

TRỰC QUAN HOÁ KẾT QUẢ TÌM ĐƯỜNG ĐI XE BUÝT DỰA TRÊN DỮ LIỆU DÒNG GIAO THÔNG KHÔNG - THỜI GIAN*

■ Hoàng Xuân Lộc**, Dương Ngọc Hiếu**, Trần Văn Hoài***, Nguyễn Thanh Dũng****

TÓM TẮT

Trực quan hóa khoa học và trực quan hóa thông tin là những lĩnh vực đa ngành mới được tập trung phát triển trong thập kỷ gần đây. Thời gian trước đây, trực quan hóa chủ yếu tập trung vào việc hiển thị và giúp đánh giá các kết quả mô phỏng. Tuy nhiên, với các dữ liệu lớn ngày nay thì trực quan hóa còn được giao một nhiệm vụ lớn hơn, đó là giúp khám phá dữ liệu để giúp các nhà khoa học hiểu hơn những khái niệm, những quan hệ và quá trình bên trong dữ liệu. Tại Việt Nam chủ đề giao thông ở các thành phố lớn như thành phố Hồ Chí Minh, Hà Nội đang được nhiều người quan tâm ở nhiều góc nhìn khác nhau. Trong những năm gần đây, xe buýt dần trở thành phương tiện công cộng phổ biến và chính yếu của người dân. Trong bài báo này, nhóm tác giả tập trung vào phân tích dữ liệu dòng giao thông xe buýt để xây dựng mô hình dữ liệu xe buýt hướng thời gian. Dựa vào mô hình dữ liệu này, nhóm tác giả phát triển giải thuật tìm đường đi xe buýt theo thời gian thực. Cuối cùng, bằng kỹ thuật WebGL, kết quả tìm đường xe buýt sẽ được hiển thị trên nền bản đồ 3D.

ABSTRACT

Visualizing results of finding bus lines based on data of the space and time traffic flow

Scientific visualization and information visualization are the interdisciplinary subfields that have attracted a great deal of attention in recent decades. In earlier time, visualization mainly focused on displaying and this was an essential tool for supporting to evaluate the simulation results. However, for bigger data, visualization has a greater mission to explore the data, concepts, relationships and processes within the data. Vietnamese traffic issues in big cities such as Ho Chi Minh City, Hanoi Capital are paid attention a lot in many different aspects. In a few years ago, the bus transport was quite popular and has gradually become the main transport of the Vietnamese people. In this research, the authors focus on analyzing bus traffic data in order to build the time oriented bus data model. Basing on the data model, the authors develop an algorithm to solve the shortest path problem of bus routing in real time. Finally, by employing WebGL technology, the shortest path will be displayed visually on the 3D map.

* Nghiên cứu này được tài trợ bởi ĐHQG TP.HCM trong khuôn khổ đề tài mã số C2014-20-07.

** ThS, Trường ĐH Bách Khoa - ĐHQG TP.HCM.

*** PGS.TS, Trường ĐH Bách Khoa - ĐHQG TP.HCM.

**** TS, Trường ĐH Văn Hiến.

1. Giới thiệu

Thực quan hóa khoa học và thực quan hóa thông tin là những lĩnh vực đa ngành mới được tập trung phát triển trong thập kỷ gần đây. Thời gian trước đó, thực quan hóa chủ yếu tập trung vào việc hiển thị và giúp đánh giá các kết quả mô phỏng. Tuy nhiên, với các dữ liệu lớn ngày nay trong rất nhiều lĩnh vực thì thực quan hóa còn được giao một nhiệm vụ lớn hơn, đó là giúp khám phá dữ liệu để giúp các nhà khoa học hiểu hơn những khái niệm, những quan hệ và quá trình bên trong dữ liệu. Trong xu thế đó, nhiều nhà khoa học đã đề xuất tách nhiệm vụ thực quan hóa ra hai nhánh khác nhau là thực quan hóa khoa học và thực quan hóa thông tin để phân biệt việc thực quan hai nhóm dữ liệu tương ứng là dữ liệu liên tục và dữ liệu rời rạc [9]. Tại Việt Nam chủ đề giao thông ở các thành phố lớn như thành phố Hồ Chí Minh, Hà Nội đang được nhiều người quan tâm ở nhiều góc nhìn khác nhau. Hiện nay đã có quá nhiều các nhận định trái chiều về giao thông Việt Nam, và cũng từ đó đã có rất nhiều các quyết sách chưa hợp lý. Theo nhận định chung của các nhà khoa học thì một trong những nguyên nhân là việc thiếu trầm trọng dữ liệu giao thông ở những thành phố lớn của Việt Nam. Ngoài ra việc thiếu những công cụ phân tích dữ liệu, ví dụ như công cụ thực quan hóa dữ liệu, cũng là một trong những nguyên nhân chính.

