

**ĐẠI HỌC HUẾ**  
**TRƯỜNG ĐẠI HỌC KHOA HỌC**

**MAI CƯỜNG THỌ**

**NGHIÊN CỨU MỘT SỐ GIẢI PHÁP ĐỊNH TUYẾN  
NHẪM NÂNG CAO HIỆU NĂNG MẠNG FANET**

Ngành: Khoa học máy tính

Mã số: 9480101

**TÓM TẮT LUẬN ÁN TIẾN SĨ**  
**NGÀNH KHOA HỌC MÁY TÍNH**

Người hướng dẫn khoa học:

1. TS. Lê Hữu Bình
2. PGS.TS. Võ Thanh Tú

HUẾ, NĂM 2026

## MỞ ĐẦU

Mạng thiết bị bay không người lái dạng ad-hoc (FANET) gồm các UAV kết nối theo mô hình ad-hoc và truyền dữ liệu về trạm điều khiển mặt đất (GCS) cho các nhiệm vụ giám sát, tìm kiếm cứu nạn và quan trắc môi trường [11], [12], [13], [23], [24]. Đặc trưng di chuyển ba chiều với tốc độ cao, mật độ nút thường thưa và vùng hoạt động rộng làm cấu trúc liên kết biến động nhanh, liên kết vô tuyến kém ổn định và nguy cơ phân mảnh tăng. Hệ quả là tuyến truyền đầu-cuối dễ gián đoạn, tỉ lệ mất gói tăng và độ trễ lớn, ảnh hưởng trực tiếp đến các ứng dụng yêu cầu chất lượng dịch vụ [5], [19], [23].

Các nghiên cứu định tuyến FANET thường được phân thành ba hướng. Hướng dựa trên tô-pô kế thừa các giao thức như AODV, OLSR, DSDV và bổ sung các tiêu chí ETX, SNR, thời gian tồn tại liên kết, tải hàng đợi, năng lượng [2], [19]. Tuy nhiên, cơ chế thiết lập và duy trì tuyến khiến chi phí điều khiển tăng khi tô-pô thay đổi nhanh. Hướng dựa trên vị trí, tiêu biểu là GPSR, cho phép quyết định chuyển tiếp dựa trên vị trí nút hiện tại, láng giềng và đích [16]. Nhiều biến thể như SF-GeoR, UF-GPSR, GPSR+ cải thiện bằng cách đưa thêm hướng di chuyển, độ ổn định liên kết, năng lượng và tải hàng đợi vào hàm đánh giá láng giềng [14], [17], [18]. Dù vậy, các hạn chế như thông tin Hello dễ lạc hậu, vùng rỗng định tuyến trong không gian 3D và mô hình chất lượng liên kết chưa gắn chặt với hành vi truyền lại và tốc độ bit vẫn là thách thức [19], [23]. Hướng học tăng

cường ứng dụng Q-learning để tăng khả năng thích nghi của chính sách chọn *next-hop*, nhưng còn các bài toán mở về thiết kế trạng thái, hàm thưởng gắn với chỉ báo thực tế và cơ chế trao đổi tri thức học đủ tươi mà không làm tăng đáng kể chi phí điều khiển [1], [3], [15].

Trong nhiều ứng dụng, dữ liệu có xu hướng hội tụ về GCS, do đó tuyến UAV về GCS (U2G) là tuyến truyền chủ đạo [12], [23]. Việc chuyên biệt hoá cơ chế định tuyến cho tuyến U2G, kết hợp tổ chức mạng để giảm biến động tô-pô và khai thác thông tin xuyên tầng nhằm tối ưu truyền lại theo điều kiện liên kết tức thời, vẫn cần được nghiên cứu sâu hơn. Từ đó, luận án đặt mục tiêu nghiên cứu và đề xuất các giải pháp định tuyến cho tuyến U2G trong FANET nhằm cải thiện các chỉ tiêu hiệu năng chính như tỉ lệ chuyển gói, độ trễ đầu-cuối và thông lượng. Các mục tiêu cụ thể gồm:

1. Đề xuất và đánh giá một giải pháp định tuyến dựa trên vị trí cho tuyến U2G, tăng cường cơ chế chọn *next-hop* bằng khai thác chất lượng liên kết và tổ chức mạng phù hợp nhằm nâng cao độ ổn định kết nối.
2. Đề xuất và đánh giá một giải pháp định tuyến dùng Q-learning cho tuyến U2G, trong đó chính sách chọn *next-hop* được điều chỉnh theo phản hồi truyền dữ liệu và thông tin chất lượng kênh từ tầng dưới nhằm tăng khả năng thích ứng với sự thay đổi tô-pô.

## **Chương 1 MẠNG FANET VÀ CÁC GIẢI PHÁP ĐỊNH TUYẾN**

### **1.1. Mạng FANET và thách thức định tuyến**

Flying Ad-hoc Network (FANET) là một dạng mạng tùy biến không dây, trong đó các phương tiện bay không người lái (UAV) kết nối với nhau theo mô hình ad-hoc và ít nhất một UAV duy trì liên lạc với trạm điều khiển mặt đất (GCS) hoặc vệ tinh [4], [5]. Trong nhiều ứng dụng như giám sát, tìm kiếm cứu nạn hay quản lý thiên tai, UAV vừa thu thập dữ liệu, vừa chuyển tiếp dữ liệu về GCS để phục vụ giám sát gần thời gian thực [11], [13], [23], [24].

So với MANET và VANET, FANET khác biệt ở môi trường hoạt động ba chiều, độ cơ động cao, mật độ nút thường thưa và liên kết vô tuyến chủ yếu theo đường thẳng (LoS). Độ cao, góc bay và điều kiện môi trường làm chất lượng liên kết biến động mạnh. UAV vừa phải dành năng lượng cho bay, vừa cho truyền thông nên tài nguyên năng lượng hạn chế. Thông tin vị trí cần được cập nhật thường xuyên để phục vụ quyết định định tuyến.

Các đặc điểm trên dẫn tới nhiều thách thức cho định tuyến. Tô-pô thay đổi nhanh làm các tuyến đầu-cuối dễ bị gián đoạn. Mật độ nút thấp và phạm vi hoạt động rộng làm tăng nguy cơ phân mảnh mạng, có thể khiến một số UAV mất kết nối tạm thời với GCS. Liên kết kém ổn định làm tăng lỗi gói và số lần truyền lại, kéo theo độ trễ và chi phí băng thông. Các cơ chế duy trì bảng định tuyến hoặc trao đổi tín hiệu điều khiển thường xuyên cũng làm tăng gánh nặng cho

kênh truyền và tiêu hao năng lượng. Trong nhiều kịch bản, phần lớn lưu lượng hội tụ về GCS nên tuyến UAV-GCS có vai trò trọng yếu.

Hiệu năng định tuyến trong FANET thường được đánh giá thông qua các chỉ số cơ bản là tỷ lệ chuyển gói thành công (PDR), độ trễ đầu-cuối và thông lượng [23].

## 1.2. Các hướng tiếp cận nâng cao hiệu năng định tuyến

Từ các đặc điểm và thách thức nêu trên, các nghiên cứu về định tuyến trong FANET chủ yếu tập trung vào ba hướng chính [1], [2], [19], [23]: (i) cải tiến các giao thức dựa trên cấu trúc liên kết, (ii) cải tiến các giao thức dựa trên vị trí và (iii) ứng dụng học tăng cường trong định tuyến. Mỗi hướng có ưu điểm riêng nhưng vẫn còn nhiều hạn chế khi áp dụng cho môi trường FANET có tô-pô biến động nhanh và yêu cầu cao về độ ổn định tuyến UAV-GCS.

### 1.2.1. Hướng giao thức dựa trên cấu trúc liên kết

Nhóm định tuyến dựa trên cấu trúc liên kết kế thừa các giao thức trong MANET như OLSR, AODV, DSDV [6], [8], [25] với cơ chế xây dựng và duy trì bảng định tuyến từ trạng thái liên kết. Khi áp dụng cho FANET, nhiều công trình đã đề xuất các giao thức cải tiến nhằm phản ánh tốt hơn chất lượng và độ ổn định liên kết.

Một số nghiên cứu tích hợp thông tin chuyển động để ước lượng thời gian tồn tại liên kết và kết hợp với các độ đo như ETX, ưu tiên các liên kết vừa có chất lượng tốt vừa ổn định hơn [32], [33]. Một số nghiên cứu khác khai thác thông tin từ tầng vật lý và MAC như RSSI, SNR, tải hàng đợi để xây dựng tiêu chí định tuyến đa yếu tố, qua đó lựa chọn các tuyến ít tắc nghẽn và đáng tin cậy hơn [26], [29]. Cũng có

các giải pháp tập trung tối ưu năng lượng trên tuyến, cân bằng năng lượng còn lại và tránh sử dụng các nút có nguy cơ cạn pin [22]. Trong [CT5], một biến thể OLSR-Vr đã được đề xuất, trong đó các tham số thời gian hiệu lực giữ thông tin liên kết và tô-pô được điều chỉnh động theo dải vận tốc UAV nhằm cải thiện độ tươi thông tin, qua đó nâng cao hiệu quả định tuyến của OLSR trong một số kịch bản FANET. Nhìn chung, các cải tiến dựa trên cấu trúc liên kết đã góp phần nâng cao độ ổn định tuyến và chất lượng truyền thông trong một số điều kiện. Tuy vậy, chúng vẫn phụ thuộc vào cơ chế duy trì bảng định tuyến nên kém hiệu quả khi tô-pô thay đổi quá nhanh hoặc mạng bị phân mảnh. Các giải pháp chủ yếu tập trung vào cải thiện liên kết cục bộ nhưng chưa giải quyết triệt để việc duy trì tuyến đầu-cuối trong các kịch bản UAV di chuyển nhanh, mật độ nút thưa hoặc không gian hoạt động rộng.

### 1.2.2. Hướng giao thức dựa trên vị trí

Định tuyến dựa trên vị trí được đánh giá là phù hợp hơn với FANET nhờ cơ chế quyết định cục bộ và chi phí điều khiển thấp [23]. Mỗi nút dựa trên vị trí của mình, của các láng giềng và của nút đích để chọn nút chuyển tiếp theo các tiêu chí cục bộ, thường dựa trên khoảng cách tới đích, hướng di chuyển hoặc một số yếu tố phụ trợ khác. GPSR [16] là giao thức điển hình với hai chế độ chính là chuyển tiếp tham lam và đi vòng khi gặp vùng trống định tuyến. Trên nền GPSR, nhiều giao thức cải tiến dành cho FANET đã được đề xuất như GPMOR [20], GPSR-PPU [27], SF-GeoR [17], UF-GPSR [18], GPSR+ [14]. Các giao thức cải tiến này bổ sung thêm các tiêu chí đánh giá láng giềng như vị trí dự đoán, thời gian tồn tại liên kết, năng lượng còn lại,

vận tốc và hướng bay, nhằm ưu tiên các nút vừa giúp gói tin tiến gần đích, vừa có liên kết ổn định và đủ năng lượng.

Các cải tiến dựa trên vị trí giúp giảm độ trễ và tăng PDR trong nhiều kịch bản FANET. Tuy nhiên, phần lớn vẫn chủ yếu khai thác đặc trưng không gian và động học, trong khi đánh giá chất lượng liên kết thường chỉ dựa trên khoảng cách hoặc các chỉ báo tức thời như RSSI và SNR. Tác động của tốc độ bit, tỷ lệ lỗi gói và cơ chế truyền lại tới hiệu năng đầu-cuối chưa được khai thác. Bên cạnh đó, chế độ đi vòng của GPSR kém phù hợp trong môi trường ba chiều với di động cao. Các giải pháp hiện tại cũng chưa chuyên biệt hóa cho tuyến UAV-GCS, lưu lượng hội tụ về một đích cố định.

