

GIẢI THUẬT TÌM ĐƯỜNG ĐI NGẮN NHẤT GIỮA HAI TẬP ĐỈNH

ALGORITHMS OF THE PROBLEM OF FINDING THE SHORTEST PATHS FROM A SET OF NODES TO ANOTHER SET OF NODES

TRẦN QUỐC CHIẾN – NGUYỄN THANH TUẤN
Trường Đại học Sư phạm, Đại học Đà Nẵng

TÓM TẮT

Bài toán tìm đường đi ngắn nhất là vấn đề quan trọng trong lý thuyết đồ thị, nó đã được nghiên cứu từ lâu và có nhiều ứng dụng trong nhiều ngành khoa học nói chung và khoa học máy tính nói riêng. Nhiều giải thuật (Dijkstra, Bellman-Ford, Floyd...) đã được phát triển để tìm đường đi ngắn nhất cho một cặp đỉnh hay cho tất cả các cặp đỉnh. Bài viết này nghiên cứu bài toán tìm đường đi ngắn nhất giữa hai tập đỉnh trên đồ thị và đề xuất một giải thuật hiệu quả để giải bài toán này. Giải thuật được cài đặt trong ngôn ngữ C# và cho kết quả thử nghiệm khả quan.

ABSTRACT

The shortest-path problem is an important issue in graph theory. It has been studied for a long time and found diverse applications in various fields. Some algorithms (e.g. Dijkstra, Bellman-Ford, Floyd...) have been designed for solving a single source shortest-path or all pair shortest-path problems. This paper treats the problem of finding shortest paths from a set of nodes to another set of nodes in a graph and presents algorithms for solving this problem. These algorithms have been programmed on the C# language with satisfactory results.

MỞ ĐẦU

Bài toán tìm đường đi ngắn nhất là bài toán quan trọng trong Lý thuyết đồ thị, nó được áp dụng để giải quyết rất nhiều bài toán trong thực tế như điều khiển tối ưu, giao thông vận tải, mạng viễn thông ...

Bài toán này có thể chia làm 2 loại:

Tìm đường đi ngắn nhất giữa một cặp đỉnh: Cho đồ thị $G(V,E)$ có trọng số cạnh và hai đỉnh u, v thuộc V tìm đường đi ngắn nhất từ đỉnh u đến đỉnh v trên đồ thị G . Các giải thuật được phát triển để giải bài toán dạng này tiêu biểu là các giải thuật: Dijkstra, Bellman-Ford,...

Tìm đường đi ngắn nhất giữa tất cả các cặp đỉnh: Cho đồ thị $G(V,E)$ có trọng số cạnh tìm đường đi từ đỉnh u đến đỉnh v , với mọi cặp đỉnh u, v thuộc V . Các giải thuật đã được phát triển để giải bài toán này là: Floyd-Warshall, Johnson,...

Trong thực tế nhiều khi ta không chỉ cần tìm đường đi ngắn nhất giữa hai đỉnh mà còn cần xác định đường đi ngắn nhất giữa một tập đỉnh này đến một tập đỉnh khác. Bài toán đó được phát biểu như sau: Cho đồ thị $G(V,E)$ có trọng số cạnh và hai tập đỉnh $A, B \subset V$ tìm đường đi ngắn nhất từ tập đỉnh A đến tập đỉnh B .

Bài viết này nghiên cứu bài toán tìm đường đi ngắn nhất giữa hai tập đỉnh trên đồ thị và đề xuất một số giải thuật hiệu quả để giải bài toán.

NỘI DUNG

1. Tìm đường đi ngắn nhất giữa một cặp đỉnh

1.1. Định nghĩa

Xét đồ thị có trọng số cạnh $G = (V, E, w)$, với hàm trọng số $w: E \rightarrow \mathbf{R}$ là ánh xạ từ tập các cạnh E đến tập số thực \mathbf{R} .

