

Thuật toán suy luận mờ và ứng dụng của nó trong chẩn đoán động cơ

(*) TS. LÊ HOÀI ĐỨC

Tóm tắt: Lý thuyết logic mờ được ứng dụng rất phổ biến trong chẩn đoán kỹ thuật nói chung. Nội dung bài báo trình bày thuật toán nhằm tăng tính hiệu quả của việc chẩn đoán bằng logic mờ. Trên cơ sở thuật toán suy luận mờ, đã tiến hành áp dụng cho một thông số cụ thể là độ rò rỉ chân không trên động cơ xăng.

1. Đặt vấn đề

Hệ thống điều khiển điện tử trên ô tô ngày một phát triển trong những năm gần đây, các hư hỏng và biến xấu trạng thái kỹ thuật cũng tăng nhanh và ngày càng phức tạp.

Mô hình mờ được ứng dụng trong chẩn đoán bằng thuật toán Fuzzy logic. Một bộ luật mờ hoàn chỉnh bao gồm n^m luật, với n là số biến vào và m là số trạng thái mờ. Điều này làm cho bài toán chẩn đoán trở nên phức tạp. Qua thuật toán huấn luyện sẽ giúp nâng cao hiệu quả công tác chẩn đoán động cơ.

2. Nội dung

2.1. Xây dựng mô hình thuật toán

Ban đầu bộ dữ liệu mờ được gán giá trị ở các điểm tùy ý và được tối ưu hoá khi thiết lập bộ luật mờ. Các luật mờ được thiết lập trên cơ sở của "luật ưu tiên trong bộ dữ liệu mẫu". Ví dụ chúng ta có luật r "if x_1 is L and x_2 is M , then y is H ", với x_1, x_2 là các biến vào, y là biến ra và L, M, H là các giá trị mờ. Với bộ dữ liệu mẫu s , nếu độ tin cậy cho biến x_1 trong giá trị mờ L là lớn hơn độ tin cậy ở các giá trị mờ khác và độ tin cậy cho biến x_2 trong giá trị mờ M là lớn hơn độ tin cậy ở các giá trị mờ khác, khi đó luật r được gọi là luật ưu tiên của dữ liệu mẫu s và mẫu s được gọi là mẫu ưu tiên của luật r . Mặc dù mẫu của biến vào s có thể không đúng ở các luật khác theo lý thuyết cơ bản, độ tin cậy của nó cho các luật mờ khác phải thấp hơn so với độ tin cậy cho luật r . Khi đó luật mờ của s được cho là ưu tiên với luật r .

Sử dụng phương pháp này, thuật toán thiết lập các luật ưu tiên cho mỗi bộ dữ liệu trong quá trình đào tạo. Nếu bộ luật cơ bản chưa có, nó sẽ được bổ sung vào. Nếu luật này rõ ràng đã tồn tại, thuật toán sẽ tăng mức độ ưu tiên của luật bằng cách thêm giá trị ưu tiên vào luật đã có. Nếu luật mới xung đột với các luật trước, thuật toán sẽ giảm mức độ ưu tiên của luật đã có. Nếu quyền ưu tiên của các luật đã có bị phủ định, chúng sẽ được thay thế bằng luật mới. Cần chú ý rằng, mức độ ưu tiên của luật mờ được bình thường hoá bằng cách sử dụng số lượng mẫu trong bộ dữ liệu ban đầu. Khi đó mức độ ưu tiên của luật thường nhỏ hơn 1.

