



Tạp chí điện tử
Khoa học và Công nghệ Giao thông
Trang website: <https://jstt.vn/index.php/vn>



Article info

Type of article:

Original research paper

DOI:

<https://doi.org/10.58845/jstt.utt.2026.vn.6.5.268-286>

*Corresponding author:

Email address:

giangnd@huce.edu.vn

Received: 08/07/2026

Received in Revised Form:
22/09/2026

Accepted: 25/09/2026

Edge-Synchronized IIOT Architecture for Multi-Device Industrial Anomaly Monitoring Using LSTM Autoencoder

Bui Duc Phuong¹, Bui The Mong¹, Pham Gia Loc¹, Nguyen Danh Giang²,
Nguyen Danh Huy¹

¹ School of Electronics - Electrical Engineering, Hanoi University of Science and Technology, No.1 Dai Co Viet Street, Bach Mai Ward, Hanoi, Vietnam

² Faculty of Mechanical Engineering, Hanoi University of Civil Engineering, No.55 Giai Phong Street, Bach Mai Ward, Hanoi, Vietnam

Abstract: In the context of industrial digital transformation and smart factories, real-time equipment status monitoring plays a crucial role in minimizing downtime and optimizing predictive maintenance. However, practical industrial systems often face challenges due to limited labeled fault data and significant differences in dynamic characteristics among various equipment types such as pressure pumps, aerodynamic fans, bucket elevators, and vibrating screens. This paper proposes an Industrial Internet of Things (IIoT) architecture that integrates a real-time data synchronization layer at the IoT Edge Gateway to collect data from SCADA systems via industrial protocols such as OPC UA and Modbus RTU. After time synchronization, resampling, and sliding window processing, the data is transmitted to an AI server via the MQTT protocol for anomaly monitoring. Unlike traditional machine learning methods that require large datasets of labeled errors, this study uses a separate LSTM Autoencoder model for each device to learn normal operating characteristics from real-world factory data. The reproduction error of each model is combined with an adaptive thresholding mechanism and time-based monitoring laws to detect anomalies early and reduce false alarms in a multi-device industrial environment. Experimental results on a semi-synthetic dataset constructed from real industrial measurements demonstrate the feasibility of the proposed architecture and its potential as a scalable platform for predictive maintenance in smart factories.

Keywords: Industrial IoT, Smart Factory, Predictive Maintenance, LSTM Autoencoder, Temporal Synchronization.



Thông tin bài viết

Dạng bài viết:

Bài báo nghiên cứu

DOI:

<https://doi.org/10.58845/jstt.utt.2026.vn.6.5.268-286>

***Tác giả liên hệ:**

Địa chỉ Email:

giangnd@huce.edu.vn

Ngày nộp bài: 08/07/2026

Ngày nộp bài sửa: 22/09/2026

Ngày chấp nhận: 25/09/2026

Kiến trúc IIOT đồng bộ thời gian tại biên cho giám sát bất thường trong hệ thống công nghiệp đa thiết bị dựa trên LSTM Autoencoder

Bùi Đức Phương¹, Bùi Thế Mong¹, Phạm Gia Lộc¹, Nguyễn Danh Giang^{2*}, Nguyễn Danh Huy¹

¹Trường Điện - Điện tử, Đại học Bách khoa Hà Nội, Số 1 Đại Cồ Việt, phường Bạch Mai, Hà Nội, Việt Nam

²Khoa Cơ khí, Đại học Xây dựng Hà Nội, Số 55 đường Giải Phóng, phường Bạch Mai, Hà Nội, Việt Nam

Tóm tắt: Trong bối cảnh chuyển đổi số công nghiệp và phát triển nhà máy thông minh, giám sát trạng thái thiết bị theo thời gian thực đóng vai trò quan trọng trong việc giảm thiểu thời gian dừng máy và tối ưu hóa công tác bảo trì dự đoán. Tuy nhiên, các hệ thống công nghiệp thực tế thường gặp nhiều thách thức do thiếu dữ liệu lỗi được gán nhãn và sự khác biệt đáng kể về đặc tính động học giữa các loại thiết bị như bơm áp lực, quạt khí động, gầu tải và sàng rung. Bài báo này đề xuất một kiến trúc Internet vạn vật công nghiệp (IIoT) tích hợp lớp đồng bộ dữ liệu thời gian thực tại cổng biên IoT (IoT Edge Gateway) nhằm thu thập dữ liệu từ hệ thống SCADA thông qua các giao thức công nghiệp. Sau khi thực hiện đồng bộ thời gian, lấy mẫu lại và xử lý bằng cửa sổ trượt, dữ liệu được truyền tới máy chủ AI thông qua giao thức MQTT để phát hiện bất thường. Khác với các phương pháp học máy truyền thống, nghiên cứu này sử dụng một mô hình LSTM Autoencoder riêng cho từng thiết bị nhằm học các đặc trưng vận hành bình thường từ dữ liệu thực tế của nhà máy. Sai số tái tạo của từng mô hình được kết hợp với cơ chế ngưỡng thích nghi và các luật giám sát theo thời gian để phát hiện sớm các bất thường và giảm thiểu cảnh báo giả trong môi trường công nghiệp đa thiết bị. Kết quả thực nghiệm trên bộ dữ liệu bán-tổng-hợp được xây dựng từ dữ liệu đo công nghiệp thực tế cho thấy tính khả thi của kiến trúc đề xuất và tiềm năng mở rộng cho các hệ thống bảo trì dự đoán trong nhà máy thông minh.

Từ khóa: Internet vạn vật công nghiệp, Nhà máy thông minh, Bảo trì dự đoán, Bộ mã hóa tự động LSTM, Đồng bộ thời gian.

1. Giới thiệu

Cuộc cách mạng công nghiệp lần thứ tư (Industry 4.0) đã thúc đẩy sự tích hợp sâu rộng giữa hệ thống điều khiển truyền thống và các nền tảng kết nối thông minh, hình thành nên khái niệm

Mạng lưới Vạn vật Công nghiệp (Industrial Internet of Things - IIoT) [1]. Trong bối cảnh đó, giám sát trạng thái thiết bị theo thời gian thực và bảo trì dự báo (predictive maintenance) trở thành hai yêu cầu cốt lõi của nhà máy thông minh: mỗi giờ ngừng

máy không được dự báo trước đều đồng nghĩa với tổn thất sản xuất đáng kể và chi phí bảo trì tăng đột biến [2]. Các phương pháp phát hiện bất thường truyền thống, chủ yếu dựa trên các quy tắc ngưỡng cố định hoặc các thuật toán học máy có giám sát như Support Vector Machine, đòi hỏi tập dữ liệu lỗi được gán nhãn đầy đủ - điều kiện khó đáp ứng trong môi trường vận hành thực tế do sự cố thiết bị chỉ chiếm tỷ lệ rất nhỏ trong toàn bộ dữ liệu vận hành [3]. Bên cạnh đó, sự không đồng nhất về đặc tính vật lý và chế độ vận hành giữa các loại thiết bị khác nhau - chẳng hạn như máy bơm áp suất, quạt ly tâm, gầu tải và sàng rung - khiến cho một mô hình đơn nhất khó đủ khả năng biểu diễn đồng thời hành vi bình thường của tất cả thiết bị [3, 4].

Trong những năm gần đây, mô hình học sâu không giám sát dựa trên tái tạo (reconstruction-based unsupervised deep learning), đặc biệt là LSTM Autoencoder, đã nổi lên như một hướng tiếp cận hứa hẹn để giải quyết bài toán phát hiện bất thường trong dữ liệu chuỗi thời gian công nghiệp [4, 5]. Nguyên lý cốt lõi của phương pháp này là huấn luyện mô hình trên dữ liệu vận hành bình thường để học cấu trúc động học đặc trưng của thiết bị; khi trạng thái thiết bị lệch khỏi vùng vận hành bình thường, sai số tái tạo tăng vọt và được sử dụng làm chỉ báo bất thường [3, 6]. Nhờ đó, phương pháp này không yêu cầu dữ liệu lỗi có nhãn và có khả năng phát hiện cả những dạng bất thường chưa từng xuất hiện trong lịch sử vận hành.

Tuy nhiên, việc triển khai các mô hình học sâu trong hệ thống IIoT đa thiết bị thực tế vẫn gặp phải nhiều thách thức chưa được giải quyết triệt để. Thứ nhất, kiến trúc đám mây tập trung truyền thống - trong đó toàn bộ dữ liệu cảm biến được truyền lên máy chủ trung tâm để xử lý - thường không đáp ứng được ràng buộc độ trễ dưới một giây của hệ thống giám sát công nghiệp thời gian thực do tích lũy độ trễ truyền thông trên đường WAN [7, 8]. Thứ hai, các luồng dữ liệu thu thập từ nhiều thiết bị qua các giao thức công nghiệp khác nhau - đặc biệt là Modbus RTU và OPC UA -

thường không đồng nhất về thời gian do hiện tượng polling jitter và clock offset, dẫn đến các dao động hình học giả (phantom oscillations) trong cửa sổ đầu vào mô hình nếu không được hiệu chỉnh [8, 9]. Thứ ba, cơ chế ngưỡng phát hiện cố định thông thường (ví dụ: $\mu + 3\sigma$) thiếu tính thích nghi với sự thay đổi điều kiện vận hành theo thời gian, dẫn đến tỷ lệ cảnh báo giả cao trong môi trường nhà máy thực tế [3, 6].

Để giải quyết đồng thời các thách thức trên, bài báo này đề xuất một kiến trúc IIoT tích hợp đồng bộ hóa thời gian tại tầng biên (edge-synchronized IIoT architecture) dành cho giám sát bất thường đa thiết bị công nghiệp sử dụng mô hình LSTM Autoencoder chuyên biệt theo thiết bị. Kiến trúc đề xuất bao gồm ba thành phần cốt lõi. (i) Tầng biên đồng bộ thời gian: Nút biên Advantech ECU-1051 thực hiện toàn bộ chuỗi xử lý tại chỗ bao gồm hiệu chỉnh clock offset bằng ước lượng trung vị, nội suy tuyến tính lên lưới thời gian thống nhất, lọc nhiễu và phân đoạn cửa sổ trượt, đảm bảo dữ liệu đến máy chủ AI đã được căn chỉnh thời gian chính xác. (ii) Mô hình LSTM Autoencoder độc lập theo thiết bị: Một mô hình riêng biệt được huấn luyện cho mỗi thiết bị công nghiệp nhằm học biểu diễn đặc trưng của trạng thái vận hành bình thường, với hàm mất mát tái tạo có trọng số độ tin cậy để giảm ảnh hưởng của các mẫu nội suy chất lượng thấp. (iii) Cơ chế ngưỡng thích nghi và bỏ phiếu theo thời gian: Ngưỡng phát hiện được xây dựng tự động từ phân phối đuôi của sai số tái tạo sử dụng lý thuyết giá trị cực trị (Extreme Value Theory - EVT), kết hợp với luật bỏ phiếu theo cửa sổ thời gian để loại bỏ cảnh báo giả phát sinh từ nhiễu ngắn hạn.

Các đóng góp chính của nghiên cứu này bao gồm:

Đóng góp 1 - Kiến trúc IIoT đồng bộ hóa biên: Đề xuất kiến trúc IIoT ba tầng, trong đó dữ liệu từ Modbus RTU và OPC UA được đồng bộ và tiền xử lý tại nút biên trước khi truyền tới máy chủ AI để phát hiện bất thường, qua đó giảm tải xử lý và nâng cao khả năng mở rộng của hệ thống.

Đóng góp 2 - Mô hình hóa sự bất nhất thời gian và bộ ước lượng clock offset nhất quán: Xây dựng mô hình toán học định lượng sai lệch thời gian đặc trưng cho từng giao thức công nghiệp, và chứng minh tính nhất quán của bộ ước lượng trung vị mẫu cho clock offset trong điều kiện jitter bị chặn.