Định nghĩa 1.1. Đường đi p từ đỉnh u đến đỉnh v là dãy các cạnh nối tiếp nhau bắt đầu từ đỉnh u kết thúc tại đỉnh v . Đường đi p từ u đến v được biểu diễn như sau: $p = (u = v_0, v_1, \dots, v_k = v)$

Định nghĩa 1.2. Độ dài của đường đi $p = (v_0, v_1, \dots, v_k)$, ký hiệu $\omega(p)$, là tổng các trọng số của các cạnh trên đường đi:

$$\omega(p) = \sum_{i=1}^k w(v_{i-1}, v_i)$$

Định nghĩa 1.3. Gọi $\wp(u, v)$ là tập tất cả đường đi từ u đến v . Độ dài đường đi ngắn nhất từ đỉnh u đến đỉnh v được xác định bởi:

$$d(u, v) = \min \{ \omega(p) \mid p \in \wp(u, v) \}$$

Định nghĩa 1.4. Đường đi ngắn nhất $p_{\min}(u, v)$ từ đỉnh u đến đỉnh v là đường đi có độ dài $d(u, v)$.

1.2. Giải thuật Dijkstra

Có rất nhiều giải thuật đã được phát triển để giải bài toán tìm đường đi ngắn nhất giữa một cặp đỉnh, trong khuôn khổ bài viết này chúng tôi chỉ xin giới thiệu giải thuật Dijkstra. Giải thuật Dijkstra là một giải thuật để giải bài toán đường đi ngắn nhất nguồn đơn trên một đồ thị có trọng số cạnh mà tất cả các trọng số đều không âm. Nó xác định đường đi ngắn nhất giữa hai đỉnh cho trước, từ đỉnh a đến đỉnh b .

Ở mỗi đỉnh v , giải thuật Dijkstra xác định 3 thông tin: k_v , d_v và p_v .

k_v : mang giá trị boolean xác định trạng thái được chọn của đỉnh v .

Ban đầu ta khởi tạo tất cả các đỉnh v chưa được chọn, nghĩa là:

$$k_v = \text{false}, \forall v \in V.$$

d_v : là chiều dài đường đi mà ta tìm thấy cho đến thời điểm đang xét từ a đến v .

Khởi tạo, $d_v = \infty, \forall v \in V \setminus \{a\}, d_a = 0$.

p_v : là đỉnh trước của đỉnh v trên đường đi ngắn nhất từ a đến b . Đường đi ngắn nhất từ a đến b có dạng $\{a, \dots, p_v, v, \dots, b\}$. Khởi tạo, $p_v = \text{null}, \forall v \in V$.

Sau đây là các bước của giải thuật Dijkstra:

-
- B1.** Khởi tạo: Đặt $k_v := \text{false} \forall v \in V; d_v := \infty, \forall v \in V \setminus \{a\}, d_a := 0$.
 - B2.** Chọn $v \in V$ sao cho $k_v = \text{false}$ và $d_v = \min \{d_t \mid t \in V, k_t = \text{false}\}$
Nếu $d_v = \infty$ thì *kết thúc*, không tồn tại đường đi từ a đến b .
 - B3.** Đánh dấu đỉnh v , $k_v := \text{true}$.
 - B4.** Nếu $v = b$ thì *kết thúc* và d_b là độ dài đường đi ngắn nhất từ a đến b .
Ngược lại nếu $v \neq b$ sang **B5**.
 - B5.** Với mỗi đỉnh u kề với v mà $k_u = \text{false}$, kiểm tra
Nếu $d_u > d_v + w(v, u)$ thì $d_u := d_v + w(v, u)$
Ghi nhớ đỉnh v : $p_u := v$.
Quay lại **B2**.
-

1.3. Độ phức tạp của giải thuật Dijkstra

a) Trường hợp sử dụng ma trận kề

Gọi $f(n)$ là số lần giải thuật Dijkstra khảo sát một cạnh của đồ thị G trong trường hợp xấu nhất. Khi đó ta có:

$$f(n) < O(V^2)$$

Chứng minh: Cho $n = |V|$, **B5** là vòng lặp chứa các bước **B2** → **B5**, vòng lặp được thực hiện đến khi $v = b$. Vì ở mỗi vòng lặp ta rút ra một đỉnh của V và khởi đầu V có n phần tử, nên vòng lặp được xử lý nhiều nhất là n lần.

Ở **B2** số đỉnh tối đa được khảo sát là $n - 1$ đỉnh

Ở **B5** số đỉnh kề tối đa được khảo sát là $n - 1$ đỉnh

Do đó: $f(n) \leq 2(n-1)n < O(V^2)$

Vậy độ phức tạp của giải thuật Dijkstra là $O(V^2)$.

b) Trường hợp sử dụng danh sách kề

Độ phức tạp của giải thuật Dijkstra là $O((|V| + |E|) \lg |V|)$.