### 1.2.3. Hướng ứng dụng học tăng cường

Gần đây, nhiều công trình đề xuất ứng dụng học tăng cường trong định tuyến FANET để tăng khả năng tự thích nghi với môi trường mạng [1], [28]. Mô hình định tuyến thường được xây dựng tương tự định tuyến dựa trên vị trí nhưng quyết định chọn nút chuyển tiếp được học thông qua tương tác giữa tác tử và môi trường.

Trong số các thuật toán học tăng cường, Q-learning [30] được sử dụng phổ biến. Mỗi nút duy trì một bảng Q, trong đó mỗi giá trị Q biểu diễn chất lượng ước lượng của việc chọn một láng giềng làm nút chuyển tiếp. Q-value được cập nhật dựa trên phần thưởng nhận được sau khi truyền gói, phản ánh các mục tiêu như giảm độ trễ, tăng PDR hay tiết kiệm năng lượng. Các giao thức như QGeo [15], QMR [21], Q-FANET [7] và QTAR [3] đã khai thác cơ chế này thông qua thiết kế hàm phần thưởng đa mục tiêu, mở rộng không gian quan sát và điều chỉnh linh hoạt các tham số học.

Các kết quả công bố cho thấy định tuyến sử dụng học

tăng cường có thể nâng cao hiệu năng trong nhiều kịch bản FANET so với các chiến lược cố định. Tuy nhiên, phần lớn giải pháp vẫn dựa nhiều vào thông tin từ gói điều khiển định kỳ nên việc cập nhật tri thức có thể bị trễ khi cấu trúc mạng biến đổi nhanh. Vấn đề truyền lại gói và lỗi liên kết chưa được mô hình hóa rõ trong hàm phần thưởng. Việc gắn kết trạng thái và phần thưởng với các chỉ báo như PER, tải hàng đợi, tín hiệu các lần truyền lại và chất lượng kênh theo tốc độ bit vẫn là những vấn đề còn chưa được giải quyết đầy đủ.

### 1.3. Giao thức GPSR-RA kết hợp điều chỉnh tốc độ bit

Bên cạnh phân tổng quan, phần này trình bày giao thức GPSR-RA [CT4]. GPSR-RA được xây dựng trên nền GPSR, giữ nguyên nguyên lý định tuyến dựa trên vị trí và kết hợp cơ chế điều chỉnh tốc độ bit theo chất lượng liên kết đo bằng SNIR. Mục tiêu là khai thác khả năng multi-bitrate của chuẩn IEEE 802.11 để cải thiện PDR, thông lượng, và giảm độ trễ.

GPSR-RA áp dụng mô hình xuyên tầng với hai mô-đun chính. Ở tầng mạng, GPSRvar mở rộng bảng láng giềng của GPSR bằng cách lưu thêm thông tin SNIR của liên kết tới từng láng giềng, thu thập từ tầng vật lý thông qua gói Hello. Ở tầng MAC, MACSnir-RA nhận *next-hop*, lấy SNIR tương ứng và quyết định tăng hoặc giảm một bậc tốc độ bit trước khi truyền gói dữ liệu. Các ngưỡng SNIR cho từng mức tốc độ bit được xác định từ quan hệ PER-SNIR, bảo đảm PER xấp xỉ ngưỡng cho phép đối với mỗi tốc độ.

Kết quả mô phỏng trên các kịch bản FANET với số lượng UAV, vận tốc và mật độ khác nhau cho thấy GPSR-RA đạt PDR cao hơn, độ trễ thấp hơn và thông lượng tốt hơn so với GPSR trong nhiều kịch bản khảo sát. Khi chất lượng kênh

suy giảm, cơ chế giảm tốc độ bit kịp thời giúp hạn chế lỗi gói và số lần truyền lại. Khi kênh tốt, việc tăng tốc độ bit cho phép khai thác băng thông hiệu quả hơn mà không cần thay đổi cơ chế định tuyến.

## Chương 2 GIẢI PHÁP ĐỊNH TUYẾN GPSR-CB DỰA TRÊN VỊ TRÍ

### 2.1. Đặt vấn đề

Định tuyến dựa trên vị trí cho phép mỗi nút chuyển tiếp gói tin theo từng chặng dựa trên vị trí của chính nó, các láng giềng và nút đích, không cần duy trì bảng định tuyến end-to-end như các giao thức dựa trên tô-pô. [16], [23].

GPSR là giao thức tiêu biểu của định tuyến vị trí, sử dụng chế độ *Greedy* và *Perimeter*. Khi áp dụng cho FANET, một số biến thể như SF-GeoR, UF-GPSR, GPSR+ đã được đề xuất với các tiêu chí bổ sung (năng lượng, xu hướng chuyển động, rủi ro đứt liên kết, v.v.) [14], [17], [18]. Tuy nhiên, các giao thức này vẫn chủ yếu dựa trên thông tin động học và các ước lượng gián tiếp, ít khai thác trực tiếp các chỉ số chất lượng liên kết (PER/SNIR) và chưa gắn chặt với kiến trúc mạng nhiều cấp backbone UAV trên tuyến UAV-GCS. Cơ chế hỗ trợ truyền lại và khai thác thông tin xuyên tầng (từ MAC/vật lý lên tầng mạng) còn hạn chế.

Trên cơ sở đó, Chương 2 xây dựng giao thức GPSR-CB, một biến thể của GPSR dành cho tuyến UAV-GCS, kết hợp (i) tổ chức mạng theo kiến trúc nhiều cấp backbone UAV, (ii) hàm đánh giá đa tiêu chí cho chế độ *Greedy*, và (iii) các cơ chế xuyên tầng hỗ trợ cập nhật láng giềng và truyền lại gói tin.

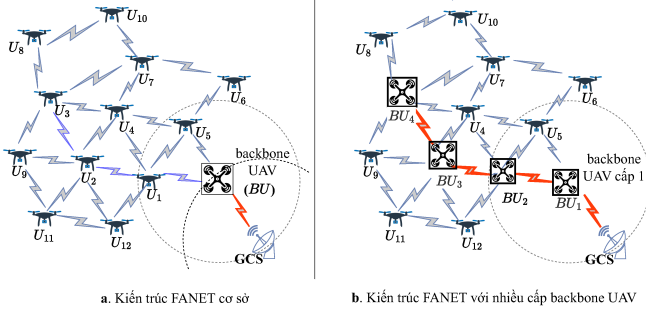
## 2.2. Phân tích một số giao thức định tuyến cải tiến GPSR

SF-GeoR [17] cải tiến GPSR bằng chỉ số ổn định liên kết, kết hợp năng lượng còn lại và xu hướng hội tụ/ly tán giữa các UAV, từ đó ưu tiên các láng giềng có liên kết ổn định hơn. UF-GPSR [18] sử dụng hàm lợi ích đa tiêu chí (năng lượng, độ gần đích, hướng láng giềng, rủi ro mất liên kết, tốc độ trung bình) để lựa chọn *next-hop*. Hai giao thức này cải thiện hiệu năng so với GPSR gốc trong một số kịch bản FANET nhưng các tiêu chí vẫn chủ yếu phản ánh đặc tính động học/vị trí, chất lượng liên kết được đánh giá gián tiếp theo khoảng cách hoặc vùng an toàn, chưa gắn trực tiếp với PER/SNIR hay phản hồi từ tầng MAC.

## 2.3. Cải thiện hiệu năng nhờ tổ chức FANET với nhiều cấp backbone UAV

Trong kiến trúc FANET cơ sở (Hình 2.2a), một UAV backbone ( $BU_1$ ) được chọn để kết nối trực tiếp với GCS; các UAV nhiệm vụ truyền dữ liệu về GCS thông qua  $BU_1$ . Khi phạm vi triển khai rộng, vận tốc lớn hoặc mật độ thưa, tuyến UAV-GCS thường qua nhiều nút trung gian, dễ bị đứt do di động nhanh hoặc phân mảnh mạng, làm giảm PDR và tăng độ trễ.

Luận án đề xuất mở rộng kiến trúc cơ sở này thành *Kiến trúc FANET nhiều cấp backbone UAV* (Hình 2.2b) nhằm khắc phục hạn chế này. Theo đó, thay vì chỉ có một BU, một dãy backbone UAV  $BU_1, BU_2, \dots, BU_n$  được bố trí dọc theo hướng từ GCS vào vùng nhiệm vụ, tạo thành trục backbone GCS- $BU_1$ - $\dots$ - $BU_n$ . Cơ chế triển khai này được thực hiện nhờ mô hình di động tham chiếu theo các điểm neo xác định từ vị trí GCS và vùng nhiệm vụ.



Hình 2.2: FANET cơ sở và FANET với nhiều cấp backbone UAV

Số lượng BU trên trục được ước lượng bởi:

$$n = \left\lceil \frac{M_x}{0.9 R_b} \right\rceil, \quad (2.6)$$

trong đó  $M_x$  là kích thước miền theo trục  $X$ ,  $R_b$  là phạm vi giao tiếp của BU, và hệ số  $0.9 R_b$  đảm bảo liên kết ổn định theo [31]. Các BU được giả định là UAV chuyên trách, năng lượng và độ ổn định cao hơn, di chuyển tham chiếu theo GCS, hình thành một "hành lang" có độ tin cậy cao cho tuyến UAV-GCS. Trong mô phỏng, cấu hình 1-BU (kiến trúc cơ sở) và 4-BU (multi-level backbone) được dùng để đánh giá GPSR-CB và tách bạch đóng góp của tổ chức mạng so với cơ chế định tuyến.

## 2.4. Đề xuất giao thức định tuyến GPSR-CB

### 2.4.1. Ý tưởng thiết kế

GPSR-CB được đề xuất như một cải tiến của GPSR cho tuyến UAV-GCS trong FANET, với hai mục tiêu chính: (i) dẫn gói tin hội tụ sớm về trục backbone UAV ổn định trước

khi đến GCS; (ii) lựa chọn *next-hop* dựa trên hàm đánh giá đa tiêu chí kết hợp thông tin vị trí, chất lượng liên kết và khả năng duy trì chế độ *Greedy*. Giao thức đồng thời tích hợp các cơ chế xuyên tầng để cập nhật nhanh thông tin láng giềng và hỗ trợ truyền lại gói tin sớm thành công.

### 2.4.2. Phương pháp chọn nút láng giềng làm *next-hop*

GPSR-CB đánh giá mỗi láng giềng  $B$  của nút hiện tại  $A$  đối với một “đích ảo”  $D$  (BU gần nhất hoặc GCS) dựa trên ba tiêu chí: độ gần đích ảo, khả năng tiếp tục *Greedy* và chất lượng liên kết.

#### \* Độ gần đích “ảo”

Độ gần đích ảo khi chuyển gói từ  $A$  sang  $B$  được lượng hóa bởi:

$$Prx_{B,D} = \frac{L_{A,D} - L_{B,D}}{txRange}, \quad (2.10)$$

trong đó  $L_{A,D}$  và  $L_{B,D}$  là khoảng cách Euclid từ  $A$  và  $B$  tới đích ảo  $D$ ,  $txRange$  là phạm vi giao tiếp cục bộ của UAV. Giá trị  $Prx_{B,D}$  càng lớn,  $B$  càng giúp gói tin tiến gần backbone/GCS.

#### \* Tỷ lệ lỗi gói

PER giữa một nút  $A$  và láng giềng  $B$  có thể được ước lượng từ BER theo công thức [10]:

$$Per_{A,B} \approx 1 - (1 - BER)^L, \quad (2.11)$$

trong đó  $L$  là kích thước gói tin (tính theo bit).

\* *Khả năng tiếp tục chế độ Greedy*

Khả năng tiếp tục *Greedy* của láng giềng  $B$  được ký hiệu  $Abi_B$ :

$$Abi_B = \frac{|Neighbors\_In\_GR(B, D)|}{|Neighbors\_Of(B)|}, \quad (2.12)$$

với  $Neighbors\_In\_GR(B, D)$  là tập láng giềng của  $B$  nằm trong vùng *Greedy* đối với đích  $D$ ,  $Neighbors\_Of(B)$  là tập toàn bộ láng giềng của  $B$ . Giá trị  $Abi_B$  càng lớn, khả năng duy trì *Greedy* từ  $B$  càng cao, giảm nguy cơ phải chuyển sang *Perimeter*.