Trên thế giới, hiện nay cũng đã có khá nhiều nghiên cứu về việc xây dựng các công cụ thực quan hóa dữ liệu giao thông. Michael và các cộng sự đã kết hợp các mô hình nghiên cứu cũ về 3-D và đưa vào dòng dữ liệu giao thông thời gian thực [1]. Tuy nhiên, chỉ có hai đại lượng chính của dòng giao thông là tốc độ và khối lượng di chuyển được cung cấp và điều này đã hạn chế khá nhiều việc thực quan hóa. Hơn nữa, các tác giả chỉ trình bày hoạt hình lại các phương tiện dựa trên hai đại lượng trên chứ không mô hình thật các phương tiện và vị trí của chúng. Nguyên mẫu này chưa hướng đến được việc phân tích thực quan mà chỉ mới đạt được mức độ hoạt hình hóa sử dụng đồ họa máy tính. Sử dụng phương cách hoạt hình đã gây ra rất nhiều bất tiện trong việc phân tích dữ liệu dòng giao thông vì nhà

phân tích khó có thể thấy được các yếu tố tại những thời điểm khác nhau cùng một lúc. Như được chỉ ra trong [6], các công cụ thực quan cổ điển trở nên kém hiệu quả trong việc phân tích thực quan để làm rõ được mối quan hệ giữa các đối tượng di chuyển, hoặc các đại lượng mô tả dòng giao thông. Nói một cách khác, các phương thức và công cụ thực quan cổ điển khó giúp ích được cho các nhà quy hoạch. Bên cạnh công trình [6] thì có rất nhiều các công bố tương tự [7]. Tuy nhiên, hầu như tất cả các nghiên cứu này đề áp dụng các phương cách hiển thị 2-D, 3-D, 4-D hướng đến một môi trường hoạt hình nhằm phục vụ mục tiêu hiển thị nhiều hơn là giúp cho phân tích dòng giao thông. Một khảo sát khá chi tiết về thực quan hóa trong lĩnh vực quy hoạch đô thị có bao gồm dữ liệu giao thông được trình bày trong [8]. Trong tài liệu này, nhóm tác giả đã khảo sát rất tốt các phương pháp thực quan hóa phục vụ cho quy hoạch ở mức quản lý vĩ mô. Tuy nhiên, hướng nghiên cứu về thực quan hóa dòng giao thông chưa được đề cập.

Tại Việt Nam, có thể nói hầu như các nghiên cứu trong nước về thực quan hóa trong giao thông là gắn chặt với các hệ thống tin địa lý. Các nghiên cứu chủ yếu là sử dụng các công cụ có sẵn để thực quan hóa các đại lượng trong một lĩnh vực quản lý cụ thể nào đó, mà chưa đào sâu vào nghiên cứu cách thực quan hợp lý và mới để phục vụ việc phân tích. Tìm kiếm trong các thư viện về các công trình nghiên cứu, cũng như trên Internet thì có thể nhận thấy đa số các nghiên cứu là trong lãnh vực GIS.

Như đã phân tích trên, việc thiếu nhận định chính xác về giao thông Việt Nam là do thiếu dữ liệu và công cụ phân tích. Tuy nhiên để giải quyết toàn diện cả hai yếu tố trên trong hoàn cảnh hiện nay là tương đối khó khăn. Tại thành phố Hồ Chí Minh, xe buýt dần trở thành phương tiện công cộng phổ biến và chính yếu của người dân. Do đó trong nghiên cứu này, nhóm nghiên cứu tập trung phân tích dữ liệu giao thông xe buýt để từ đó xây dựng mô hình dữ liệu các tuyến xe buýt theo thời gian. Dựa vào mô hình dữ liệu theo thời gian trên, nhóm nghiên cứu triển khai bài toán tìm đường đi xe buýt theo thời gian thực.