(Xem [2] - trang 530-531. Định lý 25.11)

1.4. Định lý

Giải thuật Dijkstra là đúng.

(Xem [2] - trang 529-530. Định lý 25.10)

2. Giải thuật tìm đường đi ngắn nhất giữa hai tập đỉnh

2.1. Định nghĩa

Xét đồ thị có trọng số cạnh $G=(V,E,w)$ với các tập đỉnh A, B là các tập con khác rỗng của V .

Định nghĩa 2.1 Gọi $\wp(A,B)$ là tập hợp đường đi giữa tất cả các cặp đỉnh của hai tập đỉnh A và B . Đường đi ngắn nhất giữa hai tập đỉnh A và B , ký hiệu $p_{min}(A,B)$, là đường đi có độ dài:

$$\omega(p_{min}(A,B)) = \min \{ \omega(p) \mid p \in \wp(A,B) \}$$

$\omega(p_{min}(A,B))$ gọi là khoảng cách giữa hai tập đỉnh A và B và ký hiệu là $d(A,B)$.

Lưu ý rằng, nếu $A \cap B \neq \emptyset$, thì $d(A,B) = 0$.

2.2. Giải thuật

Cho đồ thị $G(V,E,w)$. Cho

A là tập các đỉnh nguồn.

B là tập các đỉnh đích.

Kí hiệu:

Δ là đường đi ngắn nhất từ A đến B tại thời điểm đang xét.

$d(\Delta)$ là độ dài đường đi ngắn nhất từ A đến B tại thời điểm đang xét.

$\alpha = |A|, \beta = |B|$

Để tìm đường đi ngắn nhất giữa hai tập đỉnh A, B chúng ta có thể thực hiện giải thuật sau:

Giải thuật 1 (Sử dụng định nghĩa)

Đầu vào: $G(V,E)$ và 2 tập đỉnh $A, B \subset V$

Đầu ra: $a \in A, b \in B$, sao cho $d(A,B) = d(a,b)$ nhỏ nhất, và $p_{min}(A,B) = p_{min}(a,b)$

B1. Ta sắp xếp các phần tử của tích Đề-các $A \times B$ theo thứ tự nào đó

$$A \times B = \{(u_1, v_1), (u_2, v_2), \dots, (u_k, v_k)\}$$

Đặt $\Psi := A \times B, \Delta = \emptyset, d(\Delta) = +\infty$

B2. Chọn $(u_i, v_i) \in \Psi$ có chỉ số i nhỏ nhất. Đặt $(u, v) = (u_i, v_i)$

B3. Tìm đường đi ngắn nhất $p_{min}(u, v)$

Nếu $d(u, v) < d(\Delta)$, thì đặt $d(\Delta) = d(u, v)$, $\Delta = p_{min}(u, v)$

$\Psi := \Psi - \{(u, v)\}$.

B4. Nếu $\Psi \neq \emptyset$, thì quay về **B2**

Nếu $\Psi = \emptyset$, thì **kết thúc**, $p_{min}(A, B) = \Delta$, $d(A, B) = d(\Delta)$

Ở giải thuật này, để tìm được đường đi ngắn nhất giữa hai tập đỉnh ta phải tìm đường đi ngắn nhất giữa α, β cặp đỉnh, ở đây $\alpha = |A|$ và $\beta = |B|$. Trong thực tế khi mà đồ thị G có số lượng đỉnh lớn và nếu các tập đỉnh nguồn và tập đỉnh đích có nhiều đỉnh thì độ phức tạp của giải thuật này là $\alpha, \beta \cdot O(|V|^2)$ (Ở đây giải thuật để tìm đường đi ngắn nhất giữa hai cặp đỉnh ở **B3** là giải thuật Dijkstra với độ phức tạp $O(|V|^2)$).

Khái niệm Cây đường đi ngắn nhất

Cây đường đi ngắn nhất $T_{(a)} = (V_{(a)}, E')$ là đồ thị con có hướng của G sao cho:

1. $V_{(a)}$ là tập hợp các đỉnh được đánh dấu trong giải thuật tìm đường đi ngắn nhất từ đỉnh a .

$$V_{(a)} = \{v \in V \mid k_v = true\}$$

2. $T_{(a)}$ là cây có gốc a .

3. $\forall v \in V_{(a)}$, đường đi từ a đến v trong $T_{(a)}$ là đường đi ngắn nhất từ a đến v trong G .