Đóng góp 3 - LSTM Autoencoder chuyên biệt với hàm mất mát trọng số độ tin cậy: Tích hợp ma trận trọng số độ tin cậy sinh ra từ giai đoạn đồng bộ thời gian trực tiếp vào hàm mất mát tái tạo nhằm giảm mức đóng góp của các mẫu nội suy có độ tin cậy thấp trong quá trình huấn luyện. Hiệu quả của cơ chế này được đánh giá thông qua nghiên cứu triệt tiêu tại Mục 5.4.

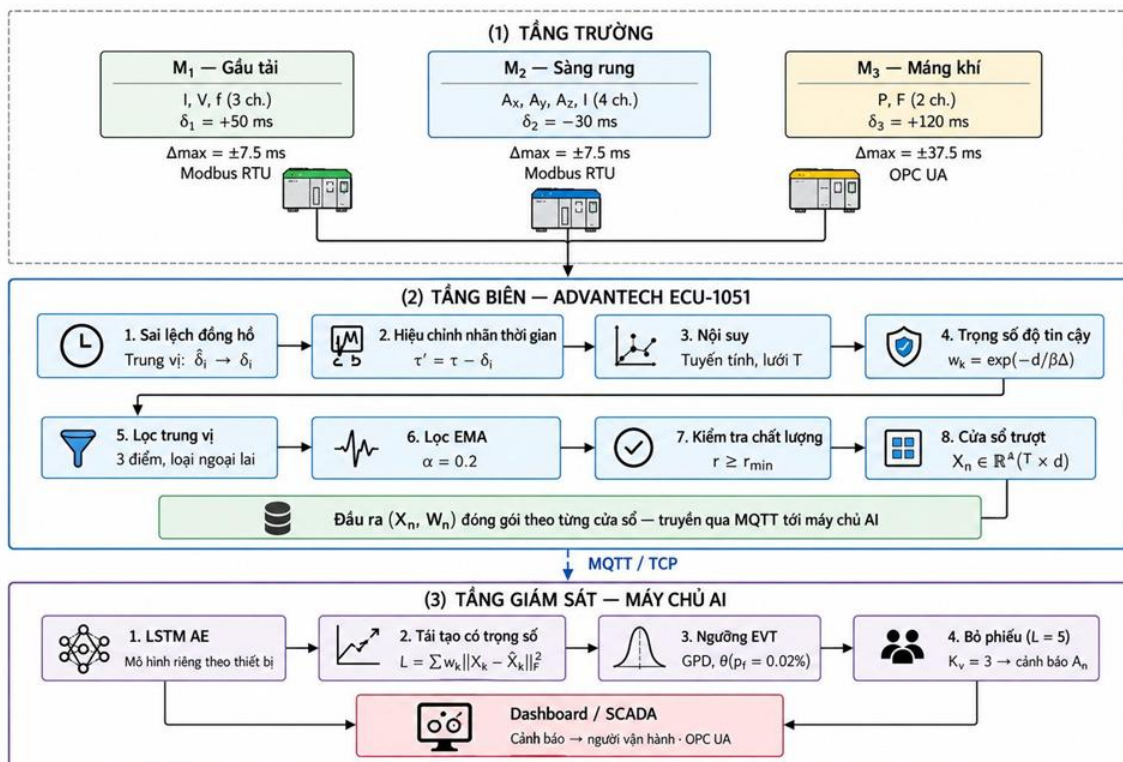
Đóng góp 4 - Cơ chế phát hiện ngưỡng thích nghi EVT + bỏ phiếu thời gian: Kết hợp Generalized Pareto Distribution (GPD) từ lý thuyết giá trị cực trị với cơ chế bỏ phiếu cửa sổ thời gian để đạt được xác suất cảnh báo giả kiểm soát được trong điều kiện tải mạng biến đổi, vượt trội rõ rệt so với ngưỡng tĩnh truyền thống khi điểm vận hành của ngưỡng được hiệu chỉnh phù hợp với đặc tính

phân phối đuôi của sai số tái tạo. Phần còn lại của bài báo được tổ chức như sau. Phần 2 trình bày kiến trúc hệ thống và các thiết bị được giám sát. Phần 3 mô hình hóa sự bất nhất thời gian và trình bày pipeline xử lý tại nút biên. Phần 4 trình bày framework phát hiện bất thường trên nền tảng AI. Phần 5 trình bày kết quả thực nghiệm. Phần 6 trình bày Kết luận và hướng phát triển.

2. Cấu trúc hệ thống

2.1. Kiến trúc hệ thống

Bài toán giám sát bất thường thiết bị công nghiệp trong thời gian thực đặt ra yêu cầu kép: một mặt cần đảm bảo độ trễ quyết định đủ thấp để can thiệp kịp thời trước khi sự cố lan rộng, mặt khác phải hoạt động liên tục trong môi trường phân xưởng với băng thông mạng hạn chế và điều kiện vận hành khắc nghiệt. Kiến trúc đám mây tập trung truyền thống không đáp ứng được đồng thời cả hai ràng buộc này: độ trễ truyền thông tích lũy từ cảm biến đến máy chủ suy luận thường vượt quá ngưỡng thời gian thực 1.000 ms, đồng thời phụ thuộc hoàn toàn vào kết nối mạng WAN [7, 8].



Hình 1. Kiến trúc IIoT được đồng bộ hóa biên ba lớp được đề xuất để theo dõi sự bất thường trong công nghiệp đa thiết bị. Lớp trường thu thập dữ liệu cảm biến qua Modbus RTU và OPC UA; lớp biên (Advantech ECU-1051) thực hiện đồng bộ hóa thời gian và tiền xử lý

Để khắc phục những giới hạn đó, nghiên cứu này áp dụng mô hình biên bản địa (edge-native), trong đó toàn bộ vòng lặp thu thập-tiền xử lý được thực thi ngay tại nút biên đặt trong phân xưởng. Hệ thống được tổ chức theo kiến trúc phân lớp ba tầng như minh họa trong Hình 1.

Thứ nhất, Tầng trường: Bao gồm ba thiết bị công nghiệp dị biệt M1, M2, M3 được trang bị các cụm cảm biến chuyên dụng. Dữ liệu đo lường được thu thập theo chu kỳ polling qua hai giao thức bus công nghiệp phổ biến là Modbus RTU và OPC UA [1]. Thứ hai, Tầng biên: Nút biên Advantech ECU-1051 đảm nhận toàn bộ chuỗi xử lý thời gian thực, bao gồm đồng bộ hóa theo thời gian, lọc tiếng ồn, phân đoạn và đóng gói cửa sổ trượt trước khi chuyển tiếp các bộ căng tính năng đến máy chủ AI giám sát, nơi thực hiện suy luận bất thường và tạo ngưỡng thích ứng. Cuối cùng là tầng giám sát: Máy chủ AI và hệ thống dashboard nhận kết quả cảnh báo đã được lọc từ tầng biên thông qua giao thức MQTT, phục vụ lưu trữ lịch sử và hiển thị cho kỹ sư vận hành. ECU-1051 được lựa chọn làm gateway biên vì đáp ứng đồng thời ba ràng buộc triển khai thực tế: (i) đủ tài nguyên xử lý để thực thi toàn bộ pipeline tiền xử lý tại chỗ - đồng bộ hóa, nội suy, đóng gói cửa sổ trượt - trước khi chuyển tiếp dữ liệu lên AI Server qua MQTT; (ii) tiêu thụ điện năng thấp phù hợp môi trường phân xưởng; (iii) tích hợp trực tiếp Modbus RTU và OPC UA mà không cần gateway bổ sung [7].

2.2. Các thiết bị được giám sát

Hệ thống giám sát đồng thời ba loại thiết bị công nghiệp có đặc tính vật lý, số lượng cảm biến

và chế độ bất thường hoàn toàn khác biệt. Sự dị biệt này là yếu tố cốt lõi thách thức tính tổng quát hóa của bất kỳ giải pháp giám sát thống nhất nào, đồng thời cũng là luận cứ khoa học cho thiết kế đơn mô hình độc lập trên từng thiết bị sẽ được trình bày ở mục 4 [3, 4].

Thiết bị M1 - Gầu tải: Thiết bị vận chuyển vật liệu rời theo phương thẳng đứng, vận hành bởi động cơ điện xoay chiều ba pha. Cụm cảm biến đo điện năng gồm ba kênh: dòng điện pha I (A), điện áp nguồn V (V) và tần số lưới f (Hz). Bất thường điển hình bao gồm kẹt cơ học (biểu hiện qua đột biến dòng điện) và sụt áp nguồn (biểu hiện qua dao động điện áp không đối xứng) [3].

Thiết bị M2 - Sàn rung: Thiết bị phân loại cỡ hạt vật liệu bằng dao động cơ học cưỡng bức. Cụm cảm biến gồm gia tốc kế ba trục Ax, Ay, Az (m/s^2) và dòng điện động cơ I (A). Bất thường điển hình bao gồm mất cân bằng khối lệch tâm (biểu hiện qua bất đối xứng biên độ rung giữa các trục) và mài mòn ổ bi (biểu hiện qua thành phần tần số cao tăng dần trong phổ rung) [10].

Thiết bị M3 - Máng khí: Hệ thống vận chuyển bột liệu bằng luồng khí nén. Cụm cảm biến gồm áp suất khí P (kPa) và lưu lượng khí F (m^3/h). Bất thường điển hình bao gồm tắc nghẽn đường ống (biểu hiện qua tăng áp suất đột ngột kèm sụt lưu lượng) và rò rỉ khí (biểu hiện qua áp suất thấp bất thường khi lưu lượng bơm cấp vẫn ổn định) [6]. Bảng 1 tóm tắt cấu hình cảm biến, giao thức truyền thông và độ lệch đồng hồ hệ thống đi của từng thiết bị, được sử dụng trực tiếp để kiểm chứng thực nghiệm Mệnh đề 1 ở Mục 3.1.

Bảng 1. Cấu hình cảm biến, giao thức truyền thông và đặc điểm jitter của thiết bị được giám sát

Thiết bị	Mô tả	Cảm biến (d _i kênh)	Giao thức	Chu kỳ lấy mẫu	$\Delta_{\max}^{(i)}$ Jitter	Độ lệch đồng hồ δ_i
M ₁	Gầu tải	I, V, f	Modbus RTU	20ms	± 7.5 ms	+50 ms
M ₂	Sàn rung	A _x , A _y , A _z , I (3 kênh)	Modbus RTU	20 ms	± 7.5 ms	-30 ms
M ₃	Máng khí	P, F (4 kênh)	OPC UA	100 ms	± 37.5 ms	+120 ms

3. Pipeline xử lý độ tin cậy tại biên

3.1. Mô hình hóa sự bất nhất thời gian và đồng bộ hóa

Một trong những thách thức nền tảng của hệ thống đa thiết bị dị biệt là sự bất nhất thời gian (temporal inconsistency) phát sinh từ bản chất của các giao thức bus công nghiệp [8, 9]. Khác với các hệ thống đồng bộ hóa thời gian chính xác như IEEE 1588 PTP, Modbus RTU hoạt động theo cơ chế master polling tuần tự: nút biên lần lượt truy vấn từng thiết bị, do đó mốc thời gian ghi nhận tại ECU-1051 tích lũy thêm độ trễ giao tiếp so với thời điểm đo lường vật lý [9]. OPC UA có cơ chế publish/subscribe linh hoạt hơn nhưng đưa vào hệ thống một nguồn jitter khác với biên độ biến thiên cao hơn [7].

Xét luồng dữ liệu thô của thiết bị thứ i

$$S_i = \left\{ \left(x_t^{(i)}, \tau_t^{(i)} \right) \right\} \quad (1)$$

trong đó $x_t^{(i)} \in \mathbb{R}^{d_i}$ là vector tín hiệu cảm biến và $\tau_t^{(i)}$ là nhãn thời gian được ghi nhận tại nút biên. Mô hình sai lệch thời gian thực tế được biểu diễn bởi

$$\tau_t^{(i)} = t_{\text{true}} + \delta_i + \varepsilon_t^{(i)} \quad (2)$$

trong đó δ_i là độ lệch đồng hồ mang tính hệ thống của thiết bị thứ i , còn $\varepsilon_t^{(i)}$ biểu diễn thành phần jitter ngẫu nhiên gây ra bởi cơ chế truyền thông [7, 9].