Bài toán tìm đường đi ngắn nhất (shortest path) đã được nghiên cứu nhiều năm và có nhiều giải thuật giải quyết cho những trường hợp có điều kiện, ràng buộc kèm theo. Trong đó, vấn đề *tìm đường đi ngắn nhất theo thời gian* tạo ra nhiều thách thức với các nhà nghiên cứu. Cuối cùng, dựa vào công nghệ WebGL, kết quả tìm tuyến xe buýt theo thời gian thực sẽ được hiển thị trực quan trên nền bản đồ 3D. Trong phần còn lại của bài báo, nhóm nghiên cứu sẽ chia làm 4 phần chính. Phần 2 sẽ giới thiệu về dữ liệu xe buýt mà nhóm sử dụng cho nghiên cứu và mô hình lưu trữ dữ liệu giao thông xe buýt theo thời gian. Phần 3 nhóm sẽ giới thiệu về giải thuật tìm đường đi xe buýt theo thời gian thực. Phần 4 nhóm sẽ giới thiệu một số kết quả đạt được. Cuối cùng là một số kết luận và dự định nghiên cứu của nhóm trong tương lai.

2. Phân tích dữ liệu

2.1 Dữ liệu dòng giao thông xe buýt

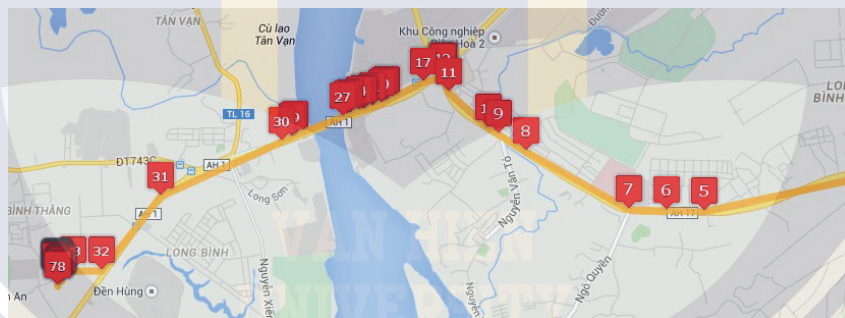
Dữ liệu dòng giao thông xe buýt được thu

thập thông qua các thiết bị GPS được gắn trên các xe buýt. Các thiết bị này được định thời để gửi những tín hiệu về máy chủ. Dữ liệu thô nhận được chỉ đơn giản với các dòng thông tin như sau:

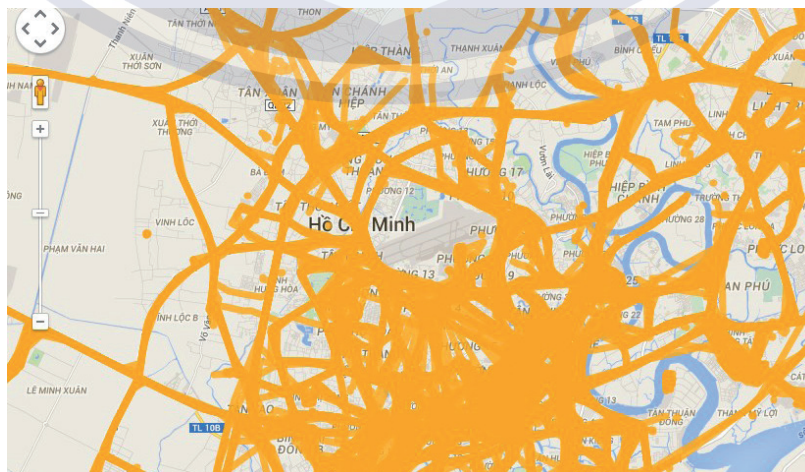
53U1917,10.751246,106.7019,0.0,0.0,0,1,0,
Wed Jun 04 00:00:19 ICT 2014

Dữ liệu này cho biết các thông tin bao gồm mã số quản lý của thiết bị, tọa độ của thiết bị và thời điểm gửi tín hiệu. Từ những thông tin trên có thể biết được vị trí của một thiết bị theo thời gian. Hình 1 thể hiện đường đi và vị trí của một thiết bị. Ở đó, thời gian giữa các tín hiệu không đều nhau, có lúc thưa hoặc dày đặc.

Với số lượng lớn thiết bị GPS được gắn cho các xe buýt trong TP.HCM (khoảng 6000 thiết bị), ta có được một mạng lưới dày đặc các đường đi của các xe buýt. Như hình 2 thể hiện một mạng lưới đường đi của các xe buýt trong khoảng thời gian từ 6 giờ đến 7 giờ.