$T_{(a)} = (V_{(a)}, E')$ gọi là *cây đường đi ngắn nhất cực đại* nếu nó là cây đường đi ngắn nhất thoả:

$$V_{(a)} = V$$

hoặc

$$V - V_{(a)} = \{v \in V \mid d(a, v) = \infty\}$$

Với khái niệm cây đường đi ngắn nhất ta có thể đưa ra một giải thuật khác hiệu quả hơn giải thuật 1:

Giải thuật 2

Đầu vào: $G(V, E)$ và 2 tập đỉnh $A, B \subset V$

Đầu ra: $a \in A, b \in B$, sao cho $d(A, B) = d(a, b)$ nhỏ nhất, và $p_{min}(A, B) = p_{min}(a, b)$

B1. Ta sắp xếp các phần tử của A theo thứ tự nào đó

$$A = \{u_1, u_2, \dots, u_k\}$$

Đặt

$$A' = A, \Delta = \emptyset, d(\Delta) = +\infty$$

B2. Chọn $u_i \in A'$ có chỉ số i nhỏ nhất. Đặt $u = u_i$.

B3. Tìm cây đường đi ngắn nhất $T_{(u)}(V_{(u)}, E')$ cho đến khi $V_{(u)} \cap B \neq \emptyset$ hoặc $T_{(u)}$ cực đại ($V_{(u)} \cap B = \emptyset$).

Nếu $V_{(u)} \cap B \neq \emptyset$, thì

Chọn $v = V_{(u)} \cap B$

Xác định đường đi ngắn nhất $p_{min}(u, v)$

Nếu $d(u, v) < d(\Delta)$, thì đặt $d(\Delta) = d(u, v)$, $\Delta = p_{min}(u, v)$

Sang **B4**.

Nếu $T_{(u)}$ là cực đại, thì sang **B4**

B4 $A' = A' - \{u\}$

Nếu $A' \neq \emptyset$, thì quay về **B2**.

Nếu $A' = \emptyset$, thì **kết thúc**, $p_{min}(A, B) = \Delta$, $d(A, B) = d(\Delta)$

Độ phức tạp của giải thuật này là $\alpha \cdot O(|V|^2)$. Cuối cùng chúng tôi xin đề xuất một giải thuật khác hiệu quả hơn.

Giải thuật 3

Đầu vào: $G(V,E)$ và 2 tập đỉnh $A, B \subset V$

Đầu ra: $a \in A, b \in B$, sao cho $d(A,B) = d(a,b)$ nhỏ nhất, và $p_{min}(A,B) = p_{min}(a,b)$

B1. Ta sắp xếp các phần tử của A theo thứ tự nào đó

$$A = \{u_1, u_2, \dots, u_k\}$$

Đặt $A' = A, \Delta = \emptyset, d(\Delta) = +\infty$

B2. Chọn $u_i \in A'$ có chỉ số i nhỏ nhất. Đặt $u = u_i$.

B3. Tìm cây đường đi ngắn nhất $T_{(u)}(V_{(u)}, E')$ cho đến khi $V_{(u)} \cap B \neq \emptyset$ hoặc $T_{(u)}$ cực đại ($V_{(u)} \cap B = \emptyset$).

Nếu $T_{(u)}$ là cực đại, thì đặt $A' = A' - V_{(u)}$, sang **B4**

Ngược lại: Chọn $v = V_{(u)} \cap B$

Xác định đường đi ngắn nhất $p_{min}(u,v)$

Tìm $u' \in p_{min}(u,v) \cap A'$ gần đỉnh v nhất

$$A' := A' - p_{min}(u,v)$$

Nếu $d(u',v) < d(\Delta)$, thì đặt $d(\Delta) := d(u',v), \Delta := p_{min}(u',v)$