\* *Hàm đánh giá láng giềng*

Chất lượng liên kết giữa  $A$  và  $B$  được phản ánh qua tỉ lệ lỗi gói  $Per_{A,B}$  (ước lượng từ BER/SNIR và kích thước gói). Ba tiêu chí được kết hợp thành hàm:

$$f(B, D)_A = \omega_1 \cdot Prx_{B,D} + \omega_2 \cdot Abi_B + \omega_3 \cdot (1 - Per_{A,B}), \quad (2.13)$$

với  $\omega_1 + \omega_2 + \omega_3 = 1$ . Nút  $A$  loại bỏ các láng giềng bị đánh dấu không tin cậy, tìm đích ảo  $D$  (BU gần nhất hoặc GCS), tính  $f(B, D)_A$  cho từng láng giềng  $B$  và chọn *next-hop* là nút có giá trị lớn nhất. Nếu  $A$  thuộc backbone, *next-hop* được ưu tiên là BU cấp thấp hơn hoặc trực tiếp GCS, bảo đảm gói tin đi dọc trục backbone về phía đích.

\* *Quản lý thông tin láng giềng*

GPSR-CB mở rộng bảng thông tin láng giềng thành, lưu trữ vị trí,  $Abi_B$  và  $Per_{A,B}$  cho từng láng giềng  $B$ . Thông tin vị trí và  $Abi$  được quảng bá qua gói Hello định kỳ. Chỉ số  $Per_{A,B}$  được cập nhật xuyên tầng khi nhận gói ở tầng dưới.

### 2.4.3. Cơ chế bổ trợ

Để giảm độ trễ cập nhật  $Abi_B$ , GPSR-CB tận dụng khung ACK của IEEE 802.11: trường Duration (mặc định không dùng trong điều khiển truy nhập) được tái sử dụng để mang giá trị  $Abi_B$  mới nhất. Mỗi khi một gói dữ liệu được truyền thành công, nút gửi trích xuất  $Abi_B$  từ ACK và cập nhật NIT, giúp hàm  $f(B, D)_A$  dựa trên thông tin tươi hơn so với chỉ dùng Hello.

Đối với các lần truyền lại, GPSR-CB áp dụng hai xử lý xuyên tầng: (i) tăng tạm thời công suất phát trong các lần retransmission theo một hệ số nhỏ, giới hạn bởi công suất cực đại, nhằm cải thiện SNIR và tăng xác suất truyền lại thành công; (ii) nếu một *next-hop* vượt quá ngưỡng số lần truyền lại mà vẫn không phản hồi ACK, liên kết tương ứng được đánh dấu tạm thời (gán  $Abi = -1$  trong NIT) để không được chọn ở các lần định tuyến kế tiếp. Khi láng giềng này xuất hiện trở lại trong các gói Hello, giá trị  $Abi$  được cập nhật theo trạng thái mới.

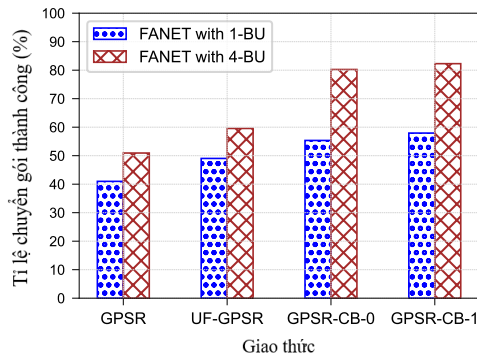
## 2.5. Mô phỏng và phân tích kết quả

### 2.5.1. Tham số và kịch bản mô phỏng

GPSR-CB được cài đặt trên OMNeT++ 6.0/INET 4 và so sánh với GPSR, UF-GPSR theo các chỉ số: tỷ lệ chuyển gói thành công (PDR), thông lượng, trễ đầu-cuối (EED) và năng lượng tiêu thụ. Hai biến thể được khảo sát là GPSR-CB-0 (không điều chỉnh công suất) và GPSR-CB-1 (có điều chỉnh công suất truyền lại, sau đây gọi là GPSR-CB). FANET được triển khai trong không gian  $1000 \times 1000 \times 1000$  m, với 50-100 UAV di chuyển theo RandomWaypoint 3D, phạm vi giao tiếp 250 m. Các kịch bản xem xét hai kiến trúc 1-BU và 4-BU, ba dải vận tốc [5-10], [15-20], [25-30] m/s.

## 2.5.2. Kết quả và phân tích

Trên kiến trúc 1-BU, khi mật độ UAV tăng, PDR của tất cả giao thức đều tăng nhưng GPSR-CB-0 và GPSR-CB-1 luôn đạt PDR và thông lượng cao hơn GPSR và UF-GPSR ở mọi mức mật độ. GPSR-CB kết hợp thông tin vị trí với PER và Ability nên ưu tiên được các liên kết có chất lượng tốt và các láng giềng có khả năng tiếp tục *Greedy* cao, giảm số lần phải chuyển sang *Perimeter* và giảm số lần truyền lại không hiệu quả. Độ trễ đầu-cuối của GPSR-CB duy trì ở mức thấp và ổn định hơn. Mức tiêu thụ năng lượng trung bình thấp hơn hoặc tương đương với UF-GPSR và GPSR, do giảm được số lần truyền lại và độ dài tuyến.



Hình 2.13. Tỷ lệ chuyển gói thành công trên FANET có 1-BU và 4-BU

Khi chuyển sang kiến trúc 4-BU, trục backbone ổn định giúp cải thiện hiệu năng của các giao thức. Tuy nhiên, mức cải thiện của GPSR-CB rõ hơn do giao thức khai thác được tiêu chí đích ảo và cơ chế ưu tiên tuyến qua BU. Hình 2.13 minh họa sự khác biệt về PDR giữa cấu hình 1-BU và 4-BU.

Kết quả ở Hình 2.13 cho thấy cấu hình FANET với nhiều cấp backbone UAV giúp cải thiện rõ rệt PDR so với kiến

trúc cơ sở 1-BU. GPSR-CB đạt mức cải thiện lớn nhất, PDR trung bình tăng từ khoảng 55–58% lên 80–82%, tức khoảng 24%. Trong khi đó, GPSR và UF-GPSR chỉ tăng khoảng 10%. Điều này cho thấy GPSR-CB khai thác hiệu quả trục backbone nhiều cấp để duy trì tuyến U2G ổn định hơn.

Trong các kịch bản thay đổi vận tốc, PDR của các giao thức đều giảm khi vận tốc UAV tăng do liên kết kém ổn định hơn, nhưng GPSR-CB vẫn duy trì kết quả tốt hơn GPSR và UF-GPSR. Tổng hợp các kết quả cho thấy GPSR-CB cải thiện PDR, thông lượng và EED trên cả kiến trúc 1-BU và 4-BU, trong khi độ phức tạp xử lý vẫn tuyến tính theo số lán giếng. Tuy nhiên, giao thức sử dụng trọng số cố định nên khả năng tự thích nghi theo điều kiện mạng còn hạn chế.

## **Chương 3 GIẢI PHÁP ĐỊNH TUYẾN QLR-FANET SỬ DỤNG HỌC TĂNG CƯỜNG**

### **3.1. Giới thiệu**

Trên cơ sở mô hình hoá bài toán định tuyến đa chặng bằng Q-learning (tác tử là gói tin, trạng thái là nút hiện tại đang giữ gói, hành động là lựa chọn *next-hop*), chương này trước hết phân tích một số giao thức tiêu biểu (QGeo, QMR, Q-FANET), từ đó chỉ ra các hạn chế chung khi áp dụng học tăng cường trong môi trường liên kết biến động nhanh. Tiếp theo, chương trình bày thiết kế giao thức QLR-FANET với hàm thưởng hai thành phần, cơ chế lan truyền tri thức theo sự kiện qua gói D-ACK và cơ chế xuyên tầng điều chỉnh bitrate. Cuối cùng là phần mô phỏng đánh giá hiệu năng.

### 3.2. Phân tích một số giao thức định tuyến sử dụng học tăng cường

Một số giao thức định tuyến cho FANET đã ứng dụng Q-learning trên nền định tuyến theo vị trí. QGeo [15] sử dụng phần thưởng dựa trên tốc độ tiến của gói tin về phía đích; mỗi nút chia sẻ vị trí, vận tốc, thông tin lỗi và  $\max Q_Y^D$  qua gói Hello, hệ số chiết khấu  $\gamma$  được đặt theo dự đoán duy trì hay đứt liên kết, và *next-hop* được chọn bằng chiến lược  $\epsilon$ -greedy. QMR [21] mở rộng QGeo theo hướng đa mục tiêu: hàm thưởng kết hợp độ trễ một chặng và năng lượng còn lại, tỉ lệ học  $\alpha$  được điều chỉnh theo độ biến động trễ trên từng liên kết, hệ số chiết khấu  $\gamma_Y$  phản ánh mức ổn định tập láng giềng, và việc chọn *next-hop* được thực hiện qua hai pha với ràng buộc vận tốc tiến theo *deadline*. Q-FANET [7] áp dụng thuật toán QNoise+ [9], trong đó giá trị Q được cập nhật từ phần thưởng có trọng số của nhiều bước gần nhất kết hợp với thành phần phụ thuộc SINR; hàm thưởng được thiết kế với ba mức (tới đích, vùng trống định tuyến, các trường hợp còn lại), và *next-hop* vẫn được chọn theo  $\epsilon$ -greedy trên tập ứng viên thoả mãn ràng buộc vận tốc tiến.

Các giao thức trên cho thấy hướng tiếp cận Q-learning có tiềm năng cải thiện PDR, EED và tiêu thụ năng lượng trong một số kịch bản FANET, nhưng cũng chia sẻ một số hạn chế chung. Thứ nhất, tri thức học được (đặc biệt là  $\max Q_Y^D$ ) chủ yếu lan truyền qua gói Hello định kỳ, nên dễ lạc hậu hoặc bị mất khi liên kết thay đổi nhanh hoặc kênh nhiễu mạnh. Thứ hai, hàm thưởng tập trung vào các đại lượng cục bộ trên một chặng (vận tốc tiến, trễ, năng lượng, độ đầy hàng đợi), chưa tách bạch đóng góp của trạng thái cục bộ với khả năng nhận và tiếp tục chuyển gói tới đích của *next-hop*. Thứ

ba, cơ chế xử lý truyền lại gói chủ yếu nằm ở tầng MAC, chưa tận dụng sự kiện truyền lại (retry) như một tín hiệu cho quá trình học, cũng như chưa khai thác khả năng điều chỉnh bitrate xuyên tầng để tăng xác suất truyền lại thành công mà không cần tăng công suất phát.

Những hạn chế này là cơ sở để luận án đề xuất giao thức QLR-FANET, trong đó hàm thưởng được thiết kế hai thành phần (phần thưởng cục bộ và phần thưởng hành động từ *next-hop*), tri thức học được chia sẻ theo sự kiện qua gói D-ACK gắn với gói dữ liệu, và cơ chế xuyên tầng điều chỉnh tốc độ bit được kích hoạt khi phát sinh truyền lại gói.

### 3.3. Đề xuất giao thức định tuyến QLR-FANET

#### 3.3.1. Ý tưởng thiết kế

QLR-FANET được xây dựng nhằm khắc phục ba hạn chế nêu trên. Thứ nhất, mô hình RL sử dụng hàm phần thưởng hai thành phần, tách bạch phần thưởng cục bộ (do nút gửi đánh giá từ quan sát tại chỗ) và phần thưởng hành động (do *next-hop* phản hồi, gắn với khả năng chuyển tiếp đến đích). Cách tách này giúp tác tử không chỉ tối ưu bước chuyển tiếp hiện tại mà còn ưu tiên những *next-hop* có tiềm năng duy trì tuyến ổn định trong các chặng kế tiếp.