Hình 1: Đường đi và vị trí của các tín hiệu trên bản đồ 2D



Hình 2: Đường đi của các xe buýt phủ khắp TP.HCM trên bản đồ 2D

Các thiết bị trả về một số lượng lớn dữ liệu, khoảng 2.5 triệu dữ liệu trong một ngày. Tuy nhiên trong đó có một số trường hợp dữ liệu không dùng được. Khi các tín hiệu được gửi đều đặn và thời gian giữa các lần gửi tín hiệu nhỏ thì có thể thấy rõ được đường đi của các phương tiện. Từ đó có thể tính toán được quãng đường đi và vận tốc của các phương tiện một cách tương đối chính xác. Nhưng trong thực tế thì nhiều trường hợp tín hiệu có thời gian ngắt quãng quá lớn. Có nhiều nguyên nhân như thiết bị hết năng lượng, tín hiệu bị mất, người điều khiển tắt thiết bị và điều này dẫn đến khó mà xác định được đường đi chính xác của xe buýt cũng như không thể xác định được vận tốc của xe buýt. Đối với những dữ liệu như vậy thì sẽ bị loại bỏ.

2.2. Dữ liệu bản đồ tuyến xe buýt

Từ dữ liệu bản đồ xe buýt hay còn gọi là mạng lưới xe buýt của trung tâm điều hành công cộng Tp. Hồ Chí Minh, có thể nhóm các thông tin về xe buýt thành các thành phần sau:

- Bản đồ xe buýt được phủ bởi hơn 110 tuyến, mỗi tuyến có lượt đi và lượt về. Trên mỗi tuyến xe buýt ta biết được lộ trình đi của tuyến, giá vé, lịch chạy.
- Có khoảng hơn 4300 trạm dừng.
- Các trạm dừng được nối với nhau bởi các tuyến đi qua đó.

Sau khi đã hiểu rõ bản đồ xe buýt ở thành phố Hồ Chí Minh, ta sẽ quan sát bản đồ xe buýt dưới góc nhìn đồ thị nhằm định ra mô hình lưu trữ trên máy tính:

- Các đỉnh là các trạm xe buýt.
- Các cạnh là đường đi của các tuyến xe buýt đi qua 2 trạm kế nhau. Mỗi cặp trạm kế nhau có thể có nhiều tuyến xe buýt đi qua, mỗi tuyến đi qua cặp trạm tạo thành một cạnh của đồ thị và cạnh này là cạnh có hướng. Mỗi cạnh có thông tin về quãng đường đi và thời gian đi. Ngoài ra, giữa những trạm ở gần nhau có thể di chuyển qua lại bằng cách đi bộ. Từ đó sẽ tạo thêm những cạnh mới cho đồ thị xe buýt. Việc thêm cạnh này để phù hợp với thực tế khi đi xe buýt.
- Có khoảng gần 10000 cạnh nối bởi xe buýt và khoảng 14000 cạnh nối bởi các trạm gần nhau.
- Đây là đồ thị thưa với số bậc trung bình là

2.6. Con số này chưa tính tới các cạnh nối bởi các trạm gần nhau.

Nhóm nghiên cứu đã tiến hành lưu trữ dữ liệu bản đồ xe buýt trên địa bàn Tp. Hồ Chí Minh trên nền cơ sở dữ liệu quan hệ và sử dụng hệ quản trị cơ sở dữ liệu SQL Server. Cho đến bước này, ta đã có được dữ liệu dòng giao thông và dữ liệu mạng lưới xe buýt tĩnh chưa có yếu tố thời gian động. Vì vậy việc cần làm tiếp theo là làm sao ánh xạ được dữ liệu dòng giao thông vào mạng lưới xe buýt tĩnh để có được dữ liệu một mạng lưới xe buýt theo thời gian.

