

ỨNG DỤNG GIẢI THUẬT DI TRUYỀN TỐI ƯU HÓA QUY LUẬT HỢP NHẤT TRONG MẠNG CẢM BIẾN PHÂN TÁN

Nguyễn Đức Minh^{*}, Phạm Văn Hùng⁺

^{*} Học viện Công nghệ Bưu chính Viễn thông

⁺ Học viện Kỹ thuật Quân sự

Tóm tắt: Bài báo trình bày phương pháp áp dụng giải thuật di truyền để tìm quy luật hợp nhất tối ưu cho hệ thống mạng cảm biến phân tán khi các phép quan sát của các cảm biến trong mạng có tương quan theo phân bố Student-t. Mô hình mạng cảm biến phân tán có cấu trúc song song được khảo sát cho thấy các kết quả tương đồng với những nghiên cứu trước đó đã khẳng định sự phù hợp của phương pháp này. Nội dung bài báo cũng đề cập đến việc khảo sát và đánh giá ảnh hưởng của hệ số tương quan giữa các cảm biến trong mạng tới tốc độ hội tụ của giải thuật GA trong một số trường hợp với các giá trị số cụ thể. Kết quả mô phỏng cho thấy xác suất lỗi P_e của toàn hệ thống bị ảnh hưởng mạnh bởi mối tương quan giữa các cảm biến. Mối tương quan giữa các cảm biến càng lớn thì xác suất lỗi P_e càng lớn và ngược lại. Đặc biệt khi các cảm biến có mối tương quan âm thì chất lượng hệ thống được cải thiện rõ rệt. Khi số cảm biến trong mạng tăng lên thì tốc độ hội tụ của thuật toán GA càng chậm đạt tới lời giải tối ưu. Số bậc tự do của phân bố Student-t không gây ảnh hưởng nhiều tới chất lượng hệ thống nhưng tốn nhiều thời gian để giải thuật đạt tới lời giải tối ưu.

Từ khóa: Giải thuật di truyền, quy luật hợp nhất, mạng cảm biến phân tán, tương quan, phân bố Student-t.

I. GIỚI THIỆU CHUNG

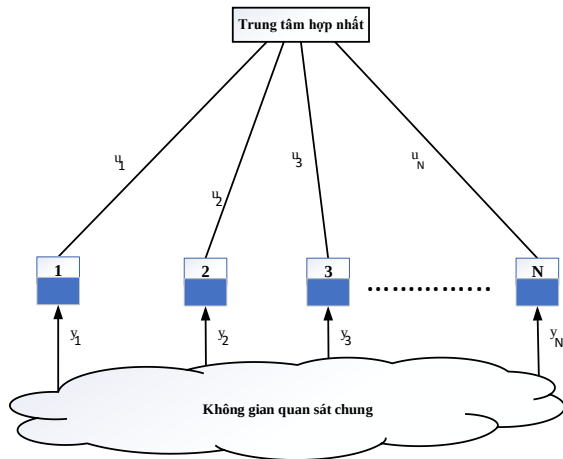
Mạng cảm biến có cấu tạo bao gồm nhiều cảm biến cùng giám sát một không gian chung, thu thập các thông tin và gửi các thông tin này về một trung tâm hợp nhất (fusion center). Tại đây, trung tâm hợp nhất sẽ sử dụng các thông tin này để phân loại hoặc ước lượng, đánh giá các thông tin thu được. Tùy thuộc vào chức năng của trung tâm hợp nhất mà bài toán hợp nhất thông tin có thể là bài toán phát hiện, ví dụ như khi phải đưa ra quyết định lựa chọn giữa các trạng thái, hoặc bài toán ước lượng chẳng hạn như khi trung tâm hợp nhất phải ước lượng một số đại lượng liên quan tới các phép quan sát của các cảm biến [12]. Do giá của các cảm biến ngày càng rẻ hơn và những ưu điểm của hệ thống mạng cảm biến nên ngày càng có nhiều các nghiên cứu về mạng cảm biến được công bố trong những năm gần đây. Ban đầu mạng cảm biến chủ yếu được dùng trong lĩnh vực