Thứ hai, thay vì chỉ lan truyền tri thức học qua Hello, QLR-FANET bổ sung kênh phản hồi theo sự kiện thông qua gói D-ACK gắn trực tiếp với gói dữ liệu. Mỗi khi một gói được chuyển tiếp thành công, *next-hop* gửi D-ACK chứa thông tin *ActionReward* và giá trị  $max Q$  hiện thời về đích, cho phép nút gửi cập nhật Q-value bám sát trạng thái liên kết hiện tại hơn.

Thứ ba, QLR-FANET tích hợp cơ chế xuyên tầng sử dụng tín hiệu truyền lại (Retry) từ tầng MAC. Các lần truyền lại

thất bại được ánh xạ thành phần thưởng âm, đồng thời kích hoạt điều chỉnh bitrate theo cơ chế multi-rate IEEE 802.11, giảm tốc độ bit tạm thời thay vì tăng công suất phát. Việc này vừa tăng xác suất truyền lại thành công, vừa hạn chế gây nhiễu và phù hợp với UAV có giới hạn phần cứng.

### 3.3.2. Mô hình định tuyến và dữ liệu học tại nút

QLR-FANET mô hình hoá tác tử theo tiến trình chuyển tiếp gói tin, trong đó trạng thái tương ứng với nút hiện tại và hành động là chọn một láng giềng làm *next-hop*. Mỗi nút duy trì hai cấu trúc dữ liệu chính. Thứ nhất là bảng quan sát (*Observation*) thu thập từ Hello và từ các đo lường liên kết, phục vụ tính thưởng. Thứ hai là bảng Q (*Q-table*) lưu các Q-value theo tập láng giềng, nhằm hỗ trợ ra quyết định phân tán. Việc cập nhật Q-value được kích hoạt theo sự kiện gắn với truyền dữ liệu, thông qua gói D-ACK hoặc tín hiệu truyền lại.

QLR-FANET sử dụng phần thưởng hai lớp gồm *Local-Reward*  $\mathcal{L}_A(B)$  và *ActionReward*  $\mathcal{N}(B)$  cho mỗi láng giềng  $B$  của nút hiện tại  $A$  đối với đích  $D$ .

Hai thành phần thưởng được pha trộn bởi hệ số  $\Lambda$ :

$$\mathcal{R}_A(B) = \Lambda \cdot \mathcal{L}_A(B) + (1 - \Lambda) \cdot \mathcal{N}(B), \quad (3.21)$$

giá trị  $\mathcal{R}_A(B)$  là phần thưởng dùng để cập nhật Q-value theo quy tắc Q-learning.

*LocalReward* phản ánh quan sát cục bộ tại nút hiện tại khi xem xét một láng giềng ứng viên, gồm bốn thành phần: tỉ lệ lỗi gói trên liên kết ( $Per_{\text{link}(A,B)}$ ), vận tốc tương đối ( $v_{B,A}$ ), mức tiến gần đích ( $d_{(B,A)}$ ) và mức độ đầy hàng đợi tại *next-hop* ( $Load_B$ ). Bốn thành phần đều được chuẩn hoá về miền

$[0, 1]$  và kết hợp tuyến tính theo bộ trọng số  $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ :

$$\begin{aligned} \mathcal{L}_A(B) = & \lambda_1 \cdot (1 - Per_{\text{link}(A,B)}) + \lambda_2 \cdot (1 - v_{B,A}) \\ & + \lambda_3 \cdot d_{(B,A) \rightarrow D} + \lambda_4 \cdot (1 - Load_B), \end{aligned} \quad (3.22)$$

*ActionReward* phản ánh phản hồi trực tiếp từ *next-hop* về khả năng tiếp tục chuyển gói về đích, dựa trên chỉ số  $\mathcal{P}_B(D)$  và chất lượng liên kết ngược  $Per_{\text{link}(B,A)}$ . Cụ thể:

$$\mathcal{N}(B) = \begin{cases} -1, & \text{nếu } \mathcal{P}_B(D) = 0, \\ 1, & \text{nếu } \mathcal{P}_B(D) > \mathcal{P}_\theta, \\ k - Per_{\text{link}(B,A)}, & \text{các trường hợp còn lại,} \end{cases} \quad (3.23)$$

trong đó  $k$  là hằng số hiệu chỉnh và  $\mathcal{P}_\theta \in (0, 1)$  là ngưỡng khả năng chuyển tiếp.

### 3.3.3. Lựa chọn nút chuyển tiếp next-hop

Việc lựa chọn *next-hop* trong QLR-FANET được thực hiện trên cơ sở kết hợp Q-value với thông tin không gian. Tại mỗi nút  $A$ , các láng giềng thuộc ba nút gần nhất mà gói vừa đi qua và các láng giềng có  $\mathcal{P}_B(D) = 0$  được loại khỏi tập ứng viên để tránh vòng lặp và loại bỏ sớm các hướng bế tắc. Với mỗi ứng viên còn lại, Q-value  $Q_A(B, D)$  được nhân với chỉ số độ gần đích  $d_{(B,A) \rightarrow D}$  để thu được Q-value có trọng số, từ đó xây dựng phân phối xác suất chọn hành động theo chính sách Softmax. *Next-hop* được chọn là nút có xác suất lớn nhất trong phân phối này. Cách tiếp cận này cho phép ưu tiên các nút vừa có giá trị học cao, vừa tiến gần đích, đồng thời vẫn duy trì được thành phần khám phá có điều hướng trong môi trường liên kết biến động.

### 3.3.4. Các cơ chế hỗ trợ

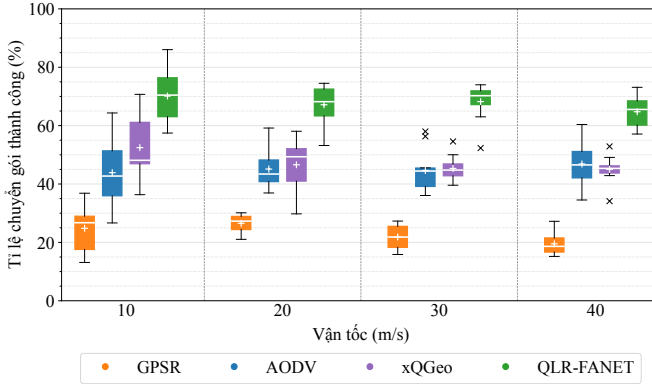
QLR-FANET bổ sung hai cơ chế hỗ trợ gắn với tín hiệu truyền lại từ tầng MAC. Thứ nhất, khi gói phải truyền lại, tác tử ghi nhận đây là một quyết định chưa tối ưu và gán cho liên kết gây truyền lại một mức phạt âm cố định, qua đó làm giảm dần xác suất chọn lại liên kết này trong các lần định tuyến tiếp theo mà không loại bỏ ngay các láng giềng vẫn còn khả dụng. Thứ hai, giao thức tận dụng khả năng multi-rate của IEEE 802.11 để điều chỉnh bitrate: nếu không nhận được ACK sau lần truyền đầu tiên, nút gửi hạ tốc độ bit xuống mức liền kề thấp hơn. Sau khi truyền lại thành công, tốc độ được khôi phục về giá trị mặc định.

### 3.4. Mô phỏng và đánh giá hiệu năng

Hiệu năng của QLR-FANET được đánh giá bằng mô phỏng trên mô hình FANET kích thước  $1500\text{ m} \times 1500\text{ m} \times 120\text{ m}$ , số lượng UAV từ 25 đến 45, vận tốc di chuyển 10–40 m/s và mô hình di động MassMobility. Liên kết vô tuyến sử dụng IEEE 802.11g multi-rate với công suất phát 10 dBm, mô hình nhiễu nền Thermal noise và ngưỡng SINR 4 dB. Tần số phát Hello là 1 Hz. Các tham số học tăng cường gồm tỉ lệ học, hệ số chiết khấu và trọng số trong hàm thưởng được giữ cố định.

QLR-FANET được đối sánh với GPSR, AODV và xQ-Geo. Trong đó, GPSR đại diện cho nhóm định tuyến dựa trên vị trí, AODV đại diện cho nhóm định tuyến dựa trên cấu trúc liên kết, còn xQGeo là giao thức đối sánh theo hướng định tuyến có sử dụng Q-learning. Cách lựa chọn này cho phép đánh giá QLR-FANET so với cả các giao thức định tuyến truyền thống và một giao thức có cơ chế học gần hơn về bản chất.

Kết quả khảo sát theo vận tốc cho thấy QLR-FANET duy trì PDR cao hơn so với GPSR, AODV và xQGeo trong toàn bộ dải vận tốc 10–40 m/s. Hình 3.9 minh họa kết quả PDR của các giao thức khi vận tốc UAV thay đổi.



Hình 3.9. Tỷ lệ chuyển gói thành công theo vận tốc

Kết quả ở Hình 3.9 cho thấy tại vận tốc 30 m/s với 25 nút, PDR trung bình của GPSR, AODV, xQGeo và QLR-FANET lần lượt là 22.8%, 44.6%, 45.3% và 68.2%. Kết quả này phản ánh khả năng thích nghi của QLR-FANET trong môi trường FANET có tô-pô biến động nhanh. Trong khi GPSR và AODV chịu ảnh hưởng bởi thông tin định tuyến lỗi thời hoặc quá trình tái thiết lập đường đi, xQGeo tuy có sử dụng Q-learning nhưng chưa khai thác đầy đủ phản hồi truyền dữ liệu từ tầng liên kết. QLR-FANET cải thiện PDR nhờ kết hợp hàm thưởng hai thành phần, phản hồi D-ACK, cơ chế phạt liên kết kém và điều chỉnh tốc độ bit khi truyền lại.

Trong khảo sát theo mật độ nút với 25, 35 và 45 UAV ở vận tốc 30 m/s, PDR và thông lượng của QLR-FANET tăng theo mật độ và luôn cao hơn các giao thức đối sánh. Với 45 nút, PDR của GPSR, AODV, xQGeo và QLR-FANET

lần lượt khoảng 31%, 60%, 67% và 93%; thông lượng trung bình tương ứng khoảng 70, 135, 155 và 211 Kb/s. Số liên kết bị gián đoạn trung bình trên toàn mạng dưới QLR-FANET cũng thấp hơn đáng kể, phản ánh khả năng lựa chọn tuyến ổn định hơn trong môi trường liên kết biến động nhanh.

### 3.5. So sánh QLR-FANET với GPSR-CB

Để đối chiếu trực tiếp hai giao thức đề xuất, luận án sử dụng lại kịch bản mô phỏng của GPSR-CB trên kiến trúc FANET cơ sở (1-BU).

Trong kịch bản khảo sát ảnh hưởng của mật độ UAV tại dải vận tốc 5–10 m/s, với ba mức mật độ 50, 75 và 100 UAV, PDR của cả hai giao thức đều có xu hướng tăng khi mật độ UAV tăng do số lượng láng giềng khả dụng và khả năng duy trì kết nối đa chặng được cải thiện. Trong hầu hết các mức mật độ khảo sát, QLR-FANET đạt PDR trung bình cao hơn GPSR-CB. Cụ thể, tại mật độ 75 UAV, QLR-FANET đạt khoảng 63%, cao hơn GPSR-CB khoảng 5%.

Khi ở dải vận tốc cao, PDR của cả GPSR-CB và QLR-FANET đều suy giảm do liên kết thay đổi nhanh và xác suất đứt tuyến tăng. Chênh lệch PDR giữa hai giao thức có xu hướng thu hẹp, cho thấy lợi thế học thích nghi của QLR-FANET giảm dần khi trạng thái liên kết biến động quá nhanh. Trong khi đó, GPSR-CB vẫn duy trì ưu thế về trễ đầu-cuối thấp và có khả năng phát huy hiệu quả hơn khi khai thác kiến trúc backbone UAV nhiều cấp. Vì vậy, hai giao thức thể hiện hai hướng tiếp cận bổ sung: GPSR-CB phù hợp với định tuyến U2G có tổ chức backbone rõ ràng, còn QLR-FANET phù hợp với kịch bản cần khả năng học

thích nghi từ phản hồi truyền dữ liệu.