Giả thiết 1: Trong khoảng thời gian quan sát ngắn, thành phần jitter được giả sử là độc lập, có trung vị bằng không và bị chặn bởi

$$\left| \varepsilon_t^{(i)} \right| \leq \frac{\Delta_{\max}^{(i)}}{2} \\ \varepsilon_t^{(i)} \sim \text{Uniform} \left(-\frac{\Delta_{\max}^{(i)}}{2}, \frac{\Delta_{\max}^{(i)}}{2} \right) \quad (3)$$

Giả thiết này phù hợp với đặc tính của Modbus RTU và OPC UA, trong đó độ trễ truyền thông chỉ dao động trong một khoảng hữu hạn và không xuất hiện hiện tượng trôi đồng hồ đáng kể trong khoảng thời gian ngắn [9]. Với Modbus RTU, biên độ jitter tối đa ước tính $\Delta_{\max}^{(M_1, M_2)} \approx 15$ ms (một khung Modbus 8 byte trên RS-485 ở 115.200 baud

[9]); với OPC UA, $\Delta_{\max}^{(M_3)} \approx 30-150$ ms tùy tải mạng Ethernet [7, 8], giá trị đại diện 75 ms được dùng trong mô phỏng. Mỗi thiết bị còn mang một clock offset δ_i riêng biệt ($\delta_{M_1} = +50$ ms, $\delta_{M_2} = -30$ ms, $\delta_{M_3} = +120$ ms - Bảng 1), tạo điều kiện kiểm chứng thực nghiệm cho bộ ước lượng trung vị dưới đây. Nếu không hiệu chỉnh, hai mẫu từ M_1 và M_3 tưởng như đồng thời có thể lệch nhau tới ~ 170 ms khi tải mạng cao - ở tần số lấy mẫu mục tiêu 50 Hz, tương đương 8-9 mẫu, đủ để tạo dao động hình học giả (phantom oscillations) trong cửa sổ đầu vào và làm tăng vọt tỷ lệ cảnh báo giả [8].

Khác với phương pháp sử dụng một thiết bị cố định làm tham chiếu, nghiên cứu này lựa chọn luồng dữ liệu có độ biến động nhãn thời gian nhỏ nhất làm mốc đồng bộ:

$$\text{ref} = \arg \min \text{Var}(\tau^{(i)}) \quad (4)$$

Việc lựa chọn reference được thực hiện trên các luồng dữ liệu đang ở trạng thái hợp lệ. Khi luồng đang được sử dụng làm reference mất kết nối hoặc không còn đáp ứng tiêu chí chất lượng thời gian, hệ thống cần lựa chọn lại reference từ các luồng còn khả dụng. Do các luồng có thể tồn tại clock offset khác nhau, việc chuyển trực tiếp giữa hai reference có khả năng tạo ra bước nhảy trên trục thời gian. Vì vậy, độ lệch tương đối giữa reference cũ và reference mới cần được ước lượng trên khoảng dữ liệu chồng lấn và được bù trước khi sử dụng reference mới. Các cửa sổ dữ liệu nằm trong giai đoạn chuyển tiếp và không bảo đảm điều kiện đồng bộ được loại bỏ khỏi quá trình tạo đầu vào cho mô hình.

Độ lệch thời gian tương đối của thiết bị thứ i so với reference được ước lượng bằng trung vị mẫu:

$$\widehat{\Delta \delta}_i = \text{median}(\tau_t^{(i)} - \tau_t^{\text{ref}}) \quad (5)$$

trong đó $\Delta \delta_i = \delta_i - \delta_{\text{ref}}$. Việc sử dụng trung vị thay cho trung bình cộng giúp giảm ảnh hưởng của các giá trị ngoại lai do tắc nghẽn mạng hoặc chu kỳ publish bất thường [7].

Mệnh đề 1. Dưới Giả thiết 1, bộ ước lượng trung vị là nhất quán:

$$\widehat{\Delta\delta_i} \xrightarrow{p} \Delta\delta_i \quad (6)$$

Chứng minh. Với giả thiết jitter có trung vị bằng không và bị chặn, thống kê trung vị mẫu hội tụ theo xác suất về trung vị của phân bố sai lệch timestamp khi số lượng mẫu tăng; khi số mẫu đủ lớn, $\widehat{\Delta\delta_i} \rightarrow \Delta\delta_i$. Mệnh đề được kiểm chứng thực nghiệm bằng cách tìm các giá trị $\Delta\delta_i$ đã biết trước cho từng thiết bị (Bảng 1) rồi so sánh với $\Delta\delta_i$ ước lượng được: sai số quan sát dưới 1 ms ở cả ba thiết bị, phù hợp với tính nhất quán đã chứng minh.

Sau khi bù độ lệch đồng hồ, nhãn thời gian hiệu chỉnh là $\tilde{t}_t^{(i)} = t_t^{(i)} - \widehat{\Delta\delta_i}$. Các luồng dữ liệu được ánh xạ lên lưới thời gian chung $T = \{t_k = t_0 + k\Delta\}_{k=0}^{K-1}$ bằng nội suy tuyến tính:

$$\hat{x}^{(i)}(t_k) = x^{(i)}(\tilde{t}_j) + \frac{t_k - \tilde{t}_j}{\tilde{t}_{j+1} - \tilde{t}_j} [x^{(i)}(\tilde{t}_{j+1}) - x^{(i)}(\tilde{t}_j)] \quad (7)$$

trong đó j là chỉ số mẫu gần nhất thỏa $\tilde{t}_j \leq t_k$. Đối với khoảng mất dữ liệu ngắn, cơ chế điền tiền được cho phép tối đa năm mẫu liên tiếp; cửa sổ có số mẫu mất vượt ngưỡng này bị loại bỏ để tránh lan truyền thông tin không đáng tin cậy [4].

Do các điểm nội suy có độ tin cậy khác nhau tùy khoảng cách tới mẫu gốc, mỗi điểm dữ liệu được gán trọng số độ tin cậy

$$w_k^{(i)} = \exp\left(-\frac{|t_k - \tilde{t}_j|}{\beta\Delta}\right), \quad \beta=1 \quad (8)$$

Nhận xét 1. Vì $\exp(-x) \in (0,1]$ với mọi $x \geq 0$, trọng số $w_k^{(i)}$ luôn thỏa $0 < w_k^{(i)} \leq 1$ - tính chất này đảm bảo tính ổn định của quá trình tối ưu ở máy chủ AI và sẽ được dùng để chứng minh hàm mất mát bị chặn ở Mục 4.1. Ma trận trọng số thu được trong quá trình đồng bộ được truyền cùng tensor dữ liệu tới máy chủ AI qua MQTT, giúp các vùng tín hiệu được nội suy nhiều hoặc chất lượng thấp đóng góp ít hơn vào quá trình học của mô hình [3, 4].

3.2. Điều hòa tín hiệu và phân đoạn cửa sổ trượt

Sau khi các luồng dữ liệu được đồng bộ và ánh xạ lên cùng lưới thời gian, tín hiệu vẫn có thể chứa giá trị đột biến, nhiễu cao tần và mẫu bị mất. Một chuỗi bước điều hòa tín hiệu được thực hiện tại nút biên trước khi dữ liệu được gửi tới máy chủ

AI [3, 4].

Trước tiên, bộ lọc trung vị ba điểm loại bỏ các giá trị ngoại lai đơn lẻ mà vẫn duy trì biên dạng tín hiệu:

$$x_t^{(i,med)} = \text{median}(x_{t-1}^{(i)}, x_t^{(i)}, x_{t+1}^{(i)}) \quad (9)$$

So với bộ lọc trung bình trượt, bộ lọc trung vị chống chịu tốt hơn với các xung nhiễu ngẫu nhiên trên đường truyền [8]. Sau đó, tín hiệu được làm trơn bằng bộ lọc trung bình mũ:

$$x_t^{(i,ema)} = \alpha x_t^{(i,med)} + (1-\alpha)x_{t-1}^{(i,ema)} \quad (10)$$

trong đó $\alpha \in (0,1)$ là hệ số làm trơn. Trong nghiên cứu này, giá trị $\alpha=0.2$ được lựa chọn nhằm đạt được sự cân bằng giữa khả năng khử nhiễu và tốc độ đáp ứng của tín hiệu.

Chất lượng mỗi cửa sổ dữ liệu được đánh giá qua tỷ lệ mẫu hợp lệ $r = N_{\text{valid}}/N_{\text{total}}$. Cửa sổ có r nhỏ hơn ngưỡng định trước sẽ bị loại bỏ [4].

Giả thiết 2. Trong điều kiện vận hành bình thường, $r \geq r_{\min}$ - mất gói kéo dài chỉ xuất hiện với xác suất rất nhỏ [11].

Sau kiểm tra chất lượng, dữ liệu được tổ chức thành các cửa sổ trượt độ dài T , tạo tensor đầu vào $X_n = [x_n, x_{n+1}, \dots, x_{n+T-1}] \in \mathbb{R}^{T \times d}$ với bước dịch nhỏ hơn T , cho phép khai thác quan hệ thời gian giữa các mẫu liên tiếp và tăng mật độ quan sát [3, 4]. Ma trận trọng số độ tin cậy từ Mục 3.1 được đóng gói và truyền cùng X_n tới máy chủ AI qua MQTT, bảo toàn thông tin chất lượng dữ liệu xuyên suốt pipeline và tích hợp trực tiếp vào hàm mất mát tái tạo ở Mục 4.1.

3.3. Phân tích độ phức tạp tính toán

Toàn bộ pipeline tiền xử lý tại ECU-1051 được thiết kế để đảm bảo vận hành thời gian thực trên nền tảng có tài nguyên hạn chế [7]. Gọi M là số thiết bị, d là tổng số kênh sau hợp nhất, T là chiều dài cửa sổ trượt và K là số mẫu mỗi chu kỳ xử lý. Vì ước lượng clock offset và nội suy tuyến tính được thực hiện độc lập, tuần tự trên từng mẫu của từng luồng ($O(MK)$), các bộ lọc trung vị/trung bình mũ chỉ yêu cầu số phép toán hằng số trên mỗi mẫu ($O(K)$), và việc tạo cửa sổ trượt chỉ gồm các thao tác cập nhật tuần tự không cần nhân ma trận

lớn ($O(Td)$), độ phức tạp thời gian tổng thể của pipeline biên xấp xỉ:

$$O(MK+Td) \quad (11)$$

Về bộ nhớ, nút biên chỉ cần lưu các mẫu trong cửa sổ quan sát hiện tại, nên yêu cầu bộ nhớ cũng là $O(Td)$, không phụ thuộc tổng thời gian vận hành hệ thống. Với cấu hình thực nghiệm ($M=3$, $d=9$, $T=50$), tải tính toán tại ECU-1051 tăng tuyến tính theo kích thước dữ liệu và thấp hơn đáng kể so với chi phí suy luận của LSTM Autoencoder - vốn được chuyển hoàn toàn sang máy chủ AI [3, 4]. Do mọi bước xử lý đều tuyến tính, kiến trúc đề xuất có khả năng mở rộng tốt khi số lượng thiết bị giám sát tăng lên.

4. Framework phát hiện bất thường trên nền tảng AI

4.1. LSTM Autoencoder chuyên biệt theo thiết bị

Sau tiền xử lý tại biên, các tensor đã đồng bộ và phân đoạn được truyền tới máy chủ AI để phát hiện bất thường. Do đặc tính vận hành, dải giá trị và động học của từng thiết bị khác biệt đáng kể, một mô hình duy nhất cho toàn hệ thống có thể làm suy giảm khả năng biểu diễn trạng thái bình thường của từng thiết bị riêng lẻ [3, 4, 12,13,14]. Vì vậy, một mô hình LSTM Autoencoder độc lập được huấn luyện cho mỗi thiết bị [5, 6].