quân sự với các ứng dụng giám sát và điều khiển, nhưng ngày nay các ứng dụng của nó đã được phổ biến rộng rãi trong các lĩnh vực khoa học kỹ thuật, công nghiệp, chăm sóc sức khỏe, nông nghiệp và các ứng dụng dân sự. Ban đầu các cảm biến trong mạng không thực hiện bất kỳ một công việc xử lý dữ liệu sơ bộ nào, tất cả các dữ liệu thu thập được đều gửi về trung tâm hợp nhất để thực hiện việc phát hiện. Một mạng như vậy được gọi là mạng cảm biến xử lý tập trung (centralized multi-sensor networks). Trở ngại lớn mà chúng ta gặp phải khi thiết kế các mạng cảm biến xử lý tập trung như vậy là hạn chế về băng thông liên lạc khi truyền thông tin từ các cảm biến cục bộ tới trung tâm hợp nhất. Có một cách để vượt qua rào cản này là thực hiện một số xử lý sơ bộ đối với dữ liệu tại mỗi cảm biến cục bộ và sau đó gửi các thông tin đã được cô đọng tới trung tâm hợp nhất [27]. Những mạng cảm biến như vậy có những nút mạng thông minh [9] và được gọi là mạng cảm biến phân tán (distributed or decentralized sensor networks). Một mạng như vậy được mô tả trong Hình 1. Ở Hình 1 các hộp vuông biểu diễn các cảm biến, phần bôi đen là bộ phận cảm nhận của các cảm biến còn phần trắng là thành phần xử lý dữ liệu. Các phép quan sát được biểu diễn bởi y_i , $i=1, \dots, N$ và các thông tin đã được xử lý sơ bộ được gửi tới trung tâm hợp nhất được ký hiệu bởi u_i , $i=1, \dots, N$. Việc xử lý sơ bộ được thực hiện tại các cảm biến thường là các thao tác nén hoặc lấy mẫu dữ liệu. Do vậy, trung tâm hợp nhất chỉ thu được một phần dữ liệu của các cảm biến thu được từ môi trường quan sát, kết quả là sẽ có những sự mất mát thông tin khi so sánh với trường hợp mạng cảm biến xử lý tập trung [28]. Mặc dầu vậy, sự mất mát hiệu năng này có thể được giảm thiểu bằng việc xử lý tối ưu các thông tin thu được từ môi trường tại các cảm biến [25]. Mạng cảm biến xử lý phân tán có ưu điểm là giảm thiểu các yêu cầu về băng thông, gia tăng độ tin cậy, điều này khiến cho nó trở nên rất hấp dẫn đối với những nhà thiết kế hệ thống cũng như các nhà nghiên cứu [25], [10], [3], [5], [14], [22], [4], [8], [16]. Mặc dầu các mô hình mạng cảm biến phân tán thường có cấu trúc bao gồm các cảm biến và một trung tâm hợp nhất, nhưng cũng có những mạng hoạt động mà không sử dụng trung tâm hợp nhất. Một trong những ví dụ về những mạng như vậy có thể kể đến mạng cảm biến phân tán với cấu hình nối tiếp [27], hay mạng sử dụng chiến lược “đàm phán” để lựa chọn quyết định cuối cùng như trong [24]. Tuy nhiên, bài báo này chỉ đề cập tới mạng cảm biến phân tán có cấu hình song song với một trung tâm hợp nhất.

Tác giả liên hệ: Nguyễn Đức Minh,

Email: minhnd@ptit.edu.vn

Đến tòa soạn: 10/2021, chỉnh sửa: 11/2021, chấp nhận đăng: 12/2021.



Hình 1. Mạng cảm biến phân tán cấu hình song song với một trung tâm hợp nhất

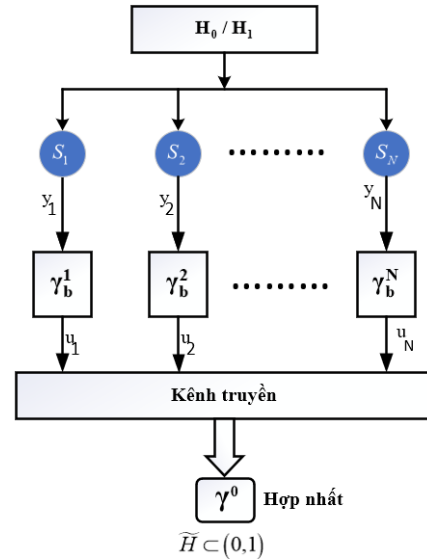
Các giả thuyết này thường được biết tới như là các giả thuyết nhị phân. Mục đích chính của việc thiết kế mạng cảm biến phân tán là tìm ra quy luật quyết định tối ưu tại từng cảm biến và quy luật hợp nhất tối ưu tại trung tâm hợp nhất. Nội dung bài báo này sẽ đi sâu vào khía cạnh tối ưu hóa việc hợp nhất dữ liệu. Bài toán có thể được chia ra làm hai trường hợp: trường hợp thứ nhất khi phép quan sát tại các cảm biến cục bộ là độc lập có điều kiện, trường hợp thứ hai là khi phép quan sát tại các cảm biến tương quan. Giả thiết về tính chất độc lập có điều kiện của các phép quan sát của các cảm biến cục bộ làm cho bài toán trở nên đơn giản hơn. Tuy nhiên, giả thiết này hiếm khi phù hợp trên thực tế khi việc đặt các cảm biến gần nhau sẽ khiến cho các phép quan sát của các cảm biến trở nên tương quan với nhau. Bài toán trong mạng cảm biến phân tán với các phép quan sát của các cảm biến có tương quan đã được nghiên cứu trong một số công bố [21], [18], [15], [6], [1], [19], [17]. Hầu hết các công bố đều cố gắng giải bài toán phát hiện với giả thiết mô hình tương quan Gauss, các cảm biến đồng nhất hoặc không đồng nhất có các quyết định nhị phân với hai mức 0 và 1 được truyền về trung tâm hợp nhất trên các kênh đa truy cập MAC (Multiple Access Chanel) không có nhiễu. Phương pháp giải thường là sử dụng các công cụ toán học như tính trực tiếp các tích phân xác suất phức tạp [15], sử dụng công thức Bahadur-Lazarsfeld [19] để biểu diễn các hàm mật độ xác suất phân bố nhiều chiều thông qua các hệ số tương quan, hay cách dùng lý thuyết Copula [23] để xây dựng các mô hình tương quan có phân bố thống kê tính toán được dựa vào các phân bố biên hoặc dùng các giải thuật tiến hóa với trường hợp tương quan phân bố Gauss [11], [7], [2]...vv. Các kết quả thu được còn rời rạc và chưa nhiều hoặc gặp phải số lượng các phép tính rất lớn. Nhằm giảm thiểu gánh nặng tính toán và tìm một cách tiếp cận khác cho vấn đề này bài báo trình bày phương pháp áp dụng giải thuật di truyền nhằm tối ưu hóa quy luật hợp nhất trong mạng cảm biến phân tán trong trường hợp nhiều tương quan, tuy nhiên nhiều được đề cập đến ở đây là nhiều tương quan có phân phối Student-t.