## KẾT LUẬN

Vấn đề nâng cao hiệu năng định tuyến trên tuyến U2G trong mạng FANET là cần thiết và có ý nghĩa thực tiễn đối với các ứng dụng giám sát, tìm kiếm cứu nạn và an ninh-quốc phòng. Qua thời gian nghiên cứu, luận án đã đề xuất và kiểm chứng một số giải pháp định tuyến mới, đồng thời mở ra một số hướng phát triển tiếp theo.

### a) Kết quả đạt được

Luận án tập trung nghiên cứu cơ chế định tuyến thích ứng cho FANET với hai hướng chính là định tuyến dựa trên vị trí và định tuyến sử dụng học tăng cường, gắn với đặc thù luồng dữ liệu hội tụ về GCS. Các kết quả đạt được đã đáp ứng các mục tiêu đề ra ban đầu, cụ thể luận án đã:

- (1) Đề xuất giao thức định tuyến GPSR-CB, cải tiến GPSR dành cho tuyến U2G, được thiết kế để khai thác kiến trúc FANET nhiều cấp backbone UAV nhằm tận dụng tốt hơn các liên kết ổn định trên tuyến U2G. Cụ thể GPSR-CB cải tiến cách đánh giá láng giềng trong chế độ *Greedy*: nút chuyển tiếp được chọn bằng một hàm đa tiêu chí, kết hợp độ gần tới “đích ảo” (BU gần nhất hoặc GCS), chỉ số *Ability* và chất lượng liên kết lượng hóa bằng PER. GPSR-CB áp dụng cơ chế tối ưu truyền lại bằng cách tăng công suất phát mỗi khi gói tin phải truyền lại, nhằm nâng xác suất truyền thành công.
- (2) Đề xuất giao thức định tuyến QLR-FANET, ứng dụng Q-learning tại bộ chọn *next-hop*. Trong giao thức này,

gói tin được xem như một tác tử học từ phản hồi truyền dữ liệu, hàm thưởng được thiết kế kết hợp phần thưởng cục bộ và phần thưởng hành động từ *next-hop*. Thông tin max  $Q$  được chia sẻ qua gói D-ACK và gắn với cơ chế phạt khi truyền lại, điều chỉnh tốc độ bit theo tín hiệu tầng dưới. Cách thiết kế này cho phép chính sách định tuyến tự thích nghi dần theo điều kiện mạng và cải thiện hiệu năng truyền thông.

- (3) Đề xuất kiến trúc FANET nhiều cấp backbone UAV, trong đó các backbone UAV được tổ chức theo cấp bậc dọc theo hướng về GCS, hình thành một trục chuyên tiếp ổn định hơn. Kết quả mô phỏng cho thấy kiến trúc backbone UAV nhiều cấp giúp cải thiện hiệu năng mạng.
- (4) Đề xuất giao thức GPSR-RA, cải tiến GPSR theo hướng xuyên tầng với cơ chế điều chỉnh tốc độ bit dựa trên các ngưỡng SNIR. Bằng cách giảm tốc độ bit khi chất lượng kênh suy giảm và tăng khi kênh cải thiện trước mỗi lần truyền, GPSR-RA góp phần nâng cao độ tin cậy truyền dẫn và giảm độ trễ đầu-cuối.

Tóm lại, luận án đã đạt được mục tiêu nghiên cứu là đề xuất và kiểm chứng các giải pháp định tuyến mới nhằm cải thiện hiệu năng mạng FANET trên tuyến U2G. Các kết quả này góp phần bổ sung cơ sở khoa học cho định tuyến FANET theo định hướng U2G và chỉ ra lợi ích của việc kết hợp tổ chức mạng có chủ đích, đánh giá chất lượng liên kết theo PER và tốc độ bit với cơ chế học thích nghi phân tán.

Tuy vậy, các kết quả và kết luận trên được xây dựng cho bài toán định tuyến đơn đích trên tuyến U2G trong mạng FANET và không xét riêng các kịch bản có yêu cầu đặc thù

về bảo mật định tuyến, hỗ trợ đa đích hay phân biệt nhiều lớp dịch vụ.

### **b) Hướng phát triển**

Trên cơ sở các kết quả đã đạt được, luận án định hướng một số hướng nghiên cứu tiếp theo như sau:

- (1) Trong GPSR-CB và QLR-FANET, các trọng số trong hàm đánh giá và hàm thưởng hiện đang được lựa chọn cố định. Việc phát triển các cơ chế tự điều chỉnh các tham số này là một trong những hướng nghiên cứu tiếp theo cần được xem xét.
- (2) Mở rộng phạm vi áp dụng các cơ chế định tuyến đề xuất trong các bối cảnh có thêm yêu cầu về bảo mật định tuyến, hỗ trợ đa đích hoặc phân biệt lớp dịch vụ. Việc kết hợp các yêu cầu này với kiến trúc FANET nhiều cấp backbone UAV là một hướng có thể được nghiên cứu nhằm xây dựng các giải pháp phù hợp hơn cho các ứng dụng FANET phức tạp.

## **DANH MỤC CÔNG BỐ KHOA HỌC CỦA TÁC GIẢ**

- [CT1] **Mai Cuong Tho**, Le Huu Binh, Tu T. Vo (2025). "GPSR-CB: A novel routing algorithm for FANET using cross-layer models in combination with multi-level backbone UAV". *Ad Hoc Networks*, 173, 103828.
- [CT2] **Mai Cuong Tho**, Ly T. H. Nguyen, Le Huu Binh, Tu T. Vo (2025). "QLR-FANET: A Q-learning and

rate control-based routing protocol for flying ad hoc network”. *ETRI Journal*.

- [CT3]**Tho C. Mai**, Ly T. H. Nguyen, Cuong Q. Nguyen, Le Huu Binh, Tu T. Vo, (2024). ”An Efficient and Low-Cost Hierarchical FANET Architecture Using the Attached Mobility Model to Create the Core Layer for Connecting UAVs and Ground Control Stations”. *Journal of Advances in Information Technology*, 15(4), 500-510.
- [CT4]**Mai Cường Thọ**, Lê Hữu Bình, Võ Thanh Tú (2025), ”GPSR-RA: Giao thức định tuyến dựa trên vị trí kết hợp điều chỉnh tốc độ bit kênh truyền cho FANET sử dụng mô hình thiết kế xuyên tầng”, *Tạp chí Khoa học, Đại học Huế*, Tập 134, Số 2A(2025).
- [CT5]**Mai Cường Thọ**, Nguyễn Thị Hương Lý, Võ Thanh Tú (2023), “Điều chỉnh tham số thời gian hiệu lực trong giao thức OLSR nhằm cải thiện hiệu năng định tuyến FANET”, *Tạp chí Khoa học và Công nghệ, Trường Đại học Khoa học, Đại học Huế*, Tập 22(1B), Trang 105-116.
- [CT6]**Mai Cường Thọ**, Nguyễn Thị Hương Lý, Nguyễn Quốc Cường, Lê Hữu Bình, Võ Thanh Tú (2023).”Cải thiện hiệu năng FANET ứng dụng trong tìm kiếm cứu nạn: Tham số mô hình di động và chiến lược triển khai”, *Toàn văn kỷ yếu hội nghị FAIR2023*.

**HUE UNIVERSITY  
UNIVERSITY OF SCIENCES**

**MAI CUONG THO**

**RESEARCH ON ROUTING SOLUTIONS FOR  
IMPROVING FANET PERFORMANCE**

Major: Computer Science

Code: 9480101

**SUMMARY OF DOCTORAL THESIS  
IN COMPUTER SCIENCE**

Scientific supervisors:

1. Dr. Le Huu Binh
2. Assoc. Prof. Dr. Vo Thanh Tu

HUE CITY, 2026

## INTRODUCTION

A Flying Ad-hoc Network (FANET) consists of UAVs connected in an ad-hoc manner and transmitting data to a Ground Control Station (GCS) for missions such as surveillance, search and rescue, and environmental monitoring [11], [12], [13], [23], [24]. The 3D mobility of UAVs at high speeds, together with typically sparse node density and large operating areas, causes rapid topology changes, unstable wireless links, and an increased risk of network fragmentation. As a result, end-to-end routes are prone to disruption, packet loss increases, and delay becomes larger, directly affecting applications with QoS requirements [5], [19], [23].

Routing studies in FANETs are commonly classified into three main approaches. The topology-based approach inherits protocols such as AODV, OLSR, and DSDV, and further incorporates metrics such as ETX, SNR, link lifetime, queue load, and energy [2], [19]. However, route establishment and maintenance mechanisms increase control overhead when the topology changes rapidly. The position-based approach, represented by GPSR, enables forwarding decisions based on the positions of the current node, its neighbors, and the destination [16]. Several variants, such as SF-GeoR, UF-GPSR, and GPSR+, improve neighbor evaluation by incorporating movement direction, link stability, energy, and queue load into the evaluation function [14], [17], [18]. Nevertheless, limitations such

as outdated Hello information, routing voids in three-dimensional space, and link-quality models that are not tightly coupled with retransmission behavior and bit rate remain challenging issues [19], [23]. The reinforcement-learning-based approach applies Q-learning to improve the adaptability of the *next-hop* selection policy. However, open issues remain in state design, reward functions associated with practical indicators, and mechanisms for exchanging sufficiently fresh learning knowledge without significantly increasing control overhead [1], [3], [15].

In many applications, data tends to converge toward the GCS; therefore, the UAV-to-GCS (U2G) route is the dominant data transmission route [12], [23]. Further research is still needed on specializing routing mechanisms for U2G transmission, combining network organization to reduce topology fluctuation, and exploiting cross-layer information to optimize retransmissions according to instantaneous link conditions. Based on these considerations, this thesis aims to study and propose routing solutions for U2G transmission in FANETs in order to improve key performance indicators, including packet delivery ratio, end-to-end delay, and throughput. The specific objectives are as follows:

1. To propose and evaluate a position-based routing solution for U2G transmission, enhancing the *next-hop* selection mechanism by exploiting link quality and an appropriate network organization to improve connection stability.
2. To propose and evaluate a Q-learning-based routing solution for U2G transmission, in which the *next-hop* se-

lection policy is adjusted according to data transmission feedback and lower-layer channel-quality information to improve adaptability to topology changes.

## **Chapter 1 FANET NETWORKS AND ROUTING SOLUTIONS**

### **1.1. FANET Networks and Routing Challenges**

A Flying Ad-hoc Network (FANET) is a type of wireless ad-hoc network in which unmanned aerial vehicles (UAVs) are interconnected in an ad-hoc manner and at least one UAV maintains communication with a Ground Control Station (GCS) or a satellite [4], [5]. In many applications, such as surveillance, search and rescue, or disaster management, UAVs both collect data and forward it to the GCS to support near-real-time monitoring [11], [13], [23], [24].

Compared with MANETs and VANETs, FANETs are characterized by three-dimensional operating environments, high mobility, typically sparse node density, and wireless links that are mainly line-of-sight (LoS). Flight altitude, flight angle, and environmental conditions cause link quality to fluctuate significantly. UAVs must allocate energy to both flight and communication, which makes energy resources limited. Position information must also be updated frequently to support routing decisions.

These characteristics lead to several routing challenges. Rapid topology changes make end-to-end routes prone to disruption. Low node density and large operating areas increase the risk of network fragmentation, which may tem-

porarily disconnect some UAVs from the GCS. Unstable links increase packet errors and retransmissions, thereby increasing delay and bandwidth cost. Routing-table maintenance mechanisms or frequent control signaling also impose additional burden on the wireless channel and consume energy. In many scenarios, most traffic converges toward the GCS; therefore, the UAV-to-GCS route plays a critical role.

Routing performance in FANETs is commonly evaluated using basic metrics, including packet delivery ratio (PDR), end-to-end delay, and throughput [23].