Với tensor đầu vào $X_n^{(i)} \in \mathbb{R}^{T \times d_i}$ của thiết bị thứ i , khối encoder gồm hai lớp LSTM tuần tự nén chuỗi tín hiệu thành vector đặc trưng số chiều thấp:

$$\begin{aligned} h_t^{(i)} &= \text{LSTM}_{\text{enc}}(x_t^{(i)}, h_{t-1}^{(i)}) \\ z^{(i)} &= f_{\text{enc}}(X_n^{(i)}) \in \mathbb{R}^{d_z} \end{aligned} \quad (12)$$

Giá trị $d_z = 8$ được chọn để duy trì tỷ lệ nén > 10 lần, tránh ánh xạ đồng nhất (identity mapping) - hiện tượng mô hình chỉ sao chép dữ liệu đầu vào mà không học được đặc trưng động học có ý nghĩa [3, 5, 15]. Khối decoder tái tạo lại chuỗi tín hiệu ban đầu:

$$\hat{X}_n^{(i)} = f_{\text{dec}}(z^{(i)}) \quad (13)$$

Khác với MSE truyền thống [4, 5], hàm mất mát sử dụng trực tiếp trọng số độ tin cậy $w_k^{(i)}$ từ Mục 3.1 để giảm ảnh hưởng của các mẫu được nội

suy nhiều hoặc chất lượng thấp. Với sai số tái tạo

$e_k^{(i)} = \|X_k^{(i)} - \hat{X}_k^{(i)}\|_F^2$, hàm mất mát tái tạo có trọng số là:

$$L^{(i)} = \frac{1}{T} \sum_{k=1}^T w_k^{(i)} e_k^{(i)} \quad (14)$$

Nhờ đó, các mẫu được nội suy từ khoảng thời gian xa dữ liệu gốc đóng góp ít hơn vào quá trình tối ưu, hạn chế ảnh hưởng của sai số truyền thông tới quá trình học [8, 9].

Nhận xét 2. Do $0 < w_k^{(i)} \leq 1$ (Nhận xét 1) và sai số tái tạo trong điều kiện vận hành bình thường có kỳ vọng hữu hạn - hợp lý vì tín hiệu vận hành bình thường bị chặn về vật lý và không xuất hiện giá trị phân kỳ [3, 4], ta có $L^{(i)} = \frac{1}{T} \sum_k w_k^{(i)} e_k^{(i)} \leq \frac{1}{T} \sum_k e_k^{(i)} < \infty$: hàm mất mát tái tạo có trọng số luôn bị chặn trên và không gây mất ổn định trong quá trình tối ưu.

Kiến trúc đề xuất cho phép thông tin chất lượng dữ liệu được bảo toàn xuyên suốt chuỗi xử lý: trọng số độ tin cậy sinh ra ở giai đoạn đồng bộ thời gian tiếp tục được tích hợp trực tiếp vào quá trình học nhằm giảm ảnh hưởng của các mẫu nội suy có độ tin cậy thấp. [8, 9].

4.2. Ngưỡng thích nghi và quyết định theo thời gian

Sau khi LSTM Autoencoder được huấn luyện trên dữ liệu vận hành bình thường, sai số tái tạo đóng vai trò chỉ báo mức độ bất thường [4, 5]. Do đặc tính biến thiên theo thời gian của hệ thống công nghiệp, một ngưỡng cố định thường không đủ linh hoạt để phản ánh thay đổi điều kiện vận hành [3, 6]. Nghiên cứu này dùng lý thuyết giá trị cực trị (Extreme Value Theory - EVT) kết hợp cơ chế bỏ phiếu theo thời gian để xây dựng ngưỡng thích nghi và hạn chế cảnh báo giả do nhiễu tức thời [3, 6].

Với mỗi cửa sổ quan sát thứ n , điểm bất thường là

$$a_n^{(i)} = \frac{1}{T} \sum_{k=1}^T w_k^{(i)} \|X_k^{(i)} - \hat{X}_k^{(i)}\|_F^2 \quad (15)$$

Ngưỡng sơ cấp u được chọn bằng phân vị

thứ p_u của tập huấn luyện: $u=Q_{p_u}(a)$, với $p_u=99.5\%$ trong cấu hình cuối cùng (xem thảo luận lựa chọn tham số ở Mục 5.4). Các tham số p_u và p_f được xác định tại giai đoạn hiệu chỉnh và không được giả thiết là cố định cho mọi điều kiện vận hành. Khi phân bố sai số tái tạo của trạng thái bình thường thay đổi, ngưỡng sơ cấp u và các tham số GPD cần được ước lượng lại trên tập dữ liệu nền mới để cập nhật ngưỡng phát hiện θ .

Trong nghiên cứu hiện tại, quá trình hiệu

Bảng 2. Tóm tắt bộ dữ liệu bán-tổng-hợp được sử dụng trong nghiên cứu

Thiết bị	Cảm biến	Giao thức	Mẫu hiệu chỉnh	Mẫu kiểm thử	Mẫu bất thường
M1	I,V,f	Modbus RTU	86,400	518,351	2,641 (0.51%)
M2	A_x,A_y,A_z,I	Modbus RTU	86,400	518,351	3,378 (0.65%)
M3	P,F	OPC UA	85,438*	518,351	1,338 (0.26%)

*M3: 962 mẫu lẫn bất thường trong khoảng hiệu chỉnh bị lọc bỏ trước khi huấn luyện.

Theo định lý Pickands-Balkema-de Haan, phân bố các phần vượt ngưỡng được xấp xỉ bằng phân bố Pareto tổng quát (GPD) [3, 6]:

$$F(y)=1-\left(1+\frac{\xi y}{\sigma}\right)^{-1/\xi}, \quad y>0 \quad (16)$$

với tham số tỷ lệ $\sigma>0$ và tham số hình dạng ξ ước lượng bằng Maximum Likelihood Estimation. Ngưỡng phát hiện thích nghi tương ứng xác suất báo động giả mong muốn p_f là:

$$\theta=u+\frac{\sigma}{\xi}\left[\left(\frac{N}{N_u}p_f\right)^{-\xi}-1\right] \quad (17)$$

trong đó N là tổng số mẫu huấn luyện và N_u là số mẫu vượt ngưỡng sơ cấp u . Nhờ đó, ngưỡng phát hiện được xác định trực tiếp từ đặc tính thống kê phần đuôi của sai số tái tạo thay vì một giá trị kinh nghiệm cố định [3, 6].

Giả thiết 4. Các phần vượt ngưỡng sau declustering được giả sử độc lập và đồng phân bố; dưới giả thiết này, xác suất báo động giả được kiểm soát xấp xỉ bởi p_f đã chọn.

Mặc dù ngưỡng EVT giảm ảnh hưởng của thay đổi điều kiện vận hành, các dao động ngắn hạn hoặc nhiễu truyền thông vẫn có thể tạo cảnh báo đơn lẻ [7, 8]. Cơ chế bỏ phiếu theo thời gian được áp dụng để nâng cao độ tin cậy quyết định

chỉnh này được thực hiện ngoại tuyến; cơ chế cập nhật trực tuyến chưa được xem xét. Do các giá trị vượt ngưỡng liên tiếp thường có tương quan mạnh theo thời gian (đặc tính của LSTM và tính liên tục của quá trình công nghệ) [4, 6], phương pháp declustering được áp dụng trước khi ước lượng phân bố cực trị: với mỗi cụm vượt ngưỡng liên tiếp $C_j=\{a_t|a_t>u\}$, giá trị đại diện là $Y_j=\max_{a_t\in C_j}(a_t-u)$, chỉ giữ lại giá trị lớn nhất trong mỗi cụm để tập dữ liệu cực trị gần thỏa giả thiết độc lập.

cuối cùng [6]. Biến quyết định nhị phân $b_n=1[a_n>\theta]$; với cửa sổ bỏ phiếu L quyết định liên tiếp, tín hiệu cảnh báo cuối cùng là:

$$A_n=1\left[\sum_{k=n-L+1}^n b_k\geq K_v\right] \quad (18)$$

trong đó K_v là số phiếu tối thiểu để kích hoạt cảnh báo ($L=5$, $K_v=3$ trong nghiên cứu này). Cơ chế này loại bỏ các cảnh báo tức thời gây ra bởi nhiễu ngẫu nhiên, chỉ coi những bất thường kéo dài trong nhiều cửa sổ liên tiếp là sự cố thực sự, cho phép hệ thống duy trì độ nhạy cao trong khi kiểm soát xác suất báo động giả [3, 6].

4.3. Phân tích độ trễ đầu cuối

Ngoài độ chính xác phát hiện, khả năng đáp ứng thời gian thực là yêu cầu quan trọng của hệ thống giám sát công nghiệp [1, 7]. Dữ liệu được xử lý tuần tự qua các giai đoạn đồng bộ thời gian, điều hòa tín hiệu, truyền MQTT, suy luận LSTM Autoencoder và ra quyết định. Độ trễ đầu cuối là tổng độ trễ từng thành phần:

$$T_{\text{total}}=T_{\text{sync}}+T_{\text{filter}}+T_{\text{MQTT}}+T_{\text{LSTM}}+T_{\text{decision}} \quad (19)$$

Theo Mục 3.3, các tác vụ tại nút biên có độ phức tạp tuyến tính và chỉ yêu cầu ít phép toán trên mỗi chu kỳ lấy mẫu, trong khi suy luận học sâu được chuyển sang máy chủ AI để giảm tải cho ECU-1051 [7]. Với chiều dài cửa sổ T , số chiều đầu

vào d_i và số nút ẩn LSTM H, chi phí suy luận xấp xỉ $O(Td_iH)$ [4, 5] - tăng tuyến tính theo chiều dài chuỗi và số nút ẩn. Các phép toán liên quan ngưỡng EVT và luật bỏ phiếu chỉ gồm cộng và so sánh đơn giản nên $T_{\text{decision}} \ll T_{\text{LSTM}}$ và có thể bỏ qua trong độ trễ tổng.

Nhận xét 3. Vì mỗi thành phần trong (19) đều bị chặn trên bởi một hằng số hữu hạn ($T_{\text{sync}} \leq c_1, \dots, T_{\text{decision}} \leq c_5$), độ trễ đầu cuối cũng bị chặn trên: $T_{\text{total}} \leq c_1 + c_2 + c_3 + c_4 + c_5 < \infty$. Việc phân tách tác vụ tiền xử lý tại tầng biên và suy luận học sâu tại máy chủ AI giúp giảm tải tính toán cho ECU-1051 và duy trì khả năng mở rộng của kiến trúc. Tuy nhiên, giới hạn $T_{\text{total}} < 1s$ cần được xác nhận bằng phép đo độ trễ đầu cuối trên hệ thống triển khai thực tế. Trong nghiên cứu hiện tại, số liệu đo đầy đủ cho từng thành phần của T_{total} chưa được thu thập [7, 8].

5. Thiết lập thí nghiệm và thảo luận

5.1. Bộ dữ liệu thực nghiệm

Do hạn chế về thời gian triển khai hạ tầng truyền thông MQTT thực tế tại thời điểm thực hiện nghiên cứu, bộ dữ liệu thực nghiệm được xây dựng theo phương pháp bán-tổng-hợp (semi-synthetic), dựa trên một đoạn dữ liệu thật thu thập liên tục từ hệ thống giám sát của ba thiết bị công nghiệp M1, M2, M3. Ngày đầu tiên (24 giờ, 86,400 mẫu/thiết bị, không chứa bất thường) được sử dụng làm tập hiệu chỉnh (calibration set) để huấn luyện mô hình LSTM Autoencoder và fit ngưỡng EVT; phần dữ liệu còn lại, gồm 518,351 mẫu/thiết bị tương ứng xấp xỉ sáu ngày vận hành liên tục, được sử dụng làm tập kiểm thử. Việc mở rộng cỡ mẫu hiệu chỉnh lên 86,400 nhằm đảm bảo số lượng điểm vượt ngưỡng (N_u) đủ lớn để ước lượng ổn định các tham số (σ, ξ) của phân bố GPD, tránh hiện tượng bất ổn của ngưỡng EVT khi cỡ mẫu hiệu chỉnh quá nhỏ.