2.3. Dữ liệu bản đồ tuyến xe buýt theo thời gian

Trước tiên, dữ liệu dòng giao thông cho biết thông tin một phương tiện đi tới điểm A ở thời điểm t_1 rồi đi tới điểm B ở thời điểm t_2 . Như vậy ta chỉ có được khoảng thời gian di chuyển giữa A và B, nhưng chưa biết đường đi thực sự giữa A và B. Như đã trình bày ở phần trước, ta chỉ quan tâm tới những trường hợp mà $(t_2 - t_1)$ là một khoảng thời gian nhỏ Δt thì có thể xem như đường đi từ A tới B là một đường thẳng. Từ đó ta có thể tính được quãng đường đi giữa A và B bằng công thức khoảng cách Euclide, cũng như vận tốc trung bình trên đoạn đường AB trong khoảng thời gian từ t_1 đến t_2 . Với cách này, ta sẽ tính được vận tốc cho tất cả các cặp vị trí liên tục.

Tiếp theo việc quan trọng cần làm là tính toán để xây dựng dữ liệu cho đồ thị xe buýt theo thời gian từ dữ liệu đã được tính toán ở trên. Ban đầu đồ thị xe buýt chỉ có thông tin về những bộ dữ liệu tĩnh ($\langle v_1, v_2 \rangle, d, r$), ở đó $\langle v_1, v_2 \rangle$ là đoạn đường đi có chiều dài d giữa trạm v_1 và v_2 bởi tuyến r , ở đây chưa cung cấp được thông tin về thời gian. Nhóm nghiên cứu chia thời gian một ngày thành những khoảng thời gian nhỏ T liên tục. Mỗi bộ dữ liệu tĩnh ($\langle v_1, v_2 \rangle, d, r$) với từng khoảng thời gian T sẽ có thông tin về vận tốc và chi phí thời gian tương ứng. Từ đó ta có những bộ dữ liệu theo thời gian gian ($\langle v_1, v_2 \rangle, d, r, T, s, t$) cho biết đoạn đường đi có chiều dài d giữa trạm v_1 và v_2 bởi tuyến r trong khoảng thời gian T với vận tốc s và chi phí thời gian để đi t . Để tính được vận tốc s và chi phí thời gian t cho từng bộ dữ liệu, nhóm nghiên cứu dùng giải thuật sau:

Bước 1: Với mỗi bộ dữ liệu ($\langle v_1, v_2 \rangle$, d , r , T), ở đó $\langle v_1, v_2 \rangle$ là đoạn đường đi có chiều dài d giữa trạm v_1 và v_2 bởi tuyến r trong khoảng thời gian T .

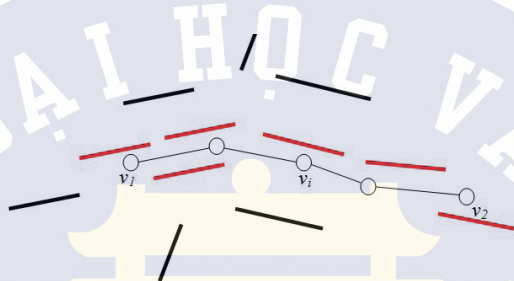
Tìm tất cả những đoạn AB đi trong khoảng thời gian từ t_1 đến t_2 , sao cho:

+ AB gần với $\langle v_1, v_2 \rangle$.

+ $[t_1, t_2]$ thuộc khoảng thời gian T .

Bước 2: Gán vận tốc trên đoạn đường $\langle v_1, v_2 \rangle$ bằng vận tốc trung bình của tất cả những đoạn AB tìm được. Từ đó tính được chi phí thời gian trên đoạn đường này.

Bước 3: Nếu không tìm được đoạn AB nào thì vận tốc sẽ bằng một giá trị vận tốc mặc định. Ta có thể sử dụng vận tốc trung bình của toàn bộ dữ liệu dòng giao thông.



Hình 3. Mô tả những đoạn màu đỏ được tính là gần với đoạn đường $\langle v_1, v_2 \rangle$

Trong Bước 1, mục tiêu là tìm ra được những đoạn AB gần với đoạn đường $\langle v_1, v_2 \rangle$, một đoạn AB được tính là gần với $\langle v_1, v_2 \rangle$, khi tồn tại một điểm v_i thuộc $\langle v_1, v_2 \rangle$ mà khoảng cách từ v_i tới trung điểm của đoạn AB nhỏ hơn một giá trị Δd cho trước. Như hình 3 mô tả những đoạn AB là những đoạn thẳng (màu đen và đỏ), trong đó những đoạn màu đỏ là những đoạn AB được tính là gần với $\langle v_1, v_2 \rangle$.