Chúng ta có thể minh họa thuật toán này bằng ví dụ sau. Bảng 1 chỉ ra tất cả các luật bởi 1000 mẫu. Chúng ta có hai biến vào là x_1 và x_2 và biến ra là y . Các dữ liệu mẫu được chia thành các nhóm có độ tin cậy cao nhất trong các giá trị mờ. Ví dụ nhóm x_1 là *LOW* và x_2 là *LOW* có 200 mẫu, chúng được gọi là mẫu ưu tiên của luật sẽ được thiết lập và luật đó sẽ gọi là luật ưu tiên. Kết quả của luật ưu tiên này xảy ra ba khả năng: y is *Low* có 60 mẫu, y is *Medium* có 100 mẫu và y is *High* có 40 mẫu. Chú ý rằng mỗi nhóm có một đặc trưng là độ tin cậy trung bình. Ví dụ, độ tin cậy trung bình cho các nhóm mà chúng ta đang xét là 0.72, 0.80, 0.65. Sử dụng thuật toán ở trên chúng ta có thể thiết lập luật mờ x_1 là *LOW* và x_2 là *LOW* and y is *Medium*.

(*) Trường Đại học Giao thông Vận tải

Bảng 1: Dữ liệu mẫu

Nhóm	Biến vào		Biến ra					
	x ₁	x ₂	y is Low		y is Medium		y is High	
			Số mẫu	Độ tin cậy TB	Số mẫu	Độ tin cậy TB	Số mẫu	Độ tin cậy TB
0	Low	Low	60	0.72	100	0.80	40	0.65
1	Medium	Low	20	0.62	150	0.77	30	0.67
2	High	Low	30	0.75	140	0.81	30	0.81
3	Low	Medium	120	0.74	20	0.72	20	0.63
4	Medium	Medium	100	0.79	20	0.69	10	0.77
5	High	Medium	0	-	0	-	0	-
6	Low	High	0	-	0	-	0	-
7	Medium	High	0	-	0	-	0	-
8	High	High	10	0.65	10	0.73	90	0.71

Từ các mẫu trong bảng 1, theo luật ưu tiên chúng ta có bộ luật và giá trị ưu tiên như bảng 2. Giá trị ưu tiên thể hiện độ chính xác của bộ dữ liệu và được xác định bằng công thức:

$$\frac{n.A - \frac{1}{k} \sum_{j=1}^k n_j . B_j}{n}$$

với A là độ tin cậy trung bình của mẫu được sử dụng và B_j là độ tin cậy trung bình của mẫu không được sử dụng và k là số luật không sử dụng và n là tổng số mẫu của bộ dữ liệu. Theo ví dụ trên ta có: $(100*0.8 - (60*0.72 + 40*0.65)/2)/100 = 0.0454$

Sau khi luật mờ được thiết lập, độ tin cậy của bộ dữ liệu mẫu bị phá bỏ và luật mờ sẽ được sử dụng trong việc tái thiết lập các nhóm cùng trọng số (hay khả năng) của nhóm.

Bảng 2: Các luật mờ thu gọn từ bảng 1

Nhóm	Biến vào		Biến ra	
	x ₁	x ₂	y	Giá trị ưu tiên
0	Low	Low	Medium	0.04540
1	Medium	Low	Medium	0.09925
2	High	Low	Medium	0.09000
3	Low	Medium	Low	0.07530
4	Medium	Medium	Low	0.06825
5	High	Medium	-	-
6	Low	High	-	-
7	Medium	High	-	-
8	High	High	High	0.05700

Đối với mỗi luật bị phá bỏ, thuật toán sẽ hoà chúng vào các luật đang tồn tại. Trong bảng 2 các luật 5,6,7 là không có giá trị bởi vì không mẫu nào thuộc về các nhóm này. Các luật không có giá trị rất có ích trong quá trình giảm số luật. Ví dụ: trong bảng 2 không có các luật từ 5 đến 7, chỉ có các luật từ 0 đến 2 được hoà lẫn. Luật (0~2) \Rightarrow "if x_2 is Low, then y is Medium".