Trên nền dữ liệu thật này, tín hiệu được mở rộng bằng mô hình nhiễu Gauss quanh giá trị thống kê gốc của từng thiết bị, đồng thời nhân thời gian được thêm jitter và clock offset theo đúng mô hình Phương trình 2 và 3 với các tham số cụ thể ở Bảng

1, nhằm tạo ra tập dữ liệu đủ lớn để huấn luyện và đánh giá mô hình trong điều kiện có kiểm soát [4]. Để đảm bảo tính ổn định thống kê của kết quả, toàn bộ thực nghiệm được lặp lại với 5 giá trị seed ngẫu nhiên khác nhau; các số liệu được báo cáo dưới dạng $\text{mean} \pm \text{std}$. Các sự kiện bất thường được thêm có chủ đích vào tín hiệu theo các kịch bản lỗi vật lý đã mô tả ở Mục 2.2 (kẹt cơ học, sụt áp nguồn, mất cân bằng khối lệch tâm, mài mòn ổ bi, tắc nghẽn đường ống, rò khí), với nhãn ground-truth được xác định chính xác tuyệt đối theo thiết kế thêm lỗi.

Bảng 2 tóm tắt cấu hình của bộ dữ liệu được sử dụng trong các thực nghiệm tiếp theo. Cần nhấn mạnh rằng phương pháp này chưa thay thế hoàn toàn việc đánh giá trên dữ liệu lỗi thực tế dài hạn với nhãn xác nhận đa nguồn (nhật ký bảo trì, SCADA, chuyên gia cơ điện); đây là hướng mở rộng trong các nghiên cứu tiếp theo khi hạ tầng truyền thông MQTT/Edge thực tế được triển khai đầy đủ [7]. Trong phạm vi bài báo này, cách tiếp cận bán-tổng-hợp cho phép đánh giá định lượng, có kiểm soát và có thể lặp lại đối với khả năng phát hiện bất thường của thuật toán đề xuất.

5.2. Các phương pháp đối sánh

Để đánh giá khách quan hiệu quả của kiến trúc đề xuất, nghiên cứu này tiến hành so sánh với nhiều phương pháp phát hiện bất thường đại diện cho các hướng tiếp cận khác nhau, bao gồm cả các thuật toán học máy truyền thống và các mô hình học sâu dựa trên tái tạo.

One-Class Support Vector Machine (OCSVM): OCSVM là một trong những phương pháp phát hiện bất thường cổ điển được sử dụng rộng rãi trong các bài toán giám sát công nghiệp [3]. Trong nghiên cứu này, các đặc trưng miền thời gian bao gồm giá trị trung bình, độ lệch chuẩn, giá trị cực đại, giá trị cực tiểu và năng lượng tín hiệu được trích xuất từ từng cửa sổ quan sát trước khi được đưa vào bộ phân loại OCSVM.

Isolation Forest (IForest): Isolation Forest là phương pháp dựa trên cây ngẫu nhiên với khả năng phát hiện các điểm dữ liệu hiếm thông qua cơ chế phân hoạch ngẫu nhiên [3]. Tương tự

OCSVM, mô hình được huấn luyện trên tập đặc trưng thống kê miền thời gian nhằm tạo ra một đối sánh đại diện cho nhóm phương pháp học máy truyền thống.

LSTM Autoencoder không trọng số độ tin cậy (w/o Reliability Weight): Mô hình này sử dụng cùng kiến trúc và cùng dữ liệu đầu vào đã đồng bộ thời gian như phương pháp đề xuất, nhưng thay trọng số $w_k^{(i)}$ trong hàm mất mát (Phương trình 14) bằng giá trị đồng nhất $w_k^{(i)}=1$. Đối sánh này tách biệt đúng ảnh hưởng riêng của thành phần trọng số độ tin cậy, độc lập với bước đồng bộ thời gian.

LSTM Autoencoder với nội suy giữ bậc không (ZOH Resampling AE): Để khảo sát ảnh hưởng của phương pháp nội suy, nghiên cứu này thay thế nội suy tuyến tính bằng phương pháp giữ bậc không (Zero-Order Hold -- ZOH). Trong phương pháp này, giá trị gần nhất trước thời điểm lấy mẫu được duy trì cho đến khi xuất hiện mẫu mới [8]. Các thành phần còn lại của hệ thống được giữ nguyên.

LSTM Autoencoder với ngưỡng tĩnh (Static-threshold AE): Bên cạnh các phương pháp đối sánh trên, một mô hình LSTM Autoencoder sử dụng ngưỡng tĩnh truyền thống $\theta=\mu+3\sigma$ cũng được xây dựng nhằm đánh giá lợi ích của cơ chế ngưỡng thích nghi dựa trên lý thuyết giá trị cực trị

(EVT) [3, 6], trong đó μ và σ là giá trị trung bình và độ lệch chuẩn của sai số tái tạo trên tập huấn luyện.

5.3. Các chỉ số đánh giá

Hiệu năng của hệ thống phát hiện bất thường được đánh giá thông qua nhiều chỉ số khác nhau nhằm phản ánh đồng thời khả năng phát hiện chính xác, mức độ báo động giả và khả năng đáp ứng thời gian thực [3, 6]. Các chỉ số được sử dụng bao gồm Precision, Recall, F1-score, diện tích dưới đường cong ROC (AUC-ROC), diện tích dưới đường cong Precision--Recall (PR-AUC), tỷ lệ báo động giả (False Alarm Rate -- FAR) và độ trễ suy luận.

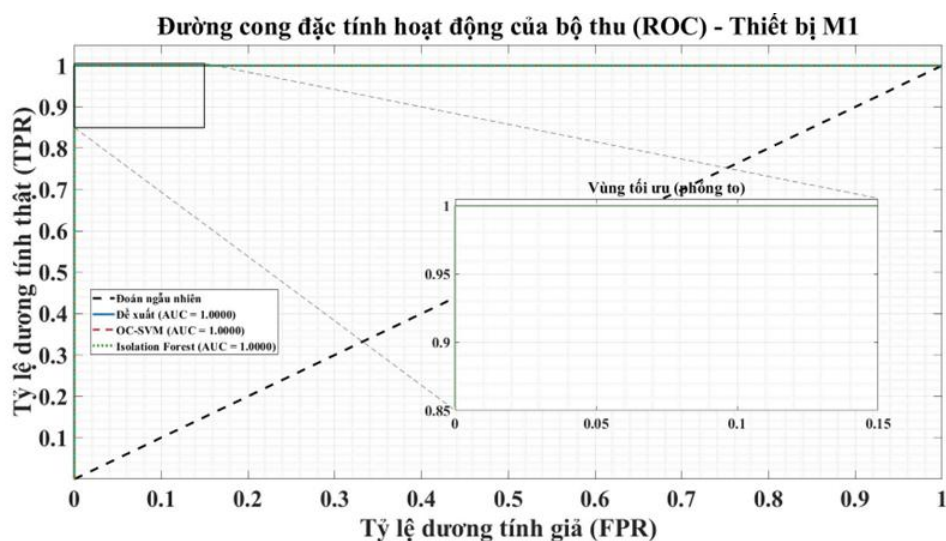
Gọi TP, TN, FP và FN lần lượt là số lượng mẫu dương tính đúng, âm tính đúng, dương tính giả và âm tính giả. Chỉ số Precision (Phương trình 20) để phản ánh tỷ lệ các cảnh báo được phát hiện thực sự tương ứng với các bất thường xảy ra trong hệ thống.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (20)$$

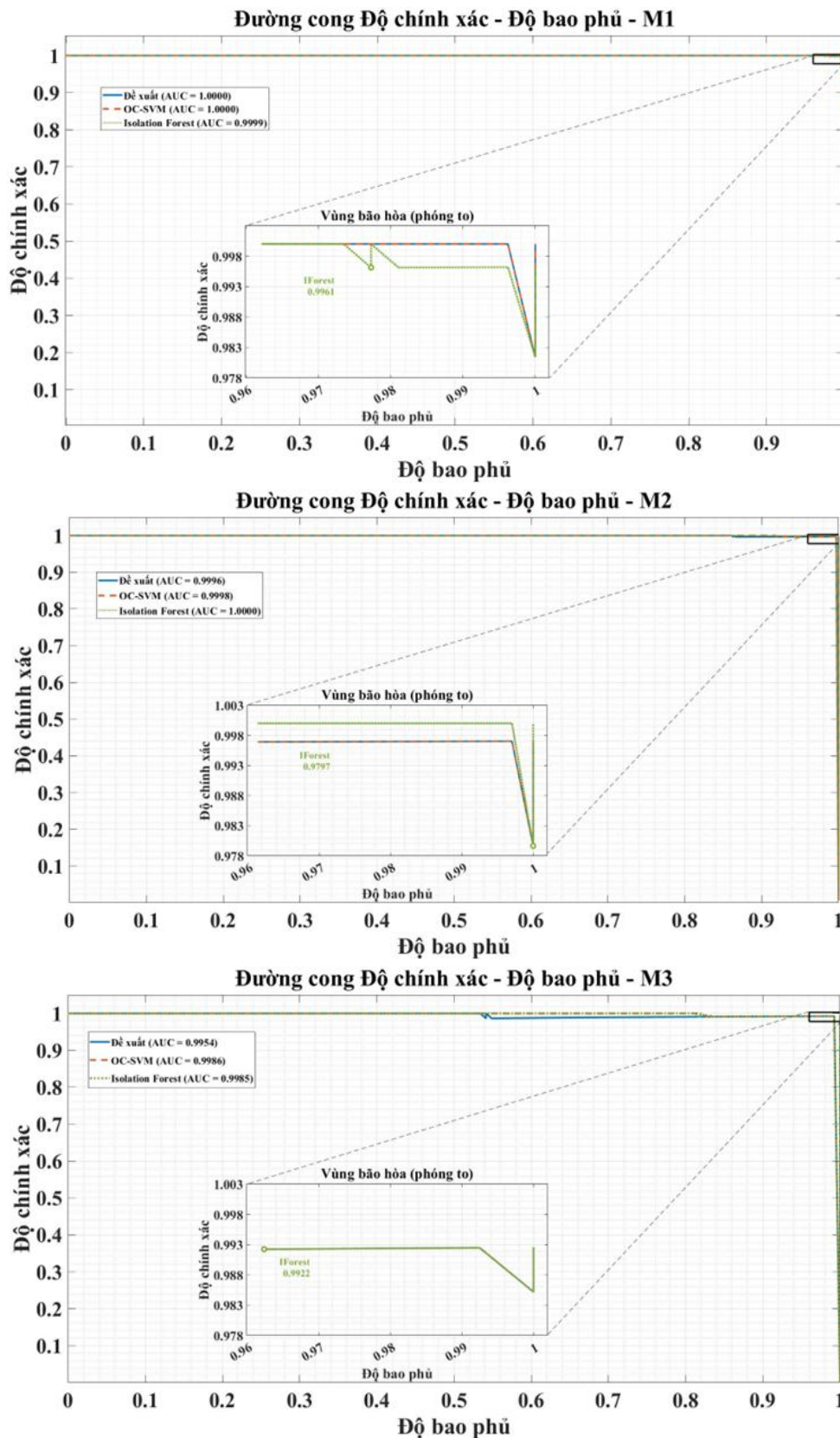
Khả năng phát hiện đầy đủ các sự kiện bất thường được đánh giá thông qua chỉ số Recall:

$$\text{Recall} = \frac{TP}{TP+FN} \quad (21)$$

trong đó giá trị lớn cho thấy mô hình ít bỏ sót các sự cố quan trọng.



Hình 2. Đường cong ROC so sánh giữa phương pháp đề xuất, các biến thể ablation và các phương pháp đối sánh trên ba thiết bị giám sát. AUC-ROC bão hòa ≈ 1.000 ở hầu hết các phương pháp do đặc tính dễ tách biệt của bất thường tiềm tổng hợp.