Sở dĩ đề cập đến phân phối Student-t trong bài báo này là do vai trò quan trọng của nó khi số lượng phép thử thống kê không đủ lớn, hay nói cụ thể hơn là số cảm biến trong mạng không lớn. Khi đó phân phối Gauss tỏ ra không

còn phù hợp nữa. Phương pháp này có thể được mở rộng cho việc tối ưu hóa tổng thể toàn hệ thống với các mô hình tương quan có phân bố thống kê khác nhau.

II. MÔ HÌNH MẠNG VÀ BÀI TOÁN TỐI ƯU

Xét mô hình mạng cảm biến phân tán với một trung tâm hợp nhất như được cho ở Hình 2 dưới đây.



Hình 2. Mô hình mạng cảm biến của bài toán tối ưu

Mạng gồm có N cảm biến và một trung tâm hợp nhất. Các cảm biến cục bộ thu thập kết quả từ các phép đo y_n , tạo ra các quyết định cục bộ u_n của mỗi cảm biến, sau đó truyền các quyết định này về trung tâm hợp nhất γ_0 thông qua một kênh truyền MAC không có nhiễu. Trung tâm hợp nhất sẽ tạo ra quyết định tổng thể cuối cùng H về trạng thái của môi trường quan sát H dựa trên các quyết định cục bộ thu được từ các cảm biến. Giả sử rằng trung tâm hợp nhất không thực hiện phép đo nào của riêng nó. Tại trung tâm hợp nhất ta có bài toán phát hiện nhị phân với các giả thuyết H_0 và H_1 tương ứng với các xác suất tiên nghiệm π_0 và π_1 . Phép quan sát y_i tại mỗi một cảm biến chịu ảnh hưởng của nhiễu tương quan có phân bố Student-t được cho bởi công thức:

$$y_i = \begin{cases} s_i + n_i & \text{dưới giả thuyết } H_1 \\ n_i & \text{dưới giả thuyết } H_0 \end{cases} \quad (1)$$

Kết quả lỗi ra của trung tâm hợp nhất là các giá trị nhị phân H_0 và H_1 . Giả sử rằng mỗi cảm biến sẽ cho kết quả là một chuỗi nhị phân. Do đó, mỗi một cảm biến trong mạng sẽ

chia phép quan sát y_n của chúng thành $L = 2^b$ lớp trong đó b là số bit được truyền đi từ mỗi cảm biến. Như vậy, mỗi cảm biến sẽ ánh xạ không gian quan sát của chúng thành một không gian phân lớp với L lớp và trung tâm hợp nhất ánh xạ N quyết định cục bộ thành một trong hai lớp ứng với hai giả thuyết H_0 và H_1 . Mỗi một quyết định cục bộ có thể có của các cảm biến được biểu diễn bởi một vector gồm N số nguyên như sau:

$$\mathbf{u} = (u_1 \ u_2 \ \dots \ u_N), \quad u_n \in \{0, 1, \dots, L-1\} \quad (2)$$

Giả sử rằng $L=2^b$ vậy $\mathbf{u}$ có thể được biểu diễn như là một chuỗi gồm bN bit như sau:

$$\mathbf{u} = (u_1^1, u_1^2, \dots, u_1^b, u_2^1, u_2^2, \dots, u_2^b, \dots, u_N^1, u_N^2, \dots, u_N^b) \quad (3)$$

Trong đó $u_n^j \in \{0,1\}$. Như vậy, không gian của tất cả các quyết định cục bộ có thể có được biểu diễn bởi một số nguyên q gồm bN bit có giá trị chạy từ 0 tới $2^{bN}-1$. Với một tổ hợp cụ thể của các quyết định cục bộ được biểu diễn bởi q , từng giá trị của các quyết định cục bộ $u_n, n=1,2,\dots,N$ có thể được trích ra bằng cách sử dụng hàm ánh xạ ngược $\psi_n(q)$. Hàm này được định nghĩa như sau:

$$\psi_n(q) = \frac{q}{2^{b(N-n)}} \bmod L \quad (4)$$

Trong đó $\bmod$ là phép toán module và thực hiện với các số nguyên. Để biểu diễn quy tắc hợp nhất ta sử dụng kiểu biểu diễn nhị phân để mô tả kết quả đầu ra của quy tắc hợp nhất với mọi sự kết hợp có thể có của các quyết định cục bộ. Nếu mạng bao gồm N cảm biến và mỗi cảm biến phân lớp phép quan sát của chúng thành L lớp thì mỗi một quy luật hợp nhất có thể sử dụng cho L^N các khả năng quyết định cục bộ và được biểu diễn như một chuỗi gồm L^N bit như sau:

$$h = (h_0, h_1, \dots, h_{L^N-2}, h_{L^N-1}) \quad (5)$$

$$h_q \in \{0,1\}, q = 0,1,\dots,L^N - 1$$

Ví dụ, nếu quy tắc hợp nhất được cho bởi chuỗi nhị phân (00...01) thì điều này có nghĩa là trung tâm hợp nhất luôn quyết định H_0 , trừ trường hợp khi $u_1=00$ và $u_2=01$ khi đó trung tâm hợp nhất sẽ quyết định H_1 .