Số mẫu hỗ trợ của luật mới là $100+150+140=390$. Cùng các luật không giá trị (5~7), luật (3~8) trong bảng 2 có thể hoà lẫn:

Luật (3~5) \Rightarrow "if x_2 is Medium, then y is Low"

Luật (6~8) \Rightarrow "if x_2 is High, then y is High"

Số mẫu hỗ trợ cho hai luật mới này sẽ là: $120+100+0=220$ và $0+0+90=90$. Bộ luật mờ sau khi thiết lập được liệt kê trong bảng 3, với chú ý là các luật không giá trị bị loại bỏ.

Bảng 3: Các luật mờ thu gọn cuối cùng trong suy luận ưu tiên

Luật	Biến vào		Biến ra	
	x_1	x_2	y	Giá trị ưu tiên
0	Không quan tâm	Low	Medium	0.23465
1	Not High	Medium	Low	0.14355
2	High	High	High	0.05700

Bước tiếp theo của thuật toán là tái thiết lập các nhóm. Việc tái thiết lập để giải quyết vấn đề các mẫu thuộc các nhóm khác nhau mà có thể trong cùng một nhóm ở bước tạo nhóm trước. Như vậy, các luật bị phá hủy từ các nhóm đó sẽ không bị ảnh hưởng.

Thủ tục tái thiết lập các nhóm sử dụng việc đánh giá sai lệch giữa mẫu và nhóm. Khi đó các luật sẽ theo mẫu: *If (x_1 is c_1 and x_2 is c_2 and ... and x_n is c_n) then (y is c);*

Đối với lớp c_i của biến điều khiển x_p tồn tại chuỗi (x_{icL}, x_{icH}) trong cùng một thành phần của lớp c_i có giá trị cao nhất. Do vậy, lớp c của biến ra sẽ là một đồ thị đa chiều (lập thể) tạo bởi các luật này. Chúng được đặc trưng bởi bộ dữ liệu (x_{icL}, x_{icH}) với $i = 1, 2, \dots, n$. Tất cả các mẫu như vậy hình thành các nhóm. Bình phương sai lệch giữa mẫu thứ s và nhóm được xác định như sau:

$$SD = \sum_{i=1}^n \varepsilon_{xi}(s)$$

$$\text{mà } \varepsilon_{xi}(s) = \begin{cases} (x_{icL} - x_i^{(s)})^2 & \text{if } (x_i^{(s)} < x_{icL}) \\ 0 & \text{if } (x_{icL} < x_i^{(s)} < x_{icH}) \\ (x_i^{(s)} - x_{icH})^2 & \text{if } (x_{icH} < x_i^{(s)}) \end{cases}$$

Sau khi tất cả các mẫu đã được tái thiết lập nhóm, tâm điểm của mỗi nhóm sẽ được tính toán. Nếu vị trí của các nhóm này gần với thông số tới hạn, luật mờ được thiết lập và thuật toán kết thúc. Trong trường hợp ngược lại, các thông số tới hạn của thành phần mờ được cập nhật và đặt tại tâm điểm các nhóm, ưu tiên của luật mờ bị phá vỡ, tái lập nhóm và quá trình cập nhật được lặp lại.

Việc giảm bớt các luật và triệt tiêu các luật không tin cậy trong quá trình hoà lẫn bằng cách kết hợp các luật cho kết quả hệ tri thức mờ mới làm tăng hiệu quả của việc điều tra lỗi. Tuy nhiên, trong quá trình suy luận mờ có thể mẫu dữ liệu vào bị phá vỡ không tuân theo bất cứ luật nào trong cơ sở tri thức. Để giải quyết vấn đề này, chúng ta phát triển sơ đồ suy luận bằng cách phá vỡ luật gần nhất đối với mẫu vào.