Hình 3. Đường cong Độ chính xác-Độ bao phủ tương ứng, dùng làm chỉ số phân biệt chính giữa các phương pháp trong điều kiện mất cân bằng nặng. PR-AUC của phương pháp đề xuất đạt 1.0000, 0.9990 và 0.9947 lần lượt tại M1, M2 và M3 (đường chấm ngang biểu diễn tỷ lệ bất thường nền của từng thiết bị). Điểm vận hành được đánh dấu bằng ký hiệu ◦ tương ứng với F1 cao nhất đạt được

Do Precision và Recall thường tồn tại sự đánh đổi lẫn nhau, nghiên cứu này sử dụng F1-score (công thức 22) như một chỉ số tổng hợp nhằm đánh giá cân bằng giữa khả năng phát hiện và mức độ báo động giả [3, 6].

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (22)$$

Bên cạnh đó, diện tích dưới đường cong đặc tính hoạt động của bộ phân loại (AUC-ROC) được sử dụng để đánh giá khả năng phân tách giữa trạng thái bình thường và trạng thái bất thường trên toàn bộ miền giá trị của ngưỡng quyết định [3, 5]. Tuy nhiên, do các bất thường tiềm trong bộ dữ liệu bán-tổng-hợp có biên độ đủ lớn để dễ dàng tách biệt, AUC-ROC đạt giá trị bão hòa ≈ 1.000 ở hầu hết các phương pháp được so sánh (xem Hình 2), khiến chỉ số này kém hữu ích để phân biệt hiệu năng tương đối giữa các phương pháp. Vì vậy, nghiên cứu này bổ sung PR-AUC (Phương trình 23).

$$\text{PR-AUC} = \int_0^1 \text{Precision}(\text{Recall}) d(\text{Recall}) \quad (23)$$

PR-AUC vốn nhạy hơn với hiệu năng trên lớp thiểu số (bất thường) trong điều kiện mất cân bằng nặng [3, 6], và được dùng làm chỉ số bổ trợ không phụ thuộc ngưỡng quyết định, song song với F1-score (phụ thuộc ngưỡng) trong các bảng kết quả. Kết quả đường PR-AUC được thể hiện trong Hình 3.

Đối với các hệ thống giám sát công nghiệp, tỷ lệ báo động giả đóng vai trò đặc biệt quan trọng do các cảnh báo sai liên tiếp có thể làm tăng chi phí bảo trì và gây gián đoạn quá trình sản xuất [2, 6]. Vì vậy, chỉ số False Alarm Rate được xác định

theo Phương trình 24:

$$\text{FAR} = \frac{\text{FP}}{\text{FP} + \text{TN}} \quad (24)$$

trong đó giá trị nhỏ tương ứng với mức độ ổn định cao hơn của hệ thống phát hiện.

Ngoài các chỉ số về độ chính xác, khả năng đáp ứng thời gian thực của hệ thống cũng được đánh giá thông qua độ trễ suy luận trung bình (Phương trình 25):

$$\text{Latency} = \frac{1}{N} \sum_{i=1}^N T_i \quad (25)$$

trong đó T_i là thời gian xử lý của mẫu thứ i và N là tổng số mẫu kiểm thử. Các chỉ số trên được tính toán trên cùng tập kiểm thử và được áp dụng đồng nhất cho tất cả các phương pháp đối sánh nhằm đảm bảo tính công bằng của quá trình đánh giá [3]. Trong nghiên cứu này, F1-score và PR-AUC được sử dụng làm các chỉ số chính để so sánh hiệu năng phát hiện giữa các mô hình, trong khi FAR được sử dụng để đánh giá khả năng kiểm soát cảnh báo giả. Độ trễ đầu cuối của hệ thống được phân tích về mặt cấu trúc tại Mục 4.3; việc đo và đánh giá thực nghiệm đầy đủ độ trễ xử lý trên hệ thống triển khai thực tế được để lại cho nghiên cứu tiếp theo.

5.4. Thiết lập nghiên cứu triệt tiêu

Nhằm đánh giá định lượng mức độ đóng góp của từng thành phần trong kiến trúc đề xuất, nghiên cứu này tiến hành một loạt thí nghiệm triệt tiêu (ablation study) bằng cách loại bỏ hoặc thay thế từng module riêng lẻ trong khi giữ nguyên các điều kiện thực nghiệm còn lại. Bảng 3 tóm tắt các cấu hình được sử dụng.

Bảng 3. Các cấu hình được sử dụng trong nghiên cứu

Cấu hình	Trọng số tin cậy	Tuyến tính	EVT
Mô hình đầy đủ	✓	✓	✓
Không dùng trọng số tin cậy	x	✓	✓
Thay thế bằng ZOH	✓	x	✓
Không EVT (ngưỡng tĩnh)	✓	✓	x

Mô hình đầy đủ: Cấu hình hoàn chỉnh, bao gồm đồng bộ thời gian, nội suy tuyến tính, trọng số độ tin cậy, ngưỡng EVT và cơ chế bỏ phiếu theo

thời gian.

Không dùng trọng số tin cậy: Giữ nguyên dữ liệu đầu vào đã đồng bộ thời gian (nội suy tuyến

tính), nhưng thay trọng số $w_k^{(i)}$ trong hàm mất mát bằng giá trị đồng nhất $w_k^{(i)}=1$, nhằm tách biệt đúng ảnh hưởng của Đóng góp 3.

Thay thế bằng ZOH: Thay nội suy tuyến tính bằng phương pháp giữ bậc không (Zero-Order Hold) [8].

Không EVT (ngưỡng tĩnh): Thay ngưỡng thích nghi EVT bằng ngưỡng tĩnh $\theta=\mu+3\sigma$ [3, 6].

Ngưỡng phát hiện cho mỗi cấu hình được fit riêng biệt trên chính phân phối sai số tái tạo của cấu hình đó, sử dụng ngưỡng sơ cấp $u=Q_{99.5}$ và xác suất báo động giả mục tiêu $p_f=0.02\%$. Điểm vận hành này được lựa chọn vì Recall của mô hình duy trì bão hòa ($\geq 99.8\%$) trên một dải rộng giá trị ngưỡng, cho phép nâng ngưỡng để giảm FAR mà không đánh đổi đáng kể khả năng phát hiện. Để đảm bảo tính đại diện thống kê, mỗi cấu hình được lặp lại với 5 seed ngẫu nhiên độc lập; kết quả được báo cáo dưới dạng $\text{mean} \pm \text{std}$.

Kết quả tại Bảng 4 cho thấy hai điểm chính. Thứ nhất, cơ chế ngưỡng thích nghi EVT vượt rõ ngưỡng tĩnh $\mu+3\sigma$ ở cả ba thiết bị: F1 cải thiện 21.7, 15.8 và 39.1 điểm phần trăm lần lượt tại M1, M2 và M3 (từ 0.775, 0.835, 0.590 lên 0.992, 0.993, 0.981), trong khi FAR giảm từ 0.30%, 0.26%, 0.36% xuống còn 0.00%, 0.01%, 0.01%, xác nhận lợi ích

thực nghiệm mạnh mẽ của Đóng góp 4 [3, 6]. Thứ hai, chênh lệch giữa Full Model và cấu hình w/o Reliability Weight là không đáng kể, với F1-score chênh lệch dưới 0,001 ở cả ba thiết bị và FAR gần như tương đương. Kết quả này cho thấy trong điều kiện truyền thông hiện tại, ảnh hưởng của cơ chế Reliability Weight còn hạn chế. Theo biểu thức $w_k^{(i)} = \exp(-\frac{|t_k - \tilde{t}_j|}{\beta\Delta})$, mức độ tác động của trọng số phụ thuộc vào tỷ lệ giữa khoảng cách nội suy và chu kỳ lấy mẫu Δ . Khi khoảng cách nội suy nhỏ so với Δ , $w_k^{(i)}$ có giá trị gần 1, do đó hàm mất mát có trọng số gần tương đương với hàm mất mát thông thường. Với biên độ jitter được sử dụng trong thí nghiệm hiện tại, nghiên cứu chưa xác định được ngưỡng jitter hoặc tỷ lệ mất gói cụ thể tại đó Reliability Weight bắt đầu cải thiện hiệu năng có ý nghĩa thống kê. Nội dung này được xem là một giới hạn của nghiên cứu hiện tại và cần được khảo sát thêm trong các điều kiện truyền thông có mức jitter và mất mẫu lớn hơn.

Tương tự, cấu hình ZOH Replacement cho kết quả xấp xỉ Full Model ở M1 và M2, trong khi tại M3 (giao thức OPC UA với jitter lớn hơn) PR-AUC giảm nhẹ (0.9843 so với 0.9947), gợi ý rằng nội suy tuyến tính có lợi thế hơn ZOH khi khoảng cách giữa các mẫu gốc lớn.

Bảng 4. Kết quả nghiên cứu triệt tiêu (ngưỡng EVT $u = Q_{99.5}$, $p_f = 0.02\%$, hiệu chỉnh riêng cho từng cấu hình; kết quả trung bình $\pm$ độ lệch chuẩn trên 5 lần chạy).

Thiết bị	Cấu hình	PR-AUC	F1	Độ bao phủ	FAR (%)
M1	Mô hình đầy đủ	1.0000 $\pm$ 0.0000	0.992 $\pm$ 0.001	0.992 $\pm$ 0.000	0.00 $\pm$ 0.00
	Không dùng trọng số tin cậy	1.0000 $\pm$ 0.0000	0.992 $\pm$ 0.001	0.992 $\pm$ 0.000	0.00 $\pm$ 0.00
	Thay thế bằng ZOH	1.0000 $\pm$ 0.0000	0.992 $\pm$ 0.001	0.992 $\pm$ 0.000	0.00 $\pm$ 0.00
	Không EVT (ngưỡng tĩnh)	1.0000 $\pm$ 0.0000	0.775 $\pm$ 0.026	1.000 $\pm$ 0.000	0.30 $\pm$ 0.04
M2	Mô hình đầy đủ	0.9990 $\pm$ 0.0013	0.993 $\pm$ 0.000	0.994 $\pm$ 0.000	0.01 $\pm$ 0.00
	Không dùng trọng số tin cậy	0.9983 $\pm$ 0.0027	0.993 $\pm$ 0.000	0.994 $\pm$ 0.000	0.01 $\pm$ 0.00
	Thay thế bằng ZOH	0.9995 $\pm$ 0.0008	0.993 $\pm$ 0.000	0.994 $\pm$ 0.000	0.01 $\pm$ 0.00

Bảng 4. (tiếp)

Thiết bị	Cấu hình	PR-AUC	F1	Độ bao phủ	FAR (%)
M2	Không EVT (ngưỡng tĩnh)	0.9990 ± 0.0013	0.835 ± 0.015	1.000 ± 0.000	0.26 ± 0.03
	Mô hình đầy đủ	0.9947 ± 0.0075	0.981 ± 0.000	0.985 ± 0.000	0.01 ± 0.00
M3	Không dùng trọng số tin cậy	0.9969 ± 0.0030	0.981 ± 0.000	0.985 ± 0.000	0.01 ± 0.00
	Thay thế bằng ZOH	0.9843 ± 0.0013	0.980 ± 0.002	0.982 ± 0.004	0.01 ± 0.00
	Không EVT (ngưỡng tĩnh)	0.9947 ± 0.0075	0.590 ± 0.011	1.000 ± 0.000	0.36 ± 0.02

5.5. So sánh hiệu năng tổng thể

Trong Bảng 5, tại điểm vận hành đã hiệu chỉnh, Precision của phương pháp đề xuất đạt 97.8-99.1% tùy thiết bị, trong khi Recall duy trì trên 98.5% ở cả ba thiết bị, cho thấy mô hình hầu như không bỏ sót sự cố thực sự dù tỷ lệ bất thường

trong dữ liệu rất thấp (0.26-0.65%) - đặc điểm cố hữu của bài toán phát hiện bất thường mất cân bằng nặng [3, 4]. FAR của phương pháp đề xuất đạt mức xấp xỉ 0.00-0.01% trên cả ba thiết bị, thấp hơn hàng trăm lần so với OCSVM và hàng chục đến hàng trăm lần so với Isolation Forest.