Như đã chỉ ra trong công bố [15], đầu tiên chúng ta xét hàm mật độ nhiều đối xứng để có thể mô tả một cách hoàn chỉnh nó bằng một hệ số tương quan ρ duy nhất. Do vậy, với hệ thống gồm 3 cảm biến ma trận hiệp phương sai của của phân bố đa biến Student-t với trung bình bằng 0 và phương sai bằng 1 có dạng như sau:

$$\lambda = \begin{pmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{pmatrix} \quad (6)$$

Tuy ở đây ta đang sử dụng hệ số tương quan giống nhau ρ (tương quan đều), tuy nhiên với cách tiếp cận bài toán của chúng ta các kết quả cũng đúng cho mọi ma trận hiệp phương sai xác định dương. Việc tối ưu các quyết định cục bộ được giả sử là kiểm định tỷ số phù hợp nhất LRT (Likelihood Ratio Test). Khi việc kiểm định tỷ số phù hợp nhất đã được dùng để tối ưu hóa cho trường hợp phép quan sát của các cảm biến không tương quan [13], [26], thì nó cũng được dùng một cách rộng rãi cho trường hợp tương quan [21], [18], [15], [6], [20], [19]. Các cảm biến cục bộ giả sử được lượng tử hóa thành L mức (do vậy có $L-1$ ngưỡng). Từ đó, các quy luật quyết định cục bộ được cho bởi công thức:

$$u_n = \begin{cases} 0 & \text{nếu } y_n \leq \lambda_{n,1} \\ 1 & \text{nếu } \lambda_{n,1} \leq y_n \leq \lambda_{n,2} \\ \dots & \dots \\ L-1 & \text{nếu } y_n > \lambda_{n,L-1} \end{cases} \quad (7)$$

Trong đó y_n là phép đo cục bộ tại cảm biến thứ n , và quyết định cục bộ tương ứng là u_n và $\lambda_{n,1}, \lambda_{n,2}, \dots, \lambda_{n,L-1}$ là $L-1$ ngưỡng lượng tử của cảm biến thứ n . Để có thể tối ưu hóa quy luật hợp nhất của một mạng cảm biến phân tán như vậy, việc tối ưu phải được thực hiện trên tất cả các quy luật hợp nhất có thể có. Bài toán có thể được xác định cụ thể như sau: *Xác định quy luật hợp nhất tối ưu h của mạng cảm biến phân tán, ở đây việc tối ưu được định nghĩa là cực tiểu hóa xác suất lỗi P_e tại trung tâm hợp nhất.*

Bài toán tối ưu của mạng cảm biến phân tán trên các khía cạnh tối ưu hóa quy luật hợp nhất là một bài toán tối ưu N-P (Neyman - Pearson) hoàn chỉnh khi giả thiết không gian quan sát là rời rạc. Bài toán này có độ phức tạp tính toán tăng theo hàm mũ tỷ lệ với số cảm biến trong mạng và do đó không thể thực hiện trên thực tế khi việc tìm kiếm lời giải tối ưu là quá khó khăn với một không gian tìm kiếm lớn. Các giải thuật tiến hóa, chẳng hạn như giải thuật di truyền Genetic Algorithms – GA là một trong những giải pháp thường được sử dụng trong những trường hợp như vậy. Giải thuật GA đã tỏ ra có hiệu quả trong việc tìm kiếm các lời giải gần đúng cho rất nhiều bài toán N-P hoàn chỉnh. Thuật toán GA sử dụng sự tiến hóa và cơ chế sống sót vì thích nghi để tìm kiếm các lời giải theo hướng phù hợp nhất. Trong phần III sau đây chúng ta sẽ đề cập đến phương pháp để giải quyết bài toán tối ưu này.

III. ÁP DỤNG GIẢI THUẬT GA CHO BÀI TOÁN TỐI ƯU

Thuật toán GA được trình bày ở đây sẽ tối ưu hóa quy luật hợp nhất tại trung tâm hợp nhất trong mạng. Trường hợp khi các cảm biến không tương quan, ban đầu, một quần thể ngẫu nhiên được tạo ra bao gồm một nhóm các nhiễm sắc thể, sự phù hợp của mỗi một nhiễm sắc thể được tính toán bởi xác suất lỗi $P_e(\lambda, h)$, xác suất này được tính toán dựa vào phương trình sau [2]:

$$P_e(\lambda, h) = \sum_{k=0}^1 \pi_k \sum_{q=0}^{L^N-1} \prod_{n=1}^N P_k^n(\psi_n(q), \lambda) \quad (8)$$

$$h_q = k$$

$$P_k^n(m, \lambda) = \int_{\lambda_{n,m}}^{\lambda_{n,m+1}} f_k(y) dy = F_k(\lambda_{n,m+1}) - F_k(\lambda_{n,m}) \quad (9)$$

Khi các cảm biến trong mạng có tương quan, chúng ta không còn có thể nhân một cách đơn giản các hàm mật độ xác suất của các quan sát của các cảm biến như khi chúng độc lập có điều kiện. Do đó, chúng ta cần tính xác suất lỗi trung bình trong trường hợp tương quan. Điều này được thực hiện như sau: xác suất lỗi trung bình tại trung tâm hợp nhất được cho bởi tổng có trọng số của sai lỗi loại I và sai lỗi loại II.

$$P_e(\lambda, h) = \sum_{k=0}^1 \pi_k P_k^0(\bar{k}, \lambda, h) \quad (10)$$

Trong đó π_k là xác suất tiên nghiệm của giả thuyết H_k , $P_k^0(\bar{k}, \lambda, h) = \Pr(u_0 = \bar{k} | H_k)$ là xác suất báo động làm nếu $k=0$ hoặc xác suất bỏ sót nếu $k=1$, và $\bar{k}$ là ký hiệu cho phép toán nhị phân NOT k . Ngoài L^N tổ hợp các quyết