Ví dụ luật mờ “if x_1 is Low and x_2 is High, then y is Medium” cho vị trí trung tâm của nhóm mờ ở tâm điểm các thành phần mờ “if x_1 is Low” và “ x_2 is High”. Trên cơ sở của phương pháp này, chúng ta xác định sai lệch giữa dữ liệu mẫu và tập mờ. Cho bộ dữ liệu mẫu vào là $I = \{a_1, \dots, a_n\}$, với a_i là giá trị tức thời của biến mờ x_i , $i = 1, \dots, n$, $\sum - \{\alpha_1, \alpha_2, \dots, \alpha_p, \Phi\}$ là tập mờ của biến điều khiển trong không gian X , ϕ là biểu thị cho mục “không quan tâm”. Với việc đưa ra ϕ , luật mờ có thể thường xuyên phải viết lại theo mẫu sau:

If x_1 is a_{1k} and x_2 is a_{2k} andand x_n is a_{nk} then z is β

Với hệ thống có ba biến điều khiển $\{x_1, x_2, x_3\}$, và các thành phần mờ là {Low, Medium, High}, nếu chúng ta có luật mờ: “if x_1 is Low and x_3 is High, then y is Medium” có thể viết lại thành: “if x_1 is Low and x_2 is ϕ and x_3 is High, then y is Medium”

Khoảng cách giữa dữ liệu mẫu I và luật mờ k được xác định theo biểu thức: $d = \sqrt{\sum_{i=1}^n (a_i - c_{ik})^2}$

với c_{ik} là tâm điểm của thành phần mờ của biến x_i cho các tập mờ có $\alpha_{ik} \neq \phi$. Luật mờ mà có khoảng cách ngắn nhất tới I sẽ bị phá vỡ.

Với hệ thống thông minh có nhiều nguồn tri thức: từ các chuyên gia, từ việc đào tạo bộ dữ liệu, từ các lý thuyết khoa học... Bộ luật mờ trên cơ sở tính ưu tiên cũng là phương pháp thuận lợi để kết hợp nhiều nguồn tri thức ở giai đoạn suy luận mờ. Suy luận mờ có thể phân chia tính ưu tiên cho các nguồn tri thức này và tổng của chúng thường là 1.

Chúng ta có thể bắt đầu cùng bộ luật được thiết lập trên cơ sở tri thức của chuyên gia hoặc bộ dữ liệu chưa hoàn chỉnh, sau đó đưa ra tri thức mới trong quá trình suy luận mờ. Nếu quá trình suy luận xác định rằng một số lượng lớn dữ liệu không phù hợp với các luật đang tồn tại, nó có thể thiết lập luật mới từ các luật ban đầu. Thuật toán giả thiết rằng các luật mờ ban đầu có giá trị ưu tiên. Sau đó sẽ tích lũy giá trị ưu tiên mới cho mỗi luật theo biểu thức:

$p * \frac{N}{N + M}$ với p là giá trị ưu tiên cũ của luật, M là số mẫu trong bộ dữ liệu đào tạo hiện thời. Tiếp theo là thiết

lập luật mới từ bộ dữ liệu đào tạo theo thuật toán đã nêu ở trên. Giá trị ưu tiên của luật mới sẽ được tính theo biểu

thức: $\frac{1}{N + M} \cdot \prod_{s=1}^M b_s$, với b_s là độ tin cậy của mẫu thứ s trong quá trình đào tạo mới khi phá bỏ luật này. Nếu luật

mới chưa có, nó sẽ thêm vào trong hệ tri thức. Nếu luật mới đã có, ta thêm giá trị ưu tiên của luật mới vào luật đã có. Nếu luật mới đối nghịch với luật đã có, ta so sánh độ ưu tiên của hai luật. Nếu độ ưu tiên của luật mới nhỏ hơn ta giảm độ ưu tiên của luật cũ bằng cách trừ đi độ ưu tiên của luật cũ. Ngược lại, luật mới sẽ thay thế luật cũ với độ ưu tiên chính là sự sai khác giữa hai độ ưu tiên cũ và mới.