Bảng 5. So sánh hiệu năng giữa phương pháp đề xuất và các phương pháp đối sánh (toàn bộ tập kiểm thử, 518,351 mẫu/thiết bị; kết quả trung bình ± độ lệch chuẩn trên 5 lần chạy).

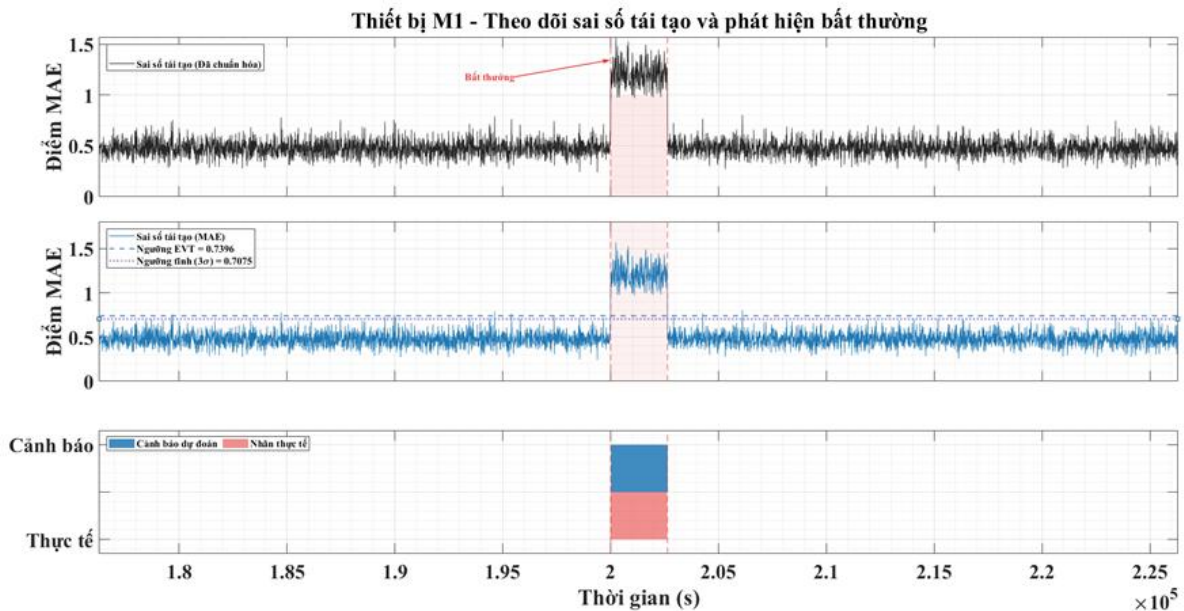
Thiết bị	Phương pháp	Độ chính xác	Độ bao phủ	F1	PR-AUC	FAR (%)
M1	Đề xuất (EVT+ Voting)	0.991 ± 0.002	0.992 ± 0.000	0.992 ± 0.001	1.0000 ± 0.0000	0.00 ± 0.00
	OCSVM	0.168 ± 0.010	1.000 ± 0.000	0.287 ± 0.014	1.0000 ± 0.0000	2.55 ± 0.17
	Isolation Forest	0.336 ± 0.004	1.000 ± 0.000	0.503 ± 0.004	0.9999 ± 0.0001	1.01 ± 0.02
	Đề xuất (EVT+ Voting)	0.991 ± 0.000	0.994 ± 0.000	0.993 ± 0.000	0.9990 ± 0.0013	0.01 ± 0.00
	OCSVM	0.213 ± 0.007	1.000 ± 0.000	0.351 ± 0.010	0.9997 ± 0.0002	2.42 ± 0.11
M2	Isolation Forest	0.399 ± 0.031	1.000 ± 0.000	0.570 ± 0.031	0.9999 ± 0.0001	0.99 ± 0.13
	Đề xuất (EVT+ Voting)	0.978 ± 0.000	0.985 ± 0.000	0.981 ± 0.000	0.9947 ± 0.0075	0.01 ± 0.00
	OCSVM	0.095 ± 0.006	1.000 ± 0.000	0.173 ± 0.009	0.9974 ± 0.0019	2.46 ± 0.16
	Isolation Forest	0.220 ± 0.036	1.000 ± 0.000	0.360 ± 0.049	0.9991 ± 0.0009	0.93 ± 0.20
	Đề xuất (EVT+ Voting)	0.991 ± 0.000	0.994 ± 0.000	0.993 ± 0.000	0.9990 ± 0.0013	0.01 ± 0.00

So với OCSVM và Isolation Forest, phương pháp đề xuất vượt trội rõ rệt về FAR (0.00-0.01% so với 0.93-2.55%) và F1 (0.981-0.993 so với

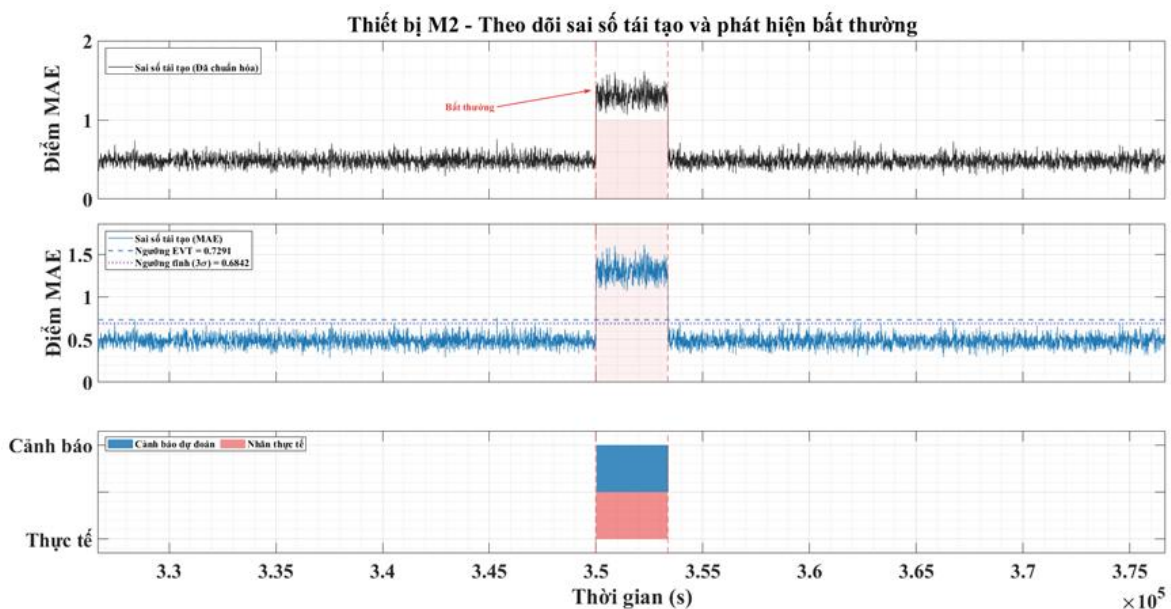
0.173-0.570). Về PR-AUC, cả phương pháp đề xuất và các phương pháp đối sánh đều đạt giá trị rất cao do bộ dữ liệu đánh giá hiện tại được xây

dựng theo phương pháp bán-tổng-hợp, trong đó các sự kiện bất thường được thêm có biên độ tương đối lớn và mức độ phân tách với trạng thái bình thường còn rõ. Vì vậy, PR-AUC trong thí nghiệm hiện tại chưa phản ánh đầy đủ độ khó của

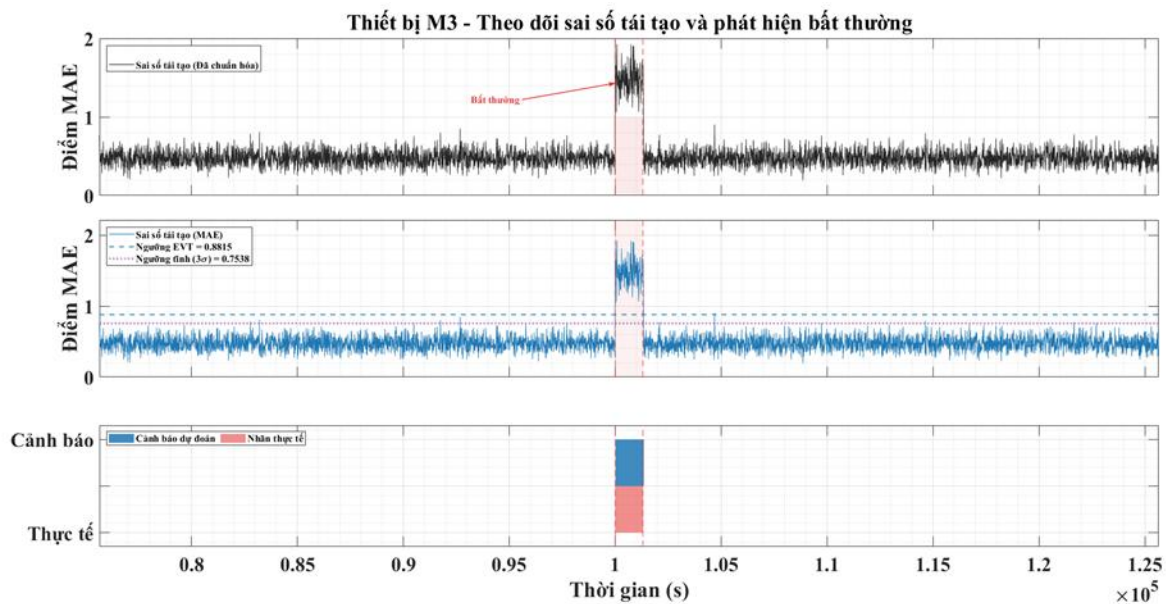
các dạng lỗi tự nhiên có biên độ nhỏ hoặc ranh giới không rõ ràng. Trong phạm vi bộ dữ liệu hiện tại, ưu thế chính của phương pháp đề xuất thể hiện ở khả năng kiểm soát FAR tại điểm vận hành đã chọn.



Hình 4. Theo dõi sai số tái tạo và phát hiện bất thường theo thời gian thực cho Thiết bị M1. Hình minh họa trên một đoạn dữ liệu đại diện ($n=2,000$ mẫu liên tục) trích từ tập kiểm thử đầy đủ (518,351 mẫu), được chọn để bao gồm giai đoạn vận hành bình thường và giai đoạn xảy ra bất thường liên tiếp. Ngưỡng EVT $\theta=0.7396$ cao hơn ngưỡng tính $3\sigma=0.7075$, phản ánh khả năng kiểm soát báo động giả tốt hơn của phương pháp đề xuất



Hình 5. Theo dõi sai số tái tạo và phát hiện bất thường theo thời gian thực cho Thiết bị M2. Hình minh họa trên một đoạn dữ liệu đại diện ($n=2,000$ mẫu liên tục) trích từ tập kiểm thử đầy đủ. Ngưỡng EVT $\theta=0.7291$ nằm cao hơn ngưỡng tính $3\sigma=0.6842$, giúp loại bỏ các cảnh báo giả phát sinh từ dao động bình thường trong tín hiệu rung động đa trục của sàng rung



Hình 6. Theo dõi sai số tái tạo và phát hiện bất thường theo thời gian thực cho Thiết bị M3. Hình minh họa trên một đoạn dữ liệu đại diện ($n=2,000$ mẫu liên tục) trích từ tập kiểm thử đầy đủ. Với dữ liệu OPC UA có jitter lớn hơn ($\Delta_{\max} \approx 75$ ms), ngưỡng EVT $\theta=0.8815$ cao hơn đáng kể so với ngưỡng tĩnh $3\sigma=0.7538$, cho thấy EVT tự động điều chỉnh theo đặc tính phân phối đuôi của từng thiết bị.