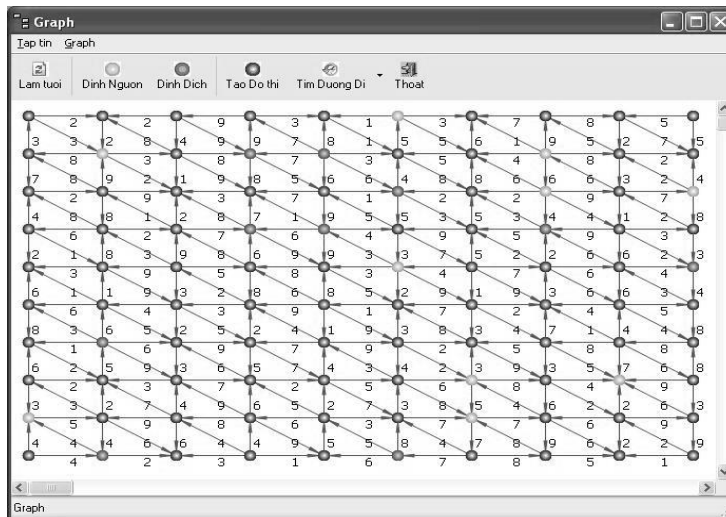
B4 Nếu $A' \neq \emptyset$. Quay về **B2**.

Nếu $A' = \emptyset$, thì **kết thúc**, $p_{min}(A,B) = \Delta, d(A,B) = d(\Delta)$

3. Kết quả thử nghiệm

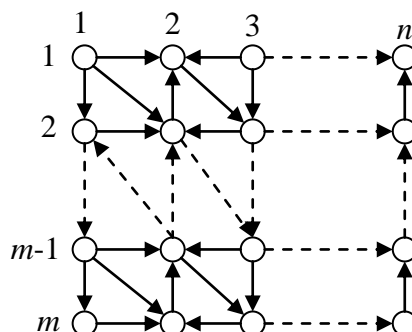
Để thử nghiệm và đánh giá các giải thuật trên chúng tôi đã viết một chương trình nhỏ cài đặt và minh họa cho các giải thuật bằng ngôn ngữ C#.

Chương trình có giao diện như sau:



Test các giải thuật:

Chương trình đã xét đến các đồ thị có dạng sau với mỗi đỉnh được xác định bởi tọa độ của nó (h,c) :



Với 100 đỉnh và 270 cạnh với hướng và trọng số được lựa chọn ngẫu nhiên từ 1 đến 9 và 20 đỉnh nguồn, 20 đỉnh đích cũng được lựa chọn ngẫu nhiên.
Kết quả trung bình sau 100 lần chạy:

Giải thuật	Thời gian (ms)
Giải thuật 1	4854.581
Giải thuật 2	41.981
Giải thuật 3	37.657

Ta thấy rằng ở đồ thị được xây dựng như trên, Giải thuật 3 hiệu quả hơn giải thuật 1: 99% và hiệu quả hơn giải thuật 2: 10%.

4. Kết luận

Qua bài báo này chúng tôi đã mở rộng bài toán tìm đường đi ngắn nhất giữa hai tập đỉnh từ đó đưa ra một số giải thuật để giải quyết bài toán này, các giải thuật đã được cài đặt trên máy tính và được so sánh đánh giá hiệu quả.

Hướng phát triển tiếp theo của đề tài này là tiếp tục mở rộng và phát triển bài toán tìm đường đi ngắn nhất giữa hai tập đỉnh đồng thời dựa trên các giải thuật đã thiết kế chúng tôi sẽ đưa ra một số giải thuật để giải nhiều bài toán khác nhau, cụ thể là các bài toán như:

- Bài toán tìm đường đi ngắn nhất với đa đỉnh nguồn, đa đỉnh đích
- Bài toán tìm đường kính giữa hai tập hợp đỉnh của đồ thị.

TÀI LIỆU THAM KHẢO

- [1] Trần Quốc Chiên, *Giáo trình Lý thuyết đồ thị*, Tài liệu lưu hành nội bộ, Đà Nẵng, 2002.
- [2] Thomas H.Cormen, Charles E.Leiserson, Ronald L.Rivest, *Introduction To Algorithms*, The MIT Press, 1999.
- [3] Kenneth H.Rosen, *Toán rời rạc ứng dụng trong tin học*, bản dịch tiếng Việt, Nxb KHKT, Hà Nội, 1997.
- [4] I-Lin Wang, *Shortest Paths and Multicommodity Network Flows*, A Thesis Presented to The Academic Faculty, School of Industrial and Systems Engineering Georgia Institute of Technology, April 2003.
- [5] K. M. Chandy, J. Misra, *Distributed Computation on Graphs, Shortest Path Algorithms*, University of Texas at Austin, November 1982.
- [6] Dijkstra.E, *A note on two problems in connection with graphs*, Numerische Mathematik, Vol.1, 1959.