## 1.2. Approaches to Improving Routing Performance

Based on the characteristics and challenges discussed above, routing studies in FANETs mainly focus on three approaches [1], [2], [19], [23]: (i) improving topology-based protocols, (ii) improving position-based protocols, and (iii) applying reinforcement learning to routing. Each approach has its own advantages, but limitations remain when applied to FANET environments with rapidly changing topology and high requirements for UAV-to-GCS route stability.

### 1.2.1. Topology-Based Protocol Approach

Topology-based routing inherits MANET protocols such as OLSR, AODV, and DSDV [6], [8], [25], which construct and maintain routing tables based on link-state information. When applied to FANETs, many studies have proposed improved protocols to better reflect link quality and link stability.

Some studies integrate mobility information to estimate link lifetime and combine it with metrics such as ETX, thereby prioritizing links that are both high-quality and more stable [32], [33]. Other studies exploit physical-layer

and MAC-layer information such as RSSI, SNR, and queue load to construct multi-factor routing criteria, thereby selecting less congested and more reliable routes [26], [29]. There are also solutions that focus on energy optimization along the route, balancing residual energy and avoiding nodes at risk of battery depletion [22]. In [CT5], an OLSR-Vr variant was proposed, in which the validity timers for link and topology information are dynamically adjusted according to UAV velocity ranges to improve information freshness, thereby enhancing the routing performance of OLSR in several FANET scenarios.

In general, topology-based improvements have contributed to enhancing route stability and communication quality under certain conditions. However, they still depend on routing-table maintenance mechanisms and therefore become less effective when the topology changes too rapidly or when the network is fragmented. Existing solutions mainly focus on improving local links but have not fully addressed the maintenance of end-to-end routes in scenarios with fast UAV mobility, sparse node density, or large operating areas.

### **1.2.2. Position-Based Protocol Approach**

Position-based routing is considered more suitable for FANETs because of its local decision-making mechanism and low control overhead [23]. Each node uses its own position, the positions of its neighbors, and the position of the destination to select the forwarding node according to local criteria, which are usually based on distance to the destination, movement direction, or other auxiliary factors. GPSR [16] is a representative protocol with two

main modes: greedy forwarding and perimeter forwarding when encountering a routing void. Based on GPSR, several improved protocols for FANETs have been proposed, such as GPMOR [20], GPSR-PPU [27], SF-GeoR [17], UF-GPSR [18], and GPSR+ [14]. These improved protocols incorporate additional neighbor-evaluation criteria, such as predicted position, link lifetime, residual energy, velocity, and flight direction, to prioritize nodes that both advance packets toward the destination and maintain stable links with sufficient energy.

Position-based improvements help reduce delay and increase PDR in many FANET scenarios. However, most of them still mainly exploit spatial and mobility characteristics, while link-quality evaluation is often based only on distance or instantaneous indicators such as RSSI and SNR. The impact of bit rate, packet error rate, and retransmission mechanisms on end-to-end performance has not been sufficiently exploited. In addition, the perimeter mode of GPSR is less suitable for highly mobile three-dimensional environments. Existing solutions are also not yet specialized for the UAV-to-GCS route, where traffic converges toward a fixed destination.

### 1.2.3. Reinforcement-Learning-Based Approach

Recently, many studies have proposed applying reinforcement learning to FANET routing to improve the ability to adapt to network environments [1], [28]. The routing model is often constructed similarly to position-based routing, but the forwarding-node selection decision is learned through interaction between the agent and the environment.

Among reinforcement learning algorithms, Q-learning

[30] is widely used. Each node maintains a Q-table, in which each Q-value represents the estimated quality of selecting a neighbor as the forwarding node. The Q-value is updated based on the reward obtained after packet transmission, reflecting objectives such as reducing delay, increasing PDR, or saving energy. Protocols such as QGeo [15], QMR [21], Q-FANET [7], and QTAR [3] have exploited this mechanism through multi-objective reward-function design, expanded observation spaces, and flexible adjustment of learning parameters.

Published results show that reinforcement-learning-based routing can improve performance in many FANET scenarios compared with fixed strategies. However, most solutions still rely heavily on information from periodic control packets, so knowledge updates may be delayed when the network structure changes rapidly. Packet retransmission and link errors have not been explicitly modeled in the reward function. Coupling states and rewards with indicators such as PER, queue load, retransmission signals, and channel quality under different bit rates remains insufficiently addressed.

### 1.3. GPSR-RA Protocol with Bit-Rate Adaptation

In addition to the general review, this chapter presents the GPSR-RA protocol [CT4]. GPSR-RA is developed based on GPSR, preserving the principle of position-based routing while incorporating a bit-rate adaptation mechanism according to link quality measured by SNIR. The objective is to exploit the multi-rate capability of IEEE 802.11 to improve PDR and throughput while reducing end-to-end delay.

GPSR-RA adopts a cross-layer design with two main

modules. At the network layer, GPSRvar extends the GPSR neighbor table by storing the SNIR information of the link to each neighbor, obtained from the physical layer through Hello packets. At the MAC layer, MACSnir-RA receives the selected *next-hop*, retrieves the corresponding SNIR value, and decides whether to increase or decrease the bit rate by one level before transmitting the data packet. The SNIR thresholds for different bit-rate levels are determined from the PER-SNIR relationship so that the PER remains approximately within the allowable threshold for each rate.

Simulation results under FANET scenarios with different numbers of UAVs, velocities, and densities show that GPSR-RA achieves higher PDR, lower delay, and better throughput than GPSR in many surveyed cases. When channel quality degrades, timely bit-rate reduction helps limit packet errors and retransmissions. This result clarifies the effectiveness of coordinating position-based routing with bit-rate adaptation according to link quality.

## **Chapter 2 POSITION-BASED GPSR-CB ROUTING SOLUTION**

### **2.1. Problem Statement**

Position-based routing allows each node to forward packets hop by hop based on its own position, the positions of its neighbors, and the destination position, without maintaining end-to-end routing tables as in topology-based protocols [16], [23].

GPSR is a representative position-based routing proto-

col, using the *Greedy* and *Perimeter* modes. When applied to FANETs, several variants such as SF-GeoR, UF-GPSR, and GPSR+ have been proposed with additional criteria, including energy, movement tendency, link-breakage risk, and other factors [14], [17], [18]. However, these protocols still rely mainly on mobility-related information and indirect estimations, while making limited use of direct link-quality indicators such as PER/SNIR. They are also not tightly coupled with a multi-level backbone UAV network architecture for the UAV-GCS route. Mechanisms for supporting retransmission and exploiting cross-layer information from the MAC/physical layers to the network layer remain limited.

Based on these observations, Chapter 2 develops GPSR-CB, a GPSR variant designed for the UAV-GCS route, which combines: (i) network organization based on a multi-level backbone UAV architecture, (ii) a multi-criteria evaluation function for the *Greedy* mode, and (iii) cross-layer mechanisms to support neighbor-information updates and packet retransmission.

## 2.2. Analysis of Some Improved GPSR-Based Routing Protocols

SF-GeoR [17] improves GPSR by using a link-stability metric that combines residual energy and the convergence/divergence tendency among UAVs, thereby prioritizing neighbors with more stable links. UF-GPSR [18] uses a multi-criteria utility function, including energy, closeness to the destination, neighbor direction, link-loss risk, and average speed, to select the *next-hop*. These two protocols improve performance compared with the original

GPSR in several FANET scenarios. However, their criteria still mainly reflect mobility and position characteristics, while link quality is evaluated indirectly through distance or safety regions and is not directly associated with PER/SNIR or feedback from the MAC layer.

### 2.3. Improving Performance by Organizing FANETs with Multi-Level Backbone UAVs

In the basic FANET architecture (Figure 2.2a), one backbone UAV ( $BU_1$ ) is selected to connect directly to the GCS; mission UAVs transmit data to the GCS through  $BU_1$ . When the deployment area is large, UAV velocity is high, or node density is sparse, the UAV-GCS route often passes through multiple intermediate nodes and is prone to disruption due to rapid mobility or network fragmentation, leading to reduced PDR and increased delay.

This thesis proposes extending the basic architecture into a *multi-level backbone UAV FANET architecture* (Figure 2.2b) to overcome this limitation. Accordingly, instead of using only one BU, a sequence of backbone UAVs  $BU_1, BU_2, \dots, BU_n$  is arranged along the direction from the GCS to the mission area, forming the backbone axis GCS- $BU_1$ - $\dots$ - $BU_n$ . This deployment mechanism is implemented using a mobility model referenced to anchor points determined from the GCS position and the mission area.

The number of BUs along the backbone axis is estimated as:

$$n = \left\lceil \frac{M_x}{0.9 R_b} \right\rceil, \quad (2.6)$$

where  $M_x$  is the size of the area along the  $X$  axis,  $R_b$  is the communication range of a BU, and the factor  $0.9 R_b$  is used

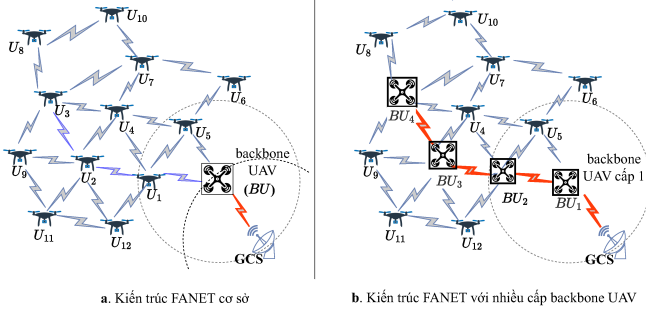


Figure 2.2: Basic FANET and FANET with multi-level backbone UAVs

to ensure link stability according to [31]. The BUs are assumed to be dedicated UAVs with higher energy availability and stability, moving with reference to the GCS and forming a reliable “corridor” for the UAV-GCS route. In the simulations, the 1-BU configuration, representing the basic architecture, and the 4-BU configuration, representing the multi-level backbone architecture, are used to evaluate GPSR-CB and to distinguish the contribution of network organization from that of the routing mechanism.

## 2.4. Proposed GPSR-CB Routing Protocol

### 2.4.1. Design Rationale

GPSR-CB is proposed as an improvement of GPSR for the UAV-GCS route in FANETs, with two main objectives: (i) to guide packets to converge early toward the stable backbone UAV axis before reaching the GCS; and (ii) to select the *next-hop* based on a multi-criteria evaluation function that combines position information, link quality, and the ability to maintain the *Greedy* mode. The protocol also integrates cross-layer mechanisms to rapidly update neighbor

information and support early successful packet retransmission.

#### 2.4.2. Method for Selecting a Neighbor as the *Next-Hop*

GPSR-CB evaluates each neighbor  $B$  of the current node  $A$  with respect to a “virtual destination”  $D$  (the nearest BU or the GCS) based on three criteria: closeness to the virtual destination, the ability to continue the *Greedy* mode, and link quality.

##### \* *Closeness to the “virtual” destination*

The progress toward the virtual destination when forwarding a packet from  $A$  to  $B$  is quantified as:

$$Prx_{B,D} = \frac{L_{A,D} - L_{B,D}}{txRange}, \quad (2.10)$$

where  $L_{A,D}$  and  $L_{B,D}$  are the Euclidean distances from  $A$  and  $B$  to the virtual destination  $D$ , respectively, and  $txRange$  is the maximum communication range of a UAV. A larger value of  $Prx_{B,D}$  indicates that  $B$  helps the packet move closer to the backbone/GCS.

##### \* *Packet error rate*

The PER between a node  $A$  and its neighbor  $B$  can be estimated from the BER using [10]:

$$Per_{A,B} \approx 1 - (1 - BER)^L, \quad (2.11)$$

where  $L$  is the packet size in bits.

