

định tuyến của T , ký hiệu là $C(T)$, là tổng chi phí định tuyến giữa mọi cặp đỉnh thuộc cây T , tức là:

$$C(T) = \sum_{u,v \in V(T)} d_T(u,v). \quad (1)$$

Bài toán đặt ra là trong số tất cả các cây khung của đồ thị G cần tìm cây khung có chi phí định tuyến nhỏ nhất. Bài toán này được đặt tên là bài toán cây khung với chi phí định tuyến nhỏ nhất (Minimum Routing Cost Spanning Tree-MRCST). Bài toán MRCST đã được chứng minh thuộc lớp bài toán NP-hard [4].

Trên cơ sở khái niệm tải định tuyến (routing load), công trình [4] đã chỉ ra cách tính chi phí định tuyến của một cây khung với độ phức tạp tính bằng cách sử dụng cấu trúc cây có gốc.

Định nghĩa 3 (Tải định tuyến một cạnh của cây khung [4]) Cho T là một cây khung của đồ thị G . Nếu loại khỏi cây T một cạnh e thì T sẽ được tách ra thành hai cây con T_1 và T_2 với hai tập đỉnh tương ứng ký hiệu là $V(T_1)$ và $V(T_2)$. Ta gọi tải định tuyến của cạnh e , ký hiệu là $l(T,e)$, là giá trị $2|V(T_1)| \times |V(T_2)|$.

Từ định nghĩa, dễ thấy tải định tuyến của cạnh e chính là bằng số lượng đường đi trên cây T có chứa cạnh e .

Định lý sau cho ta cách tính chi phí định tuyến của cây khung thông qua tải định tuyến của các cạnh.

Định lý. Cho T là một cây khung của G , ta có:

$$C(T) = \sum_{e \in E(T)} l(T,e) \times w(e) \quad (2)$$

và chi phí định tuyến của cây khung T có thể tính được trong thời gian $O(n)$ [4].

Có thể tìm thấy các ứng dụng của bài toán MRCST trong lĩnh vực thiết kế mạng truyền thông và trong tin sinh học [4,7,11,12].

1.2 Các nghiên cứu liên quan

Hiện đã có hơn hai mươi công trình giải bài toán MRCST được đề xuất theo các hướng tiếp cận giải đúng, giải gần đúng cận tỉ lệ, heuristic và metaheuristic.

Các thuật toán giải đúng phát triển dựa trên phương pháp nhánh cận kết hợp phương pháp sinh cột được đề xuất bởi Matteo Fischetti, Giuseppe Lancia, Paolo Serafini vào năm 2002 [6]; tiếp cận hướng này chỉ có thể giải được các bài toán MRCST có không quá 30 đỉnh và do đó tính ứng dụng thực tiễn của nó không cao.

Một thuật toán gần đúng cận tỉ lệ α nghĩa là trong trường hợp xấu nhất thì lời giải tìm được của thuật toán đó cũng không tệ hơn α lần so với lời giải tối ưu; tuy nhiên các cận tỉ lệ tìm được của nhiều thuật toán đề xuất trong thực tế thường là kém hơn rất nhiều so với chất lượng lời giải tìm được bởi nhiều thuật toán gần đúng khác dựa trên thực nghiệm [3,4] – điển hình theo hướng này là thuật toán WONG với cận tỉ lệ 2 và có độ phức tạp thời gian tính $O(nm + n^2 \log n)$ [4]; đây là thuật toán định tuyến mạng chạy ổn định trên mọi loại dữ liệu và đã được sử dụng trong thực tế ứng dụng [4,11].

Các thuật toán heuristic có chất lượng lời giải chấp nhận được với một loại dữ liệu cụ thể nào đó của bài toán; chẳng hạn như thuật toán ADD được đề xuất bởi Vic Grout vào năm 2005 [12] với độ phức tạp $O(n \log n)$ và chỉ tương đối hiệu quả với các đồ thị đồng nhất có kích thước nhỏ [9], thuật toán CAMPOS được đề xuất bởi Rui Campos và Manuel Ricardo vào năm 2008 [11] với độ phức tạp $O(m + n \log n)$ và chỉ hiệu quả đối với các đồ thị đầy đủ euclid [9]; cả hai heuristic này đều không hiệu quả đối với các đồ thị thưa về chất lượng lời giải [9]; tuy nhiên đây là hai thuật toán nhanh nhất hiện nay trong việc giải bài toán MRCST; điều này có ý nghĩa quan trọng; đặc biệt là trong các mạng viễn thông di động, nơi mà cấu trúc của các mạng thường xuyên thay đổi và cần phải có một cơ chế hiệu quả để tìm MRCST mới trong thời gian nhanh nhất mặc dù có thể chấp nhận một mức chi phí nhiều hơn.

Các thuật toán metaheuristic giải bài toán MRCST đã được đề xuất như các thuật toán di truyền ESCGA [5], BCGA [5], các thuật toán tìm kiếm địa phương SHC [5], PBLS [1], REPRI [8], REPIR [8], TABU [10], các thuật toán ong PABC [2], ABC+LS [2], BEE [9],... Hiện tại, các tiếp cận giải bài toán MRCST theo hướng metaheuristic là hiệu quả nhất [9]; tuy nhiên các thuật toán trên chỉ hiệu quả đối với loại đồ thị đầy đủ euclid và đồ thị ngẫu nhiên có kích thước nhỏ (tất cả các thực nghiệm đã công bố hiện nay với đồ thị đầy đủ đều không quá 300 đỉnh); các metaheuristic gần đây nhất như PABC, ABC+LS, BEE không hiệu quả về thời gian tính và cả chất lượng lời giải đối với các đồ thị thưa – nhất là với các đồ thị thưa có kích thước lớn.

Trong số các thuật toán đã khảo sát trên, chỉ có các thuật toán WONG, SHC, PBLS, REPRI, REPIR là khả thi đối với các đồ thị thưa có kích thước lớn