5.6. Phân tích định tính chuỗi thời gian giám sát thiết bị thực tế

Nhằm cung cấp cái nhìn chi tiết hơn về cơ chế kích hoạt cảnh báo theo thời gian thực, Hình 4-6 minh họa biểu đồ sai số tái tạo (e_k) kết hợp với đường biên ngưỡng động θ tính toán liên tục dựa trên phân phối EVT đối với từng thiết bị. Quan sát hành vi phân phối sai số trên đoạn dữ liệu minh họa, có thể rút ra nhận xét định tính: trong giai đoạn vận hành ổn định, sai số tái tạo dao động ở mức nền thấp và nằm dưới đường ngưỡng động θ ; tại thời điểm sự cố được thêm, sai số tăng vọt vượt ngưỡng và được xác nhận qua cơ chế Voting trước khi kích hoạt cảnh báo, phù hợp với cơ chế thiết kế ở Mục 4.2 [3, 6].

Đáng chú ý, ngưỡng EVT luôn cao hơn ngưỡng tĩnh $\mu+3\sigma$ ở cả ba thiết bị ($\theta_{\text{EVT}}=0.7396, 0.7291, 0.8815$ so với $\theta_{3\sigma}=0.7075, 0.6842, 0.7538$), xác nhận rằng phương pháp EVT tự động hiệu chỉnh ngưỡng theo đặc tính phân phối đuôi nặng của sai số tái tạo, từ đó giảm đáng kể tỷ lệ báo động giả so với ngưỡng cố định truyền thống [3, 6]. Do nhận bất thường trong nghiên cứu này được tiêm có kiểm soát (Mục 5.1) thay vì xác nhận từ SCADA/nhật ký

bảo trì thực tế, các kết luận về độ chính xác thời điểm phát hiện cần được xác nhận thêm trên dữ liệu lỗi thực tế trong các nghiên cứu tiếp theo.

6. Kết luận và hướng phát triển

6.1. Kết luận

Nghiên cứu đã đề xuất kiến trúc IIoT đồng bộ thời gian tại biên cho giám sát bất thường trong hệ thống công nghiệp đa thiết bị sử dụng LSTM Autoencoder. Kiến trúc tích hợp cơ chế hiệu chỉnh sai lệch thời gian, tiền xử lý tại nút biên, trọng số độ tin cậy, ngưỡng thích nghi EVT và bỏ phiếu theo thời gian. Kết quả thực nghiệm trên bộ dữ liệu bán-tổng-hợp xây dựng từ dữ liệu đo công nghiệp thực tế cho thấy phương pháp đề xuất đạt hiệu quả phát hiện cao và kiểm soát tốt tỷ lệ cảnh báo giả trên ba thiết bị có đặc tính vận hành khác nhau. Các kết quả này cho thấy tính khả thi của kiến trúc đề xuất đối với bài toán giám sát bất thường và bảo trì dự đoán trong môi trường IIoT đa thiết bị.

6.2. Hạn chế và hướng phát triển

Phiên bản thực nghiệm hiện tại còn các hạn chế sau, được để lại như hướng phát triển cho các nghiên cứu tiếp theo:

(i) Một ablation tách biệt hoàn toàn cho bước đồng bộ thời gian, sử dụng trực tiếp nhãn thời gian

chưa hiệu chỉnh clock offset để tạo cửa sổ đầu vào, chưa được triển khai trong nghiên cứu hiện tại. Vì vậy, kết quả hiện tại chưa được xem là bằng chứng định lượng riêng cho mức cải thiện do bước hiệu chỉnh clock offset. Việc thay đổi reference có thể tạo ra bước nhảy thời gian nếu tồn tại clock offset khác nhau giữa các luồng. Cơ chế failover chưa được triển khai và đánh giá thực nghiệm riêng trong nghiên cứu hiện tại. Trong các nghiên cứu tiếp theo, cơ chế này sẽ được triển khai và đánh giá theo hướng ước lượng và bù độ lệch tương đối giữa reference cũ và reference mới trước khi chuyển đổi, đồng thời loại bỏ các cửa sổ dữ liệu không bảo đảm tính liên tục thời gian trong giai đoạn chuyển tiếp.

(ii) Điểm vận hành EVT ($p_u=99.5\%$, $p_f=0.02\%$) được lựa chọn dựa trên khảo sát độ nhạy trên bộ dữ liệu bán-tổng-hợp hiện tại. Do các bất thường được tiêm có mức độ phân tách tương đối rõ so với trạng thái bình thường, các giá trị này chưa thể xem là cố định cho mọi điều kiện vận hành. Khi triển khai trên dữ liệu thực tế, các tham số EVT cần được hiệu chỉnh lại dựa trên phân bố sai số tái tạo của dữ liệu vận hành bình thường. Việc hiệu chỉnh có thể được thực hiện định kỳ bằng cách cập nhật tập dữ liệu nền, ước lượng lại ngưỡng sơ cấp u và các tham số của phân bố GPD, sau đó tính lại ngưỡng phát hiện tương ứng với xác suất báo động giả mục tiêu p_f . Cơ chế tự hiệu chỉnh trực tuyến và phản hồi từ chuyên gia vận hành chưa được triển khai trong nghiên cứu hiện tại và được xem là hướng phát triển tiếp theo.

(iii) Bộ dữ liệu đánh giá hiện tại được xây dựng theo phương pháp bán-tổng-hợp trên nền dữ liệu đo thực tế. Các sự kiện bất thường được tiêm có kiểm soát nhằm bảo đảm nhãn tham chiếu, tuy nhiên một số kịch bản có biên độ tương đối lớn và mức độ phân tách với trạng thái bình thường còn rõ. Do đó, các kết quả AUC/PR-AUC hiện tại chưa đủ để khẳng định hiệu năng của phương pháp đối với các dạng lỗi tự nhiên có biên độ nhỏ, nhiễu nền lớn hoặc ranh giới không rõ ràng. Việc đánh giá trên dữ liệu vận hành dài hạn với nhãn xác nhận

đa nguồn từ SCADA, nhật ký bảo trì và chuyên gia vận hành được xem là hướng phát triển tiếp theo.

Ngoài ra, giới hạn hoạt động của cơ chế Reliability Weight theo mức jitter và tỷ lệ mất mẫu chưa được xác định định lượng trong nghiên cứu hiện tại. Do biên độ jitter trong bộ dữ liệu thực nghiệm còn tương đối nhỏ, chênh lệch hiệu năng giữa Full Model và cấu hình không sử dụng Reliability Weight chưa thể hiện rõ. Trong các nghiên cứu tiếp theo, ảnh hưởng của tỷ lệ jitter so với chu kỳ lấy mẫu và tỷ lệ mất gói sẽ được khảo sát có hệ thống nhằm xác định miền điều kiện mà cơ chế Reliability Weight mang lại cải thiện có ý nghĩa.

Tuyên bố về việc sử dụng AI tạo sinh

Trong quá trình chuẩn bị bài báo này, (các) tác giả đã sử dụng ChatGPT nhằm tạo ra hình 1 trong bài báo dựa trên ý tưởng của các thành viên. Sau khi sử dụng công cụ này, (các) tác giả đã rà soát và chỉnh sửa nội dung khi cần thiết, đồng thời chịu hoàn toàn trách nhiệm về nội dung của bài báo được công bố.

Tài liệu tham khảo

- [1]. D. Yang, A. Mahmood, S.A. Hassan, M. Gidlund. (2022). Guest editorial: Industrial IoT and sensor networks in 5G-and-beyond wireless communication. *IEEE Transactions on Industrial Informatics*, 18(6): 4122–4137. <https://doi.org/10.1109/TII.2022.3142149>
- [2]. W. Zhang, Y. Zhao, J.S. Dhupia, Y. Xu. (2026). A novel approach for remaining useful life prediction with anomaly-driven detection and deep feature transfer. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, 12(4): 041106. <https://doi.org/10.1115/1.4070870>
- [3]. Q. Li, K. Li, X. Gao, J. Fu, L. Zhang. (2024). Anomaly detection based on temporal attention network with adaptive threshold adjustment for electrical submersible pump. *IEEE Transactions on Instrumentation and Measurement*, 73: 1–14.

- <https://doi.org/10.1109/TIM.2024.3436113>
- [4]. S. Benahmad, H. Mlayeh, A. Khedher. (2025). LSTM autoencoder-based approach for anomaly detection of industrial process in presence of incomplete data. *Proceedings of the International Conference on Control and Computer Engineering (STA)*: 1–6. <https://doi.org/10.1109/STA66620.2025.11364820>
- [5]. L. Vu, V.L. Cao, Q.U. Nguyen, D.N. Nguyen, D.T. Hoang, E. Dutkiewicz. (2022). Learning latent representation for IoT anomaly detection. *IEEE Transactions on Cybernetics*, 52(5): 3769–3782. <https://doi.org/10.1109/TCYB.2020.3013416>
- [6]. K. Li, S. Li, Q. Li, Z. Jiao, J. Fu, X. Gao, L. Zhang. (2025). Dual spatio-temporal contrastive learning network with adaptive threshold generation for anomaly detection of electric submersible pump. *IEEE Transactions on Instrumentation and Measurement*, 74: 1–14. <https://doi.org/10.1109/TIM.2025.3587359>
- [7]. J. Matias, R.J. de Araújo Macêdo. (2026). Resilient IIoT communication via edge-based store-and-forward and SLO monitoring. *Proceedings of the XXVII Workshop de Testes e Tolerância a Falhas (WTF)*. <https://doi.org/10.5753/wtf.2026.24074>
- [8]. P. Jia, X. Wang, X. Shen. (2023). Accurate and efficient digital twin construction using concurrent end-to-end synchronization and multi-attribute data resampling. *IEEE Internet of Things Journal*, 10(6), 4857–4870. <https://doi.org/10.1109/JIOT.2022.3221012>
- [9]. F. Luckinger, T. Sauter. (2022). Software-based AUTOSAR-compliant precision clock synchronization over CAN. *IEEE Transactions on Industrial Informatics*, 18(10), 7341 – 7350. <https://doi.org/10.1109/TII.2022.3149923>
- [10]. K. Ye, J. Jiao, F. Chen, Q. Xu, G. Liu, & J. Wang. (2025). The attention enhanced autoencoder for balise anomaly detection in potting process. *IEEE Access*, 13, 91860–91874. <https://doi.org/10.1109/ACCESS.2025.3569756>
- [11]. C.T. Karagul, M.B. Akgun, H.U. Yildiz, B. Tavli. (2025). Mitigating energy cost of connection reliability in UWSNs through nonuniform k-connectivity. *IEEE Internet of Things Journal*, 12(22), 47817 - 47826. <https://doi.org/10.1109/JIOT.2025.3603829>
- [12]. V.Q. Nguyen, V.H. Nguyen, V.L. Cao, N.-A. Le Khac, & N. Shone. (2021). A robust PCA feature selection to assist deep clustering autoencoder-based network anomaly detection. *2021 8th NAFOSTED Conference on Information and Computer Science (NICS)*, 335–341. <https://doi.org/10.1109/NICS54270.2021.9701456>
- [13]. V.Q. Nguyen, V.H. Nguyen, N.-A. Le Khac, V.L. Cao. (2021). Automatically estimate clusters in autoencoder-based clustering model for anomaly detection. *Proceedings of the International Conference on Knowledge Systems Engineering (KSE)*, 1–6. <https://doi.org/10.1109/RIVF51545.2021.9642120>
- [14]. C.T. Bui, L.C. Van, M. Hoang, & Q.U. Nguyen. (2020). A double-shrink autoencoder for network anomaly detection. *Journal of Computer Science and Cybernetics*, 36(2): 159–172. <https://doi.org/10.15625/1813-9663/36/2/14578>
- [15]. H.N. Nguyen, N.N. Tran, T.H. Hoang, & V.L. Cao. (2022). Denoising latent representation with SOMs for unsupervised IoT malware detection. *SN Computer Science*, 3, 474. <https://link.springer.com/article/10.1007/s42979-022-01344-1>