2.2. Hệ thống chẩn đoán mờ động cơ

Chúng ta chọn việc chẩn đoán độ rò rỉ chân không trong động cơ để minh họa cho thuật toán trên. Việc điều tra rò rỉ chân không là một vấn đề phức tạp chịu ảnh hưởng bởi nhiều nhân tố không chắc chắn. Do vậy hệ thống mờ là một giải pháp tốt cho việc chẩn đoán rò rỉ. Để thiết lập mô hình chẩn đoán, đầu tiên chúng ta phải xác định các biến vào ra. Một thông số điều khiển điện tử chịu ảnh hưởng nhiều nhất bởi độ rò rỉ chân không là tỷ lệ không khí/nhiên liệu lý tưởng A/F “Lambda” (λ). Tỷ lệ A/F được đo bằng 1 hoặc 2 cảm biến ô xy tùy theo từng hệ thống. Hệ thống kép điều khiển nhiên liệu cho từng công xy lanh, Lambda 1 cho công xy lanh số 1 và Lambda 2 cho công xy lanh số 2. Nếu có rò rỉ chân không, không khí sẽ vào buồng cháy nhiều hơn so với giá trị của cảm biến lưu lượng không khí ghi lại, kết quả là cháy với hỗn hợp nghèo. Trong trường hợp này giá trị của hỗn hợp sẽ được làm giàu $\lambda < 1$. Nếu cảm biến lưu lượng khí ghi lại giá trị nhiều hơn thực tế, nó sẽ điều khiển giảm lượng khí cấp vào và hỗn hợp cháy sẽ giàu. Trong trường hợp này, giá trị của Lambda sẽ bị làm nghèo. Ở hệ thống kép, sai lệch Lambda ở một hoặc hai công sẽ gây ra những vấn đề khác nhau.

Bên cạnh λ_1 , λ_2 , ba nhân tố khác cũng cần được xét tới: vị trí bướm ga, tốc độ không tải và lưu lượng khí nạp. Lượng khí đi vào buồng cháy được điều khiển bởi độ mở của cánh gạt. Cảm biến vị trí bướm ga xác định góc mở của cánh gạt và chuyển thành tín hiệu điện áp và chuyển về bộ điều khiển trung tâm. Nếu có rò rỉ chân không, vị trí bướm ga sẽ thấp. Tốc độ không tải liên quan tới việc điều khiển cuộn solenoid. Lưu lượng khí nạp cho ta biết lượng khí đi vào xy lanh. Như vậy, chẩn đoán rò rỉ chân không có 5 biến vào: vị trí bướm ga, λ_1 , λ_2 , tốc độ không tải, lưu lượng khí nạp và một biến ra là độ rò rỉ chân không. Mỗi biến có ba thành phần mờ: Low, Medium và High.

Cơ sở tri thức mờ bao gồm các luật mờ và các thành phần được xác định dựa trên phương pháp đã trình bày ở trên. Các dữ liệu được lấy từ Ford Motor Company [3]. Trong cơ sở dữ liệu bao gồm các động cơ chất lượng tốt và động cơ chất lượng kém. Động cơ mẫu được coi là tốt khi độ rò rỉ chân không là Low và ngược lại là kém khi độ rò rỉ chân không là High. Các chuyên gia của công ty đưa ra 6 luật mờ với ba thành phần cho mỗi biến được thể hiện trên bảng 4 và 5. Ứng dụng tập luật từ cơ sở tri thức của các chuyên gia trong việc chẩn đoán độ rò rỉ chân không trong động cơ cho độ chính xác là 50% với mẫu thử nghiệm [3], [4].

Theo thuật toán đã trình bày chúng ta có tập mờ và các thành phần thể hiện trên bảng 6 và 7. Kiểm tra với các mẫu từ cơ sở dữ liệu cho độ chính xác là trên 95%.