định cục bộ có thể có, tổng hợp số tổ hợp các quyết định cục bộ cho kết quả là quyết định $u_0 = \bar{k}$ tại trung tâm hợp nhất như sau:

$$P_e(\lambda, h) = \sum_{k=0}^1 \pi_k \sum_{\substack{q=0 \\ h_q = \bar{k}}}^{L-1} \Pr(u_1 = \Psi_1(q), \dots, u_N = \Psi_N(q)) \quad (11)$$

Trong đó $\Pr(u_1 = \Psi_1(q), \dots, u_N = \Psi_N(q))$ là xác suất liên kết của cảm biến 1 quyết định $\Psi_1(q)$, cảm biến 2 quyết định $\Psi_2(q)$,... và tương tự. Vì các cảm biến hoạt động như một bộ lượng tử nên xác suất liên kết này có thể được coi là tập hợp của nhiều tích phân theo công thức sau:

$$\Pr(u_1 = \Psi_1(q), \dots, u_N = \Psi_N(q)) = \int_{\lambda_{\Psi_1}}^{\lambda_{\Psi_1+1}} \dots \int_{\lambda_{\Psi_N}}^{\lambda_{\Psi_N+1}} f_k(y_1, \dots, y_N) dy_1 \dots dy_N \quad (12)$$

Trong đó: $f_k(y_1, \dots, y_N)$ là hàm mật độ xác suất kết hợp của phép quan sát $y_1, \dots, y_N$.

Ở đây chúng ta sử dụng một tích phân lặp cho hàm mật độ xác suất liên kết trong việc đánh giá xác suất lỗi trung bình trong trường hợp tương quan.

IV. CÁC KẾT QUẢ MÔ PHỎNG

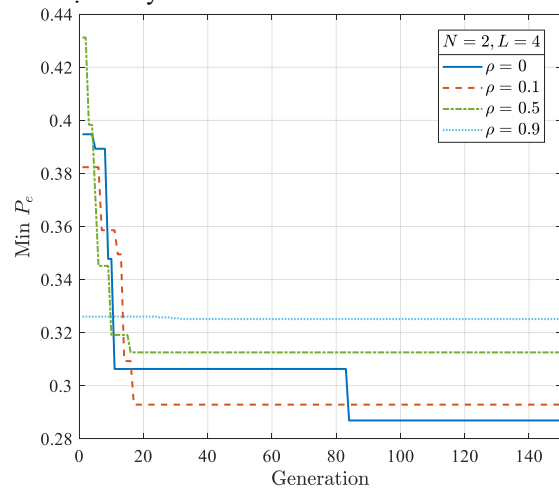
Trong bài toán của mình chúng ta xem xét trường hợp các cảm biến đồng nhất, đồng thời chỉ xét mạng cảm biến với 2 hoặc 3 cảm biến, số cảm biến nhỏ cũng phù hợp cho phân bố Student-t được khảo sát ở đây. Số bậc của tích phân tăng lên cùng với số cảm biến và gia tăng độ phức tạp của phép tính nhưng với cách tiếp cận của chúng ta giải thuật vẫn sử dụng được tốt cho mạng với số lượng cảm biến lớn. Số bit của mỗi cảm biến được giả thiết bằng 2 nên mỗi cảm biến sẽ phân vùng quan sát thành 4 lớp. Nói cách khác, chúng ta xem xét các quyết định cục bộ phi nhị phân. Chọn xác suất tiên nghiệm $\pi_0 = 0,5$, tỷ lệ lai tạo chéo bằng 0,9; tỷ lệ biến dị bằng 0,05 và số nhiễm sắc thể bằng 20 cho mọi trường hợp. Việc mô phỏng được thực hiện với các giá trị khác nhau của ρ trong ma trận tương quan. Số bậc tự do của phân bố Student-t được chọn bằng 3 để có thể thấy rõ sự khác biệt so với phân bố Gauss.

ρ	Quy luật hợp nhất tối ưu	P_e
0	0 0 0 1 0 0 1 1 0 0 1 1 1 1 1 1	0,2868
0,1	0 0 0 1 0 0 1 1 0 1 1 1 0 1 1 1	0,29285
0,5	0 0 0 1 0 0 0 1 0 0 1 1 0 1 1 1	0,31255
0,9	0 0 0 1 0 0 0 1 0 0 1 1 0 1 1 1	0,32509

Bảng 1: Quy luật hợp nhất tối ưu với hệ số tương quan dương

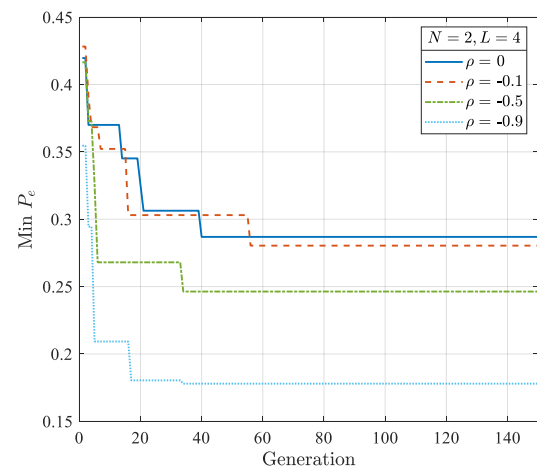
Hình vẽ 3 cho thấy sự hội tụ của xác suất lỗi sau 150 thế hệ di truyền đối với các giá trị dương khác nhau của hệ số tương quan ρ ($\rho=0; 0,1; 0,5; 0,9$). Với $\rho=0$ biểu diễn trường hợp không tương quan, khi đó tất cả các phép quan sát của các cảm biến là độc lập có điều kiện theo giả thuyết H_i . Trong trường hợp này, xác suất lỗi tổng thể nhỏ nhất đạt đến 0,2868 sau 84 thế hệ di truyền. Khi $\rho=0,1$ xác suất lỗi tổng thể nhỏ nhất đạt đến 0,29285 sau 17 thế hệ di

