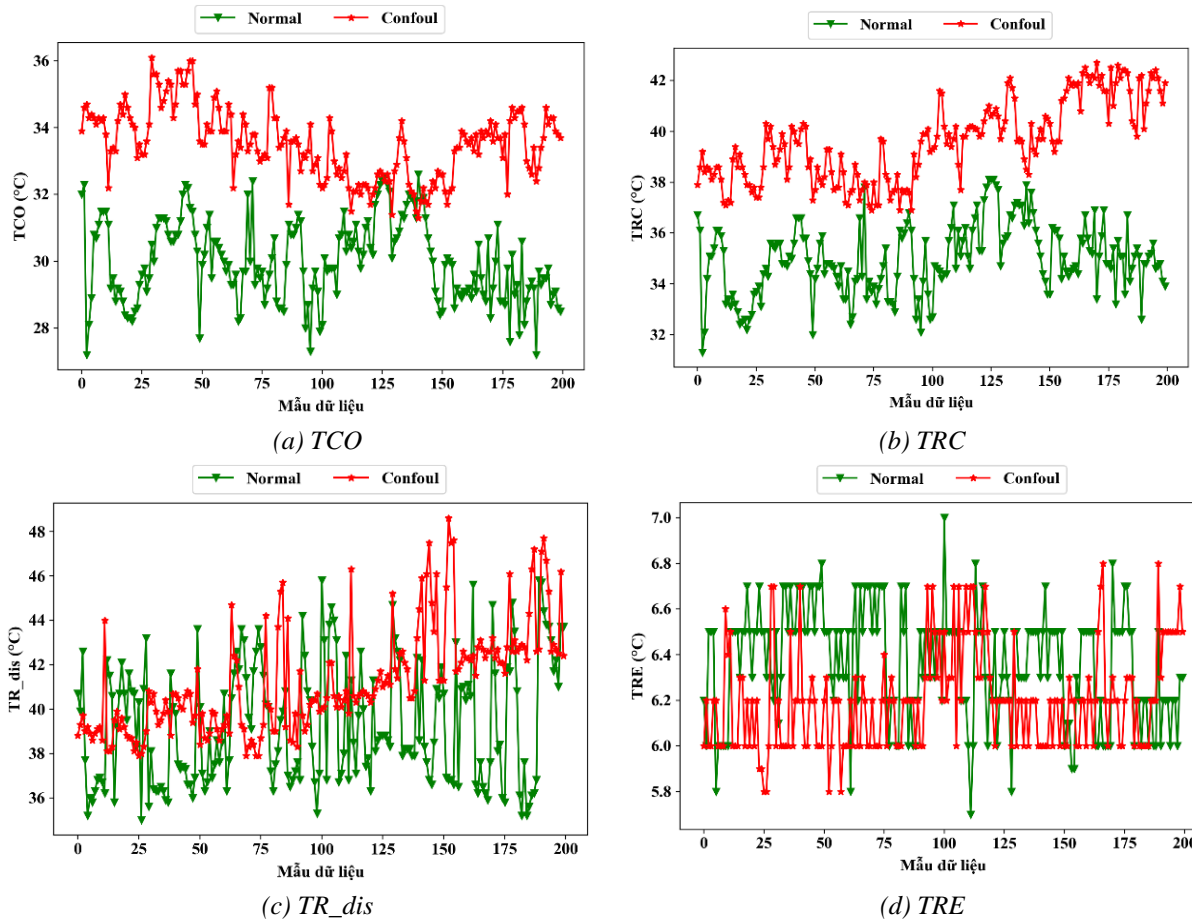
Các kết quả so sánh cho thấy mô hình FDD-KNN có khả năng chẩn đoán chính xác cao và đồng đều giữa các lỗi ứng với giá trị $K=1$. Mô hình này khắc phục được nhược điểm của các nghiên cứu trước và có độ tin cậy cao. Ngoài ra, khả năng chẩn đoán mức độ lỗi của mô hình được duy trì bất kể mức độ nghiêm trọng của lỗi giảm. Điều này còn có ý nghĩa quan trọng trong thực tế vì càng phát hiện sớm các lỗi càng ngăn chặn được hư hỏng cho thiết bị và tránh lãng phí năng lượng.