\* *Ability to continue the Greedy mode*

The ability of neighbor  $B$  to continue the *Greedy* mode is denoted by  $Abi_B$ :

$$Abi_B = \frac{|Neighbors\_In\_GR(B, D)|}{|Neighbors\_Of(B)|}, \quad (2.12)$$

where  $Neighbors\_In\_GR(B, D)$  is the set of neighbors of  $B$  located in the *Greedy* region with respect to destination  $D$ , and  $Neighbors\_Of(B)$  is the set of all neighbors of  $B$ . A larger value of  $Abi_B$  indicates a higher ability to maintain *Greedy* forwarding from  $B$ , thereby reducing the risk of switching to the *Perimeter* mode.

\* *Neighbor evaluation function*

The link quality between  $A$  and  $B$  is reflected by the packet error rate  $Per_{A,B}$ , which is estimated from BER/SNIR and packet size. The three criteria are combined into the following function:

$$f(B, D)_A = \omega_1 \cdot Prx_{B,D} + \omega_2 \cdot Abi_B + \omega_3 \cdot (1 - Per_{A,B}), \quad (2.13)$$

where  $\omega_1 + \omega_2 + \omega_3 = 1$ . Node  $A$  removes neighbors marked as unreliable, determines the virtual destination  $D$  (the nearest BU or the GCS), computes  $f(B, D)_A$  for each neighbor  $B$ , and selects the node with the highest value as the *next-hop*. If  $A$  belongs to the backbone, the *next-hop* is prioritized as a lower-level BU or directly the GCS, ensuring that packets are forwarded along the backbone axis toward the destination.

\* *Neighbor information management*

GPSR-CB extends the Neighbor Information Table (NIT) to store the position,  $Abi_B$ , and  $Per_{A,B}$  for each neighbor  $B$ . Position information and  $Abi$  are advertised through periodic Hello packets. The metric  $Per_{A,B}$  is updated in a cross-layer manner when packets are received from the lower layer.

### 2.4.3. Supporting Mechanisms

To reduce the update delay of  $Abi_B$ , GPSR-CB exploits the IEEE 802.11 ACK frame: the Duration field, which is not used by default for access control in this context, is reused to carry the latest value of  $Abi_B$ . Whenever a data packet is successfully transmitted, the sender extracts  $Abi_B$  from the ACK and updates the NIT, allowing the function  $f(B, D)_A$  to rely on fresher information than when using only Hello packets.

For retransmissions, GPSR-CB applies two cross-layer operations: (i) temporarily increasing the transmit power during retransmission attempts by a small factor, bounded by the maximum transmit power, in order to improve SNIR and increase the probability of successful retransmission; and (ii) if a *next-hop* exceeds the retransmission threshold without returning an ACK, the corresponding link is temporarily marked by assigning  $Abi = -1$  in the NIT so that it will not be selected in subsequent routing decisions. When this neighbor reappears in Hello packets, its  $Abi$  value is updated according to the new state.

## 2.5. Simulation and Result Analysis

### 2.5.1. Simulation Parameters and Scenarios

GPSR-CB is implemented on OMNeT++ 6.0/INET 4 and compared with GPSR and UF-GPSR in terms of packet delivery ratio (PDR), throughput, end-to-end delay (EED), and energy consumption. Two variants are evaluated: GPSR-CB-0, without transmit-power adjustment, and GPSR-CB-1, with retransmission power adjustment, hereafter referred to as GPSR-CB. The FANET is deployed in a  $1000 \times 1000 \times 1000$  m space, with 50–100 UAVs moving according to the 3D RandomWaypoint model and a communication range of 250 m. The scenarios consider two architectures, 1-BU and 4-BU, and three velocity ranges: [5–10], [15–20], and [25–30] m/s.

### 2.5.2. Results and Analysis

In the 1-BU architecture, as UAV density increases, the PDR of all protocols increases; however, GPSR-CB-0 and GPSR-CB-1 consistently achieve higher PDR and throughput than GPSR and UF-GPSR at all density levels. By combining position information with PER and Ability, GPSR-CB prioritizes links with better quality and neighbors with a higher ability to continue *Greedy* forwarding, thereby reducing the number of transitions to the *Perimeter* mode and avoiding ineffective retransmissions. The end-to-end delay of GPSR-CB remains low and more stable. The average energy consumption is lower than or comparable to that of UF-GPSR and GPSR, due to the reduction in retransmissions and route length.

When moving to the 4-BU architecture, the stable backbone axis improves the performance of all protocols. How-

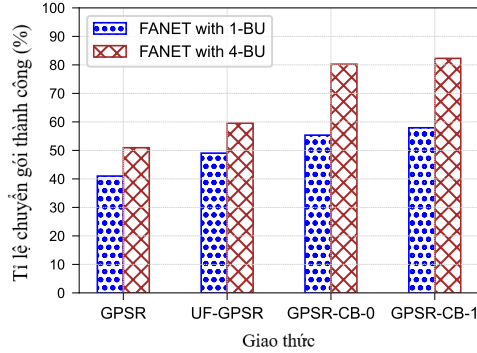


Figure 2.13. Packet delivery ratio in FANETs with 1-BU and 4-BU

ever, the improvement of GPSR-CB is more pronounced because the protocol exploits the virtual-destination criterion and the route-prioritization mechanism through BUs. Figure 2.13 illustrates the difference in PDR between the 1-BU and 4-BU configurations.

The results in Figure 2.13 show that the FANET configuration with multi-level backbone UAVs significantly improves PDR compared with the basic 1-BU architecture. GPSR-CB achieves the largest improvement, with the average PDR increasing from approximately 55–58% to 80–82%, corresponding to an increase of about 24%. In contrast, GPSR and UF-GPSR increase by only about 10%. This indicates that GPSR-CB effectively exploits the multi-level backbone axis to maintain a more stable U2G route.

In the velocity-varying scenarios, the PDR of all protocols decreases as UAV velocity increases because the links become less stable. Nevertheless, GPSR-CB still maintains better results than GPSR and UF-GPSR. Overall, the results

show that GPSR-CB improves PDR, throughput, and EED in both the 1-BU and 4-BU architectures, while its processing complexity remains linear with the number of neighbors. However, the protocol uses fixed weights, and therefore its ability to adapt to network conditions remains limited.

## **Chapter 3 QLR-FANET ROUTING SOLUTION USING REINFORCEMENT LEARNING**

### **3.1. Introduction**

Based on the modeling of the multi-hop routing problem using Q-learning, where the agent is a packet, the state is the current node holding the packet, and the action is the selection of a *next-hop*, this chapter first analyzes several representative protocols, including QGeo, QMR, and Q-FANET. From this analysis, common limitations in applying reinforcement learning to environments with rapidly varying links are identified. The chapter then presents the design of the QLR-FANET protocol, including a two-component reward function, an event-driven knowledge propagation mechanism through D-ACK packets, and a cross-layer bit-rate adaptation mechanism. Finally, simulation-based performance evaluation is presented.

### **3.2. Analysis of Some Reinforcement-Learning-Based Routing Protocols**

Several FANET routing protocols have applied Q-learning on top of position-based routing. QGeo [15] uses a reward based on the packet progress speed toward the destination. Each node shares its position, velocity, error informa-

tion, and  $maxQ_Y^D$  through Hello packets. The discount factor  $\gamma$  is set according to the predicted maintenance or breakage of the link, and the *next-hop* is selected using an  $\epsilon$ -greedy strategy. QMR [21] extends QGeo toward a multi-objective design: the reward function combines one-hop delay and residual energy, the learning rate  $\alpha$  is adjusted according to the delay variation on each link, the discount factor  $\gamma_Y$  reflects the stability of the neighbor set, and *next-hop* selection is performed through two phases with a progress-velocity constraint under a *deadline*. Q-FANET [7] applies the QNoise+ algorithm [9], in which Q-values are updated from a weighted reward of several most recent steps combined with an SINR-dependent component. Its reward function is designed with three levels, corresponding to reaching the destination, encountering a routing void, and the remaining cases, while the *next-hop* is still selected using  $\epsilon$ -greedy over the candidate set satisfying the progress-velocity constraint.

These protocols show that the Q-learning approach has the potential to improve PDR, EED, and energy consumption in several FANET scenarios. However, they also share several limitations. First, learned knowledge, especially  $maxQ_Y^D$ , is mainly propagated through periodic Hello packets, and therefore may become outdated or lost when links change rapidly or when the channel is strongly affected by noise. Second, the reward functions mainly focus on local one-hop metrics, such as progress speed, delay, energy, and queue occupancy, without clearly separating the contribution of the local state from the ability of the *next-hop* to receive and continue forwarding packets toward the destination. Third, packet retransmission is mainly

handled at the MAC layer, while retransmission events, or Retry events, have not been exploited as learning signals. In addition, cross-layer bit-rate adaptation has not been sufficiently used to increase the probability of successful retransmission without increasing transmit power.

These limitations form the basis for proposing the QLR-FANET protocol, in which the reward function is designed with two components, namely the local reward and the action reward from the *next-hop*. Learned knowledge is shared in an event-driven manner through D-ACK packets associated with data packets, and the cross-layer bit-rate adaptation mechanism is triggered when packet retransmissions occur.

### 3.3. Proposed QLR-FANET Routing Protocol

#### 3.3.1. Design Rationale

QLR-FANET is developed to address the three limitations discussed above. First, the RL model uses a two-component reward function, separating the local reward, which is evaluated by the sender from its local observation, from the action reward, which is fed back by the *next-hop* and associated with its ability to forward packets toward the destination. This separation allows the agent not only to optimize the current forwarding step, but also to prioritize *next-hop* nodes that have the potential to maintain stable routes in subsequent hops.

Second, instead of propagating learned knowledge only through Hello packets, QLR-FANET introduces an event-driven feedback channel through D-ACK packets directly associated with data packets. Whenever a packet is successfully forwarded, the *next-hop* sends a D-ACK containing

the *ActionReward* and the current max  $Q$  value toward the destination, enabling the sender to update its Q-value more closely according to the current link state.

Third, QLR-FANET integrates a cross-layer mechanism that uses Retry signals from the MAC layer. Failed retransmission attempts are mapped to negative rewards, while also triggering bit-rate adaptation according to the IEEE 802.11 multi-rate mechanism, temporarily reducing the bit rate instead of increasing transmit power. This both increases the probability of successful retransmission and limits interference, making the mechanism suitable for UAVs with hardware constraints.

### 3.3.2. Routing Model and Learning Data at Each Node

QLR-FANET models the packet forwarding process as an RL problem, in which the state corresponds to the current node and the action is to select one neighbor as the *next-hop*. Each node maintains two main data structures. The first is the *Observation* table, which is collected from Hello packets and link measurements and is used for reward calculation. The second is the *Q-table*, which stores Q-values over the neighbor set to support distributed decision-making. Q-value updates are triggered by events associated with data transmission, through D-ACK packets or retransmission signals.

QLR-FANET uses a two-layer reward consisting of *LocalReward*  $\mathcal{L}_A(B)$  and *ActionReward*  $\mathcal{N}(B)$  for each neighbor  $B$  of the current node  $A$  with respect to destination  $D$ .

The two reward components are combined using the co-

efficient  $\Lambda$ :

$$\mathcal{R}_A(B) = \Lambda \cdot \mathcal{L}_A(B) + (1 - \Lambda) \cdot \mathcal{N}(B), \quad (3.21)$$

where  $\mathcal{R}_A(B)$  is the reward used to update the Q-value according to the Q-learning rule.