truyền. Tương tự, với $\rho=0,5$ và $\rho=0,9$ xác suất lỗi nhỏ nhất đạt được là 0,31255 và 0,29285 sau lần lượt là 16 và 9 thế hệ di truyền.



Hình 3. Xác suất lỗi P_e với tương quan dương

Từ đây chúng ta có thể thấy rằng xác suất lỗi tổng thể P_e giảm xuống khi mối tương quan giữa các cảm biến giảm. Điều này là do khi hệ số tương quan càng lớn thì mạng càng tiến gần tới dạng một cảm biến duy nhất. Ngoài ra, chú ý rằng khi hệ số tương quan dương và càng lớn thì sẽ mất ít thời gian (các thế hệ di truyền) hơn để hội tụ tới lời giải tối ưu khi so với trường hợp có hệ số tương quan nhỏ.



Hình 4. Xác suất lỗi P_e với tương quan âm

Trong hình vẽ 4, khi tương quan âm ta thấy, xác suất lỗi nhỏ nhất cho 3 trường hợp tương quan khi $\rho = -0,1; -0,5$ và $-0,9$ giảm xuống thành 0,28038, 0,24631 và 0,17792 một cách tương ứng. Từ đây chúng ta nhận ra có một sự cải thiện đáng kể về mặt hiệu năng của mạng cảm biến phân tán khi mối tương quan âm giữa các cảm biến (nghịch tương quan) tăng lên.

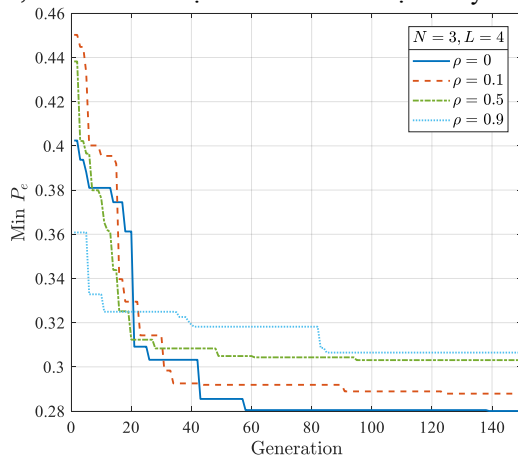
ρ	Quy luật hợp nhất tối ưu	P_e
0	0 0 0 1 0 0 1 1 0 0 1 1 1 1 1 1	0,2868
-0,1	0 0 0 1 0 0 0 1 0 0 1 1 1 1 1 1	0,28038
-0,5	0 0 0 0 0 0 0 1 0 1 1 1 0 1 1 1	0,24631
-0,9	0 0 0 1 0 0 0 1 0 1 1 1 1 1 1 1	0,17792

Bảng 2: Quy luật hợp nhất tối ưu với hệ số tương quan âm

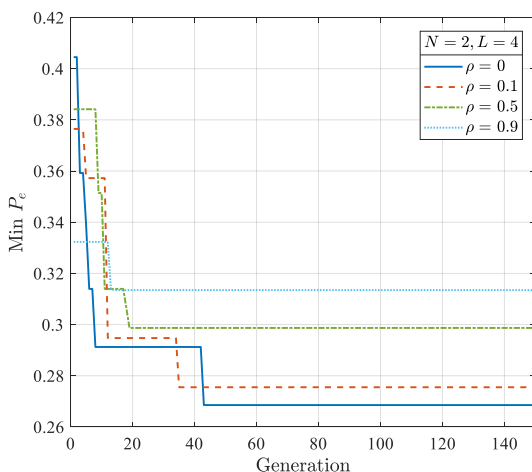
Tương quan nghịch tăng về cơ bản có nghĩa là nếu nhiều trên một cảm biến này đây phép quan sát về phía giả thuyết sai, thì nhiều trên cảm biến kia sẽ đây sẽ đây phép quan sát về phía giả thuyết đúng. Do đó, nếu một cảm biến đưa ra quyết định sai lầm thì cơ hội cho cảm biến kia đưa ra quyết định đúng sẽ nhiều hơn. Trường hợp xấu nhất sẽ là khi nhiều trên cả 2 cảm biến đều nhỏ. Trong trường hợp này, hiệu suất sẽ tương đương với trường hợp không có tương quan. Như vậy, mối tương quan âm trên các phép quan sát sẽ luôn khiến mạng hoạt động tốt hơn hoặc ít nhất bằng trường hợp không tương quan. Quy luật hợp nhất tối ưu được tìm thấy như bảng 2.