Bảng 4: Tập luật từ các chuyên gia

STT	Biến vào					Biến ra
	Lambda 1	Lambda 2	Tốc độ không tải	Lưu lượng khí nạp	Vị trí bướm ga	Rò rỉ chân không
0	Low	Low	Low	Medium	Low	High
1	Low	Low	Low	Medium	Low	High
2	Low	Medium	Không quan tâm	Không quan tâm	Không quan tâm	Low
3	Low	High	Không quan tâm	Không quan tâm	Không quan tâm	Low
4	Medium	Low	Không quan tâm	Không quan tâm	Không quan tâm	Low
5	High	Low	Không quan tâm	Không quan tâm	Không quan tâm	Low

Bảng 5: Tổng hợp các thành phần mờ từ các chuyên gia

STT	Biến vào					Biến ra
	Lambda 1	Lambda 2	Tốc độ không tải	Lưu lượng khí nạp	Vị trí bướm ga	Rò rỉ chân không
0	0.8	0.8	0.2	0	150	0
1	0.95	0.95	0.34	0.75	200	0.5
2	1.1	1.1	0.44	2	250	1

Bảng 6: Tập luật mờ qua đào tạo

STT	Biến vào					Biến ra
	Lamda 1	Lamda 2	Tốc độ không tải	Lưu lượng khí nạp	Vị trí bướm ga	Rò rỉ chân không
0	High	High	Medium	High	Medium	Low
1	High	High	Medium	High	High	Low
2	High	High	High	High	Medium	Low
3	Medium	Medium	High	High	High	High
4	Low	Low	Low	High	High	High
5	Low	Low	Low	Medium	High	High
6	Low	Low	Low	Medium	Medium	High
7	Low	Low	High	Medium	High	High
8	High	High	High	High	High	Low
9	Medium	High	High	High	High	High
10	Medium	Medium	Medium	High	High	Low
11	High	High	High	Medium	High	Low
12	Medium	Medium	Medium	Medium	High	High
13	Low	Low	Medium	Medium	High	High
14	Low	Medium	Low	Medium	High	High
15	Medium	Medium	High	Low	High	High
16	High	Low	High	High	Medium	High
17	Medium	Medium	Medium	High	Medium	High
18	Không quan tâm	Low	High	High	Low	High

Bảng 7: Tổng hợp các thành phần mờ qua đào tạo

STT	Biến vào					Biến ra
	Lamda 1	Lamda 2	Tốc độ không tải	Lưu lượng khí nạp	Vị trí bướm ga	Rò rỉ chân không
0	0.748385	0.737143	0.213	0.009	156	0
1	0.828	0.808375	0.337143	0.7645	191.783	0.5
2	0.995903	0.97529	0.3938	0.89245	201.828	1

3. Kết luận:

Bài báo trình bày thuật toán suy luận mờ dựa trên cơ sở luật ưu tiên. Mô hình tính cho phép nâng cao hiệu quả của công tác chẩn đoán. Việc suy luận mờ cho phép gia tăng cơ sở tri thức, tích lũy và phá bỏ các luật trong cơ sở dữ liệu ban đầu. Thuật toán chẩn đoán này được áp dụng cho việc chẩn đoán độ rò rỉ chân không trong động cơ đã chứng tỏ hiệu quả của mô hình. ■

Tài liệu tham khảo:

1. Lê Hoài Đức, Ứng dụng lý thuyết tập mờ trong chẩn đoán kỹ thuật động cơ diesel, Tạp chí Giao thông vận tải, 2008.
2. Johan de Kleer, Focusing on probable diagnoses, AAAI, 1991.
3. L.S.Tedesco, Service Bay Diagnostic System. Ford Motor Co, Society of Automotive Engineers, 1986.
4. H.Genther, A. König, M.Glesner, Rule weight generation for fuzzy classification system base on fuzzy cluster methods, in Proc. 3rd IEEE Int. Conf. Fuzzy Syst., Orlando, USA, 1994.
5. Yi Lu, Tie Qi Chen, A Fuzzy Diagnostic Model and Its Application in Automotive Engineering Diagnosis, Applied Intelligence, 1998.