*LocalReward* reflects the local observation at the current node when considering a candidate neighbor, and consists of four components: the packet error rate on the link ( $Per_{\text{link}(A,B)}$ ), the relative velocity ( $v_{B,A}$ ), the progress toward the destination ( $d_{(B,A) \rightarrow D}$ ), and the queue occupancy at the *next-hop* ( $Load_B$ ). These four components are normalized to the range  $[0, 1]$  and linearly combined using the weight vector  $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ :

$$\begin{aligned} \mathcal{L}_A(B) = & \lambda_1 \cdot (1 - Per_{\text{link}(A,B)}) + \lambda_2 \cdot (1 - v_{B,A}) \\ & + \lambda_3 \cdot d_{(B,A) \rightarrow D} + \lambda_4 \cdot (1 - Load_B), \end{aligned} \quad (3.22)$$

*ActionReward* reflects direct feedback from the *next-hop* regarding its ability to continue forwarding packets toward the destination, based on the metric  $\mathcal{P}_B(D)$  and the reverse-link quality  $Per_{\text{link}(B,A)}$ . Specifically:

$$\mathcal{N}(B) = \begin{cases} -1, & \text{if } \mathcal{P}_B(D) = 0, \\ 1, & \text{if } \mathcal{P}_B(D) > \mathcal{P}_\theta, \\ k - Per_{\text{link}(B,A)}, & \text{otherwise,} \end{cases} \quad (3.23)$$

where  $k$  is an adjustment constant and  $\mathcal{P}_\theta \in (0, 1)$  is the forwarding-capability threshold.

### 3.3.3. Next-Hop Selection

The selection of the *next-hop* in QLR-FANET is performed by combining the Q-value with spatial information. At each node  $A$ , the neighbors belonging to the three most recent nodes traversed by the packet and the neighbors with  $\mathcal{P}_B(D) = 0$  are removed from the candidate set to avoid loops and eliminate blocked forwarding directions early. For each remaining candidate, the Q-value  $Q_A(B, D)$  is multiplied by the progress-to-destination metric  $d_{(B,A) \rightarrow D}$  to obtain a weighted Q-value, from which the action-selection probability distribution is constructed using the Softmax policy. The *next-hop* is selected as the node with the highest probability in this distribution. This approach enables the protocol to prioritize nodes that both have high learned values and make progress toward the destination, while still maintaining a directed exploration component in an environment with fluctuating links.

### 3.3.4. Supporting Mechanisms

QLR-FANET introduces two supporting mechanisms associated with retransmission signals from the MAC layer. First, when a packet must be retransmitted, the agent records this as a suboptimal decision and assigns a fixed negative penalty to the link causing the retransmission. This gradually reduces the probability of selecting that link again in subsequent routing decisions without immediately removing neighbors that are still available. Second, the protocol exploits the multi-rate capability of IEEE 802.11 to adapt the bit rate: if no ACK is received after the first transmission, the sender reduces the bit rate to the next lower level. After a successful retransmission, the bit rate is restored to

the default value.

### 3.4. Simulation and Performance Evaluation

The performance of QLR-FANET is evaluated through simulation on a FANET model with a size of  $1500\text{ m} \times 1500\text{ m} \times 120\text{ m}$ , 25 to 45 UAVs, a mobility speed of 10–40 m/s, and the MassMobility model. The wireless links use IEEE 802.11g multi-rate with a transmit power of 10 dBm, the Thermal Noise model, and an SINR threshold of 4 dB. The Hello transmission frequency is 1 Hz. The reinforcement learning parameters, including the learning rate, discount factor, and reward-function weights, are kept fixed.

QLR-FANET is compared with GPSR, AODV, and xQ-Geo. GPSR represents the group of position-based routing protocols, AODV represents the group of topology-based routing protocols, while xQGeo is used as a comparison protocol in the direction of Q-learning-based routing. This selection enables QLR-FANET to be evaluated against both traditional routing protocols and a protocol with a learning mechanism that is closer in nature.

The velocity-based evaluation results show that QLR-FANET maintains a higher PDR than GPSR, AODV, and xQ-Geo over the entire velocity range of 10–40 m/s. Figure 3.9 illustrates the PDR results of the protocols as UAV velocity varies.

The results in Figure 3.9 show that at a velocity of 30 m/s with 25 nodes, the average PDR values of GPSR, AODV, xQGeo, and QLR-FANET are 22.8%, 44.6%, 45.3%, and 68.2%, respectively. This result reflects the adaptability of QLR-FANET in a FANET environment with rapidly

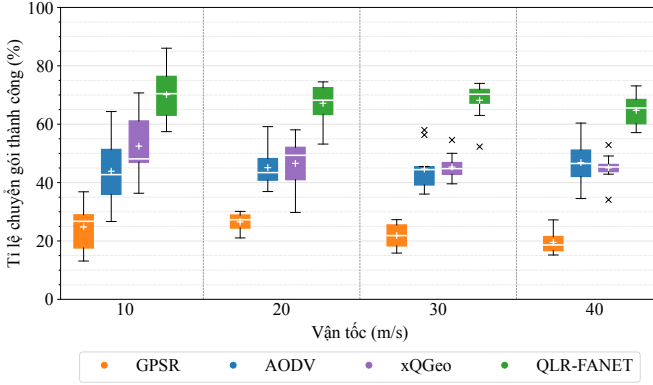


Figure 3.9. Packet delivery ratio with respect to velocity

changing topology. While GPSR and AODV are affected by outdated routing information or route re-establishment processes, xQGeo, although using Q-learning, does not fully exploit data transmission feedback from the link layer. QLR-FANET improves PDR by combining a two-component reward function, D-ACK feedback, a penalty mechanism for poor links, and bit-rate adjustment during retransmission.

In the node-density evaluation with 25, 35, and 45 UAVs at a velocity of 30 m/s, the PDR and throughput of QLR-FANET increase with node density and remain higher than those of the comparison protocols. With 45 nodes, the PDR values of GPSR, AODV, xQGeo, and QLR-FANET are approximately 31%, 60%, 67%, and 93%, respectively. The corresponding average throughputs are approximately 70, 135, 155, and 211 Kb/s. The average number of broken links across the network under QLR-FANET is also significantly lower, reflecting its ability to select more stable routes in an environment with rapidly varying links.

### **3.5. Comparison between QLR-FANET and GPSR-CB**

To directly compare the two proposed protocols, the thesis reuses the simulation scenario of GPSR-CB under the basic FANET architecture, namely the 1-BU configuration.

In the scenario evaluating the effect of UAV density within the velocity range of 5–10 m/s, with three density levels of 50, 75, and 100 UAVs, the PDR of both protocols tends to increase as UAV density increases, since the number of available neighbors and the capability to maintain multi-hop connectivity are improved. At most surveyed density levels, QLR-FANET achieves a higher average PDR than GPSR-CB. Specifically, at the density of 75 UAVs, QLR-FANET achieves approximately 63%, about 5% higher than GPSR-CB.

At high velocity ranges, the PDR of both GPSR-CB and QLR-FANET decreases because links change rapidly and the probability of route breakage increases. The PDR gap between the two protocols tends to narrow, indicating that the adaptive learning advantage of QLR-FANET gradually decreases when link states fluctuate too rapidly. Meanwhile, GPSR-CB still maintains an advantage in lower end-to-end delay and can achieve better effectiveness when exploiting a multi-level backbone UAV architecture. Therefore, the two protocols represent two complementary approaches: GPSR-CB is suitable for U2G routing with a clearly organized backbone, while QLR-FANET is suitable for scenarios that require adaptive learning capability from

data transmission feedback.

## CONCLUSION

Improving routing performance on the U2G route in FANETs is necessary and has practical significance for applications such as surveillance, search and rescue, and security–defense. Through the research process, this thesis has proposed and validated several new routing solutions, while also identifying a number of directions for future development.

### a) Achieved Results

The thesis focuses on adaptive routing mechanisms for FANETs along two main directions: position-based routing and reinforcement-learning-based routing, both associated with the specific characteristic of data flows converging toward the GCS. The achieved results have fulfilled the initially defined objectives. Specifically, the thesis has:

- (1) Proposed the GPSR-CB routing protocol, an improved GPSR variant for the U2G route, designed to exploit the multi-level backbone UAV FANET architecture in order to better utilize stable links along the U2G route. Specifically, GPSR-CB improves neighbor evaluation in the *Greedy* mode: the forwarding node is selected using a multi-criteria function that combines closeness to the “virtual destination” (the nearest BU or the GCS), the *Ability* metric, and link quality quantified by PER. GPSR-CB applies a retransmission optimization mechanism by increasing the transmit power whenever

a packet needs to be retransmitted, thereby improving the probability of successful transmission.

- (2) Proposed the QLR-FANET routing protocol, which applies Q-learning to the *next-hop* selector. In this protocol, a packet is regarded as an agent that learns from data transmission feedback, and the reward function is designed by combining a local reward and an action reward from the *next-hop*. The max  $Q$  information is shared through D-ACK packets and is associated with a retransmission penalty mechanism and bit-rate adjustment based on lower-layer signals. This design enables the routing policy to gradually adapt to network conditions and improve communication performance.
- (3) Proposed a multi-level backbone UAV FANET architecture, in which backbone UAVs are hierarchically organized along the direction toward the GCS, forming a more stable forwarding axis. Simulation results show that the multi-level backbone UAV architecture improves network performance.
- (4) Proposed the GPSR-RA protocol, an improved GPSR variant following a cross-layer approach with a bit-rate adjustment mechanism based on SNIR thresholds. By reducing the bit rate when channel quality degrades and increasing it when the channel improves before each transmission, GPSR-RA contributes to improving transmission reliability and reducing end-to-end delay.

In summary, the thesis has achieved its research objective of proposing and validating new routing solutions to improve FANET performance on the U2G route. These re-

sults contribute to strengthening the scientific foundation for U2G-oriented FANET routing and demonstrate the benefits of combining purposeful network organization, link-quality evaluation based on PER and bit rate, and distributed adaptive learning mechanisms.

Nevertheless, the above results and conclusions are established for the single-destination routing problem on the U2G route in FANETs, and do not separately consider scenarios with specific requirements for routing security, multi-destination support, or differentiated service classes.

## **b) Future Research Directions**

Based on the achieved results, the thesis identifies several future research directions as follows:

- (1) In GPSR-CB and QLR-FANET, the weights in the evaluation function and reward function are currently selected as fixed parameters. Developing mechanisms to automatically adjust these parameters is one of the future research directions that should be considered.
- (2) Extending the applicability of the proposed routing mechanisms to contexts with additional requirements for routing security, multi-destination support, or differentiated service classes. Integrating these requirements with the multi-level backbone UAV FANET architecture is a potential research direction toward developing solutions more suitable for complex FANET applications.

**LIST OF THE AUTHOR'S SCIENTIFIC PUBLICATIONS**

- [CT1]Mai Cuong Tho**, Le Huu Binh, Tu T. Vo (2025). "GPSR-CB: A novel routing algorithm for FANET using cross-layer models in combination with multi-level backbone UAV". *Ad Hoc Networks*, 173, 103828.
- [CT2]Mai Cuong Tho**, Ly T. H. Nguyen, Le Huu Binh, Tu T. Vo (2025). "QLR-FANET: A Q-learning and rate control-based routing protocol for flying ad hoc network". *ETRI Journal*.
- [CT3]Tho C. Mai**, Ly T. H. Nguyen, Cuong Q. Nguyen, Le Huu Binh, Tu T. Vo (2024). "An Efficient and Low-Cost Hierarchical FANET Architecture Using the Attached Mobility Model to Create the Core Layer for Connecting UAVs and Ground Control Stations". *Journal of Advances in Information Technology*, 15(4), 500–510.
- [CT4]Mai Cuong Tho**, Le Huu Binh, Vo Thanh Tu (2025). "GPSR-RA: A Cross-Layer Position-Based Routing Protocol with Dynamic Bitrate Adaptation for FANETs". *Hue University Journal of Science*, Vol. 134, No. 2A(2025).

**[CT5]Mai Cuong Tho**, Nguyen Thi Huong Ly, Vo Thanh Tu (2023). "Changing holding time parameters in the OLSR protocol for FANET performance enhancement". *Journal of Science and Technology, University of Sciences, Hue University*, Vol. 22(1B), pp. 105–116.

**[CT6]Mai Cuong Tho**, Nguyen Thi Huong Ly, Nguyen Quoc Cuong, Le Huu Binh, Vo Thanh Tu (2023). "Improving FANET performance in search and rescue applications: mobility model parameters and implementation strategy". *Proceedings of FAIR2023*.