3.4. Đánh giá bằng dữ liệu thực tế

Trong phần này, hai biến số vật lý TCO và TRC được kiểm chứng xem chúng có thật sự có ý nghĩa trong việc cô lập lỗi bám bẩn bình ngưng hay không. Để thực hiện kiểm tra này, một phép thử về độ dao động của các biến TCO, TRC, TR_dis và TRE ở trạng thái bình thường Normal và bám bẩn bình ngưng ConFoul được thực hiện. Dữ liệu sử dụng kiểm tra là của hệ thống điều hòa không khí trung tâm, giải nhiệt bằng nước của tòa nhà Saigon Centre, thành phố Hồ Chí Minh, Việt Nam được thu thập vào năm 2019 từ hệ thống quản lý tòa nhà. Sự dao động của các biến số vật lý được thể hiện trong hình 6.

Hình 6(a) và hình 6(b) cho thấy phạm vi dao động của TCO và TRC trong hai trạng thái Normal và ConFoul được phân biệt rõ ràng. Do đó, hai biến số này có khả năng thể hiện quy tắc phát hiện lỗi ConFoul. Khi lỗi mới bắt đầu xảy ra, lớp cặn được hình thành với bề dày rất mỏng, sự cản trở truyền nhiệt là không đáng kể nên các biến số thay đổi ít. Trải qua thời gian dài, chiều dày lớp cặn tăng lên, sự cản trở trao đổi nhiệt trong bình ngưng tăng khiến TRC và TCO cao hơn nhiều. Hình 6(c) và hình 6(d) thể hiện dao động của biến số TR_dis và TRE trong hai trạng thái Normal và ConFoul không được phân biệt rõ ràng. Có rất nhiều mẫu dữ liệu của hai trạng thái bị lẫn vào nhau khiến không thể xây dựng

quy luật nào để phân biệt. Kết quả này chứng minh rằng hai biến số TCO và TRC thật sự có ý nghĩa trong việc chẩn đoán lỗi bám bẩn bình ngưng. Đồng thời, các xu hướng thay đổi của biến số khi lỗi xảy ra đã được phân tích trong bài báo là phù hợp với thực tế.



Hình 6. Dao động của các biến số ở trạng thái Normal và ConFoul: (a) TCO; (b) TRC; (c) TR_dis; (d) TRE

4. Kết luận

Nghiên cứu này đã đề xuất một mô hình phát hiện và chẩn đoán lỗi FDD-KNN cho hệ thống điều hòa không khí trung tâm dựa trên thuật toán KNN được đề xuất. Kết quả nghiên cứu cho thấy tại giá trị $K=1$ cho hiệu suất chẩn đoán cao nhất. Kết quả nghiên cứu và so sánh với các mô hình nghiên cứu trước như AC [2], mô hình SI-PB-BNC [3] và mô hình SP-CNN7-16 [4] cho thấy mô hình FDD-KNN với $K=1$ đã khắc phục được nhược điểm của các nghiên cứu trước khi có tỷ lệ chẩn đoán chính xác cao và đồng đều giữa các lỗi, không có lỗi nào được chẩn đoán chính xác dưới 99.15%. Bên cạnh đó, so sánh với GDW-SVDD [5], mô hình CBA [6] và mô hình EBD-DBN [7] ở các lỗi nghiêm trọng LV3 và LV4, khả năng chẩn đoán mức độ lỗi của mô hình FDD-KNN với $K=1$ có đạt kết quả cao nhất. Kết quả kiểm chứng bằng dữ liệu thực của của hệ thống điều hòa không khí trung tâm của tòa nhà Saigon Centre, thành phố Hồ Chí Minh, Việt Nam cho thấy hai biến số TCO và TRC có ý nghĩa trong việc cô lập lỗi bám bẩn bình ngưng. Các xu hướng biến thiên của hai biến số này được phân tích trong bài báo là phù hợp với thực tế.