Trong hình vẽ 5, với mỗi tương quan dương và khi số cảm biến trong mạng tăng lên $N=3$, xác suất lỗi P_e cho ba trường hợp tương quan khi $\rho = 0,1; 0,5$ và $0,9$ có giá trị tương ứng bằng $0,287; 0,303$ và $0,306$. Rõ ràng, so với trường hợp 2 cảm biến các giá trị xác suất lỗi giảm đi sau lần lượt 124, 95 và 85 thế hệ di truyền. Như vậy có thể thấy, khi tối ưu quy luật hợp nhất với số lượng các cảm biến có trong mạng càng tăng thì xác suất lỗi càng giảm nhưng thời gian hội tụ tới lời giải tối ưu cũng tăng lên.

Trong hình vẽ 6 ta thấy khi $\rho=0,1$ và số bậc tự do của phân bố Student-t được cho bằng 10, xác suất lỗi nhỏ nhất đạt đến 0,2755 sau 35 thế hệ di truyền. Tương tự, cũng với $\rho=0,5$ và $\rho=0,9$ xác suất lỗi nhỏ nhất đạt được là 0,2986 và 0,3139 sau lần lượt là 19 và 13 thế hệ di truyền



Hình 5. Xác suất lỗi P_e khi có 3 cảm biến với tương quan dương



Hình 6. Xác suất lỗi P_e khi số bậc tự do của phân bố Student-t bằng 10

Như vậy, khi so sánh với hình vẽ 3 trong trường hợp số bậc tự do của phân bố là 3 thì xác suất lỗi hầu như không thay đổi nhưng số thế hệ di truyền để giải thuật đạt được lời giải tối ưu thì tăng lên khá nhiều. Có thể kết luận rằng số bậc tự do trong phân bố Student-t không ảnh hưởng nhiều đến hiệu năng hệ thống mà chỉ khiến cho việc tìm kiếm lời giải tối ưu cho bài toán trở nên khó khăn hơn.

V. KẾT LUẬN

Bài toán tối ưu hóa quy luật hợp nhất bằng giải thuật GA trong mạng cảm biến phân tán khi các cảm biến có trong mạng chịu ảnh hưởng của nhiễu tương quan phân bố Student-t cho thấy hiệu năng của hệ thống giảm đi khi mỗi tương quan giữa các cảm biến càng lớn. Khi mỗi tương quan giữa các đài là tương quan âm thì hiệu năng của hệ thống được cải thiện rất nhiều. Số cảm biến trong mạng càng lớn thì càng mất nhiều thời gian để giải thuật đạt tới lời giải tối ưu. Và với nhiễu tương quan phân bố Student-t thì số bậc tự do của phân bố không gây ảnh hưởng nhiều tới hiệu năng hệ thống mà chỉ khiến cho giải thuật tốn nhiều thời gian (các thế hệ di truyền) để đạt tới lời giải tối ưu. Bài toán có thể còn được mở rộng nghiên cứu và tính toán cho trường hợp tối ưu tổng thể toàn mạng với các ngưỡng tại các cảm biến là tối ưu và quy luật hợp nhất tối ưu khi các cảm biến là đồng nhất hoặc không và chịu ảnh hưởng của nhiễu tương quan với các phân bố khác ví dụ như phân bố K, phân bố Weibull.

TÀI LIỆU THAM KHẢO

1. Aalo V. and Viswanathan R. (Jan. 1992), "Asymptotic performance of a distributed detection system in correlated gaussian noise", *IEEE Transactions on Signal Processing*, 40, pp. 211-213.
2. Aldosari S. A. and Moura J. M. F. (April, 2004), Fusion in sensor networks with communication constraints, *Proceedings IEEE/ACM Symposium on Information Processing in Sensor Networks (IPSN'04)*, pp. 108-115.
3. Aldosari S. E. and Moura J. M. F. (May 2004), Detection in decentralized sensor networks, *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 277-80.
4. Aziz A. M., Tummala M., and Cristi R. (1997), Optimal data fusion strategies using multiple-sensor detection systems, *Conference Record of the Thirty-First Asilomar Conference on Signals, Systems & Computers*, pp. 941-945.
5. Chamberland J. F. and Veeravalli V. V. (2003), "Decentralized detection in sensor networks", *IEEE Transactions on Signal Processing*, 51, pp. 407-416.
6. Drakopoulos E. and Lee C. C. (1991), "Optimum multisensor fusion of correlated local decisions", *IEEE Transactions on Aerospace and Electronic Systems*, 27, pp. 593-606.
7. Gnanapandithan N. and Natarajan B. (2006), Impact of local decision rules in distributed sensor networks, *Proceedings of IEEE Vehicular Technology Conference*.
8. Irving W. W. and Tsitsiklis J. N. (1994), "Some properties of optimal thresholds in decentralized detection", *IEEE Transactions on Automatic Control*, 39, pp. 835-838.
9. Iyengar S. S., Kashyap R. L., and Madan R. N. (Sept. 1991), "Distributed sensor networks introduction to the special section", *IEEE Transactions on Systems, Man and Cybernetics*, 21, pp. 1027-1031.
10. Mansouri N. and Fathi M. (June 2003), Simple counting rule for optimal data fusion, pp. 1186-1191.