Lời cảm ơn

Xin trân trọng cảm ơn trường Đại học Sư phạm Kỹ thuật thành phố Hồ Chí Minh đã hỗ trợ nhóm tác giả trong suốt quá trình thực hiện nghiên cứu này.

TÀI LIỆU THAM KHẢO

- [1] S. Katipamula and M. Brambley, "Review Article: Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems—A Review, Part I," *HVAC&R Research*, vol. 11, no. 1, pp. 3-25, 2005.
- [2] R. Huang *et al.*, "An effective fault diagnosis method for centrifugal chillers using associative classification," *Applied Thermal Engineering*, vol. 136, pp. 633-642, 2018.
- [3] S. He, Z. Wang, Z. Wang, X. Gu, and Z. Yan, "Fault detection and diagnosis of chiller using Bayesian network classifier with probabilistic boundary," *Applied Thermal Engineering*, vol. 107, pp. 37-47, 2016.
- [4] J. Gao, H. Han, Z. Ren, and Y. Fan, "Fault diagnosis for building chillers based on data self-production and deep convolutional neural network," *Journal of Building Engineering*, vol. 34, 2021, doi: 10.1016/j.jobee.2020.102043.
- [5] K. Chen, Z. Wang, X. Gu, and Z. Wang, "Multicondition operation fault detection for chillers based on global density-weighted support vector data description," *Applied Soft Computing*, vol. 112, 2021, doi: 10.1016/j.asoc.2021.107795.
- [6] J. Liu *et al.*, "Data-driven and association rule mining-based fault diagnosis and action mechanism analysis for building chillers," *Energy and Buildings*, vol. 216, 2020, doi: 10.1016/j.enbuild.2020.109957.
- [7] Y. Wang, Z. Wang, S. He, and Z. Wang, "A practical chiller fault diagnosis method based on discrete Bayesian network," *International Journal of Refrigeration*, vol. 102, pp. 159-167, 2019.
- [8] N. Settouti, M. E. A. Bechar, and M. A. Chikh, "Statistical Comparisons of the Top 10 Algorithms in Data Mining for Classification Task," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 4, no. 1, 2016.
- [9] Z. Zhang, H. Han, X. Cui, and Y. Fan, "Novel application of multi-model ensemble learning for fault diagnosis in refrigeration systems," *Applied Thermal Engineering*, vol. 164, 2020, doi: 10.1016/j.applthermaleng.2019.114516.
- [10] K. Yan, A. Chong, and Y. Mo, "Generative adversarial network for fault detection diagnosis of chillers," *Building and Environment*, vol. 172, 2020, doi: 10.1016/j.buildenv.2020.106698.
- [11] A. S. Glass, P. Gruber, M. Roos, and J. Todtli, "Qualitative model-based fault detection in air-handling units," *IEEE Control Systems*, vol. 15, no. 4, pp. 11-22, 1995.
- [12] J. E. B. M. C. Comstock, "Development of Analysis Tools for the Evaluation of Fault Detection and Diagnostics in Chillers ASHRAE Research Project RP-1043," *Purdue University, Ray W. Herrick Laboratories, West Lafayette*, 1999.



Le Ha Dong Quan received her B. S in Heat and Refrigeration Technology from Industrial University of Ho Chi Minh City in 2020. She is master student in Thermal Engineering Technology in 2022, Ho Chi Minh City University of Technology and Education, Vietnam. Email: lehadongquan@gmail.com.



Le Minh Nhut is Vice Dean of Faculty of Vehicle and Energy Engineering at Ho Chi Minh City University of Technology and Education, Vietnam. He received his B. S in Heat and Refrigeration Technology from Danang University of Technology in 2003 and M. S degrees (in Thermal Engineering) from Vietnam National University, Ho Chi Minh City University of Technology in 2007, respectively. He then received his Ph.D degree from Jeju National University, Republic of Korea. His fields of interest include refrigeration and air conditioning, thermal systems, and renewable energy. Email: nhutlm@hcmute.edu.vn.