11. Nithya G. and Balasubramaniam N. (2006), Parallel Genetic Algorithm based Optimal Fusion in Sensor Networks, *IEEE CCNC 2006 proceedings*, IEEE Communications Society.
12. Poor H. V. (1988), *An Introduction to Signal Detection and Estimation*, Springer-Verlag.
13. Reibman A. R. (1987), *Performance and fault-tolerance of distributed detection networks*, Department of Electrical Engineering, Duke University, Durham, NC.
14. Tsitsiklis J. N. (1988), "Decentralized detection by a large number of sensors", *Mathematics of Control, Signals, and Systems*. 1(2), pp. 167-182.
15. Aalo V. and Viswanathan R. (May 1989), "On distributed detection with correlated sensors: Two examples", *IEEE Trans. Aerospace Elect. Syst.* 25, pp. 414-421.
16. Barkat M. and Varshney P. K. (Mar. 1989), "Decentralized CFAR signal detection", *IEEE Trans. Aerospace Electron. Syst.* 25, pp. 141-149.
17. Blum R. S. and Kassam S. A. (May 1992), "Optimum distributed detection of weak signals in dependent sensors", *IEEE Trans. Inform. Theory*. 38, pp. 1066-1079.
18. Geraniotis E. and Chau A. Y. (Dec. 1988), "Distributed detection of weak signals from multiple sensors with correlated observations", *Conf. on Decision and Control*. (in Proc. 27th, Austin, TX), pp. 2501-2506.
19. Kam M., Zhu Q., and Gray W. Steven (July 1992), "Optimal Data Fusion of Correlated Local Decisions in Multiple Sensor Detection Systems", *IEEE Trans. Aerospace and Electron. Syst.* 28, No 3, pp. 916-920.
20. L'Ecuyer P. B., et al. (2010), "Asymptotic robustness of estimators in rare-event simulation", *ACM Trans. Modlmg Comput. Simuln.* 20, p. article 6.
21. Lauer G. S. and Sandell N. R. (1982), "Distributed detection of known signal in correlated noise", *Internal report*. ALPHATECH, Burlington, MA.
22. Liu W., Lu Y., and Fu J. (2002), "Data fusion of multiradar system by using genetic algorithm", *IEEE Transactions on Aerospace and Electronic Systems*. 38, pp. 601-612.
23. Nithya G. (2015), *Data Detection and Fusion in Decentralized Sensor Networks*, Department of Electrical and Computer Engineering, KANSAS STATE UNIVERSITY, Manhattan, Kansas, USA.
24. Swaszek P. F. and Willett P. (Jan. 1995), "Parley as an approach to distributed detection", *IEEE Transactions on Aerospace and Electronic Systems*. 31, pp. 447-457.
25. Tsitsiklis J. N. and Athans M. (1985), "On the complexity of Decentralized Decision making and Detection problems", *IEEE Trans. Aerospace and Electron. Syst.* AES-23, pp. 644-653.
26. Thomopoulos S. C. A., Viswanathan R., and Bougoulas D. K. (Sept. 1989), "Optimal distributed decision fusion", *IEEE Trans. Aerospace Elect. Syst.* 25, pp. 761-765.
27. Varshney P.K. (1997), *Distributed detection and data fusion*, Signal processing and digital filtering, Vol. I, Springer-Verlag Inc, New York, 286.
28. Viswanathan R. and Varshney P. K. (Jan. 1997), "Distributed detection with multiple sensors: Part I- Fundamentals", *Proceedings of the IEEE*. 85(No. 01), pp. 54-63.

APPLY GENETIC ALGORITHM TO OPTIMIZE FUSION RULE IN DISTRIBUTED SENSOR NETWORKS

Abstract: This paper presents a method of applying genetic algorithms to find the optimal merging rules for distributed sensor networks when the observations of the sensors in the network are correlated according to the Student-t distribution. The examined parallel structured distributed sensor network model showed similar results with previous studies that confirmed the suitability of this

method. The content of the article also mentions the survey and evaluation of the influence of the correlation coefficient between the sensors in the network on the convergence speed of the GA algorithm in some cases with specific numerical values. The simulation results show that the error probability P_e of the whole system is strongly influenced by the correlation between sensors. The greater the correlation between the sensors, the greater the probability of P_e error and vice versa. Especially when the sensors have a negative correlation, the system quality is significantly improved. When the number of sensors in the network increases, the convergence speed of the GA algorithm becomes slower to reach the optimal solution. The number of degrees of freedom of the Student-t distribution does not affect the quality of the system much, but it takes more time for the algorithm to reach the optimal solution.

Keywords: Genetic algorithm, fusion rule, distributed sensor network, correlation, Student-t distribution.



Nguyễn Đức Minh tốt nghiệp đại học chuyên ngành Vật lý-Vô tuyến năm 2000 tại Đại học Khoa học tự nhiên- Đại học Quốc gia Hà Nội. Tốt nghiệp Thạc sỹ ngành Công nghệ Điện tử - Viễn thông năm 2006 tại Đại học Công nghệ - Đại học Quốc gia Hà Nội. Nhận bằng tiến sỹ kỹ thuật ngành Kỹ thuật Radar-Dẫn đường năm 2019 tại Học viện Kỹ thuật Quân sự. Hiện đang giảng dạy tại Học viện Công nghệ Bưu chính Viễn thông.

Lĩnh vực nghiên cứu: Lý thuyết phát hiện trong radar nhiều vị trí, xử lý tín hiệu, điện tử máy tính.



Phạm Văn Hùng tốt nghiệp Đại học ngành Điện – Điện tử năm 2011 và Thạc sỹ chuyên ngành Kỹ thuật radar – dẫn đường năm 2016 tại Học viện Kỹ thuật quân sự. Hiện đang công tác tại Khoa Vô tuyến điện tử, Học viện Kỹ thuật quân sự.

Lĩnh vực nghiên cứu: Lý thuyết phát hiện, lý thuyết phát hiện radar nhiều vị trí, xử lý tín hiệu và dữ liệu radar, thiết kế phân cứng.