Còn Bước 2 chỉ đơn giản là tính ra vận tốc trung bình của tất cả những đoạn AB tìm được, rồi gán giá trị vận tốc này cho đoạn đường $\langle v_1, v_2 \rangle$. Bước B3 dùng để xử lý cho những đoạn đường không tìm được những đoạn AB nào gần nó, nên sẽ gán cho nó một giá trị vận tốc trung bình của dòng giao thông. Sau khi thực hiện theo giải thuật này, nhóm nghiên cứu đã xây dựng được dữ liệu đồ thị xe buýt theo thời gian. Trong phần sau, nhóm trình bày về mô hình đồ thị theo thời gian và giải thuật tìm đường đi xe buýt theo thời gian.

3. Giải thuật tìm đường đi xe buýt theo

thời gian thực

Trong phần này nhóm sẽ trình bày những khái niệm liên quan, mô tả bài toán cũng như giải thuật để giải quyết bài toán này.

3.1. Đồ thị phụ thuộc thời gian

Đồ thị phụ thuộc thời gian (G_T , E , V) (hoặc viết tắt là G_T) được đề cập chi tiết trong [10], được định nghĩa:

- $V = \{v_i\}$ là tập các đỉnh của đồ thị
- $E \subseteq V \times V$ là tập các cạnh của đồ thị.
- W là tập các hàm có giá trị dương.
- Với mỗi cạnh $(v_i, v_j) \in E$, có một hàm $w_{ij}(t) \in W$, với t là biến thời gian trong một khoảng thời gian.
- Hàm độ trễ-cạnh (edge-delay function) $w_{ij}(t)$ xác định thời gian để di chuyển từ đỉnh v_i đến đỉnh v_j nếu xuất phát từ đỉnh v_i vào thời điểm t .

3.2 Bài toán tìm đường đi ngắn nhất với đồ thị phụ thuộc thời gian

Định nghĩa: Bài toán tìm đường đi ngắn nhất với đồ thị phụ thuộc thời gian là tìm đường đi có

thời gian di chuyển nhỏ nhất từ điểm bắt đầu đến điểm đích với thời điểm bắt đầu trên đồ thị phụ thuộc thời gian. Thời gian di chuyển là thời điểm đến điểm đích trừ thời điểm bắt đầu, gọi tắt là bài toán TDSP (time-dependent shortest path).

Trong [10] cũng đề cập tới bài toán TDSP tìm ra đường đi với thời gian di chuyển nhỏ nhất và thời điểm bắt đầu để đi cho vấn đề vận chuyển. Còn ở đây, tìm ra đường đi với thời gian di chuyển nhỏ nhất ở mỗi thời điểm được biết trước.

Ở mỗi đỉnh có các đại lượng:

- $\tilde{w}(v_i)$ kí hiệu cho thời gian đợi (waiting time) tại đỉnh
- $arrive(v_i)$ kí hiệu cho thời điểm đến đỉnh
- $depart(v_i)$ kí hiệu cho thời điểm xuất phát từ đỉnh

Mối quan hệ của ba đại lượng trên được thể hiện qua công thức sau:

$$depart(v_i) = arrive(v_i) + \tilde{w}(v_i)$$

Cho một đường đi

$$p = (v_1, v_2, \dots, v_{k-1}, v_k)$$

và thời điểm bắt đầu là $arrive(v_1) = t$,

$$arrive(v_2) = depart(v_1) + w_{1,2}(depart(v_1))$$

...

$$arrive(v_k) = depart(v_{k-1}) + w_{k-1,k}(depart(v_{k-1}))$$

$$g_p(t) = arrive(v_k)$$

$g_p(t)$ là hàm thời gian đến từ v_1 tới v_k theo đường đi p , với thời điểm bắt đầu t . Từ đó ta có hàm thời gian di chuyển theo đường đi p là $g_p(t) - t$. Mục tiêu của bài toán TDSP là tìm ra đường đi có thời gian di chuyển ngắn nhất p^* :

$$g_{p^*}(t) - t = \min_{p, \tilde{w}^*} \{g_p(t) - t\}$$

Do có thêm yếu tố thời gian nên không gian nghiệm bài toán TDSP lớn hơn nhiều so với bài toán tìm đường đi ngắn nhất không có yếu tố thời gian. Giải thuật sau dùng phương pháp gán nhãn (labeling method) sử dụng để giải quyết bài toán

TDSP dựa trên đề xuất từ [11], nhưng giải thuật này hoạt động trên đồ thị thỏa mãn giả định là các cạnh của đồ thị đều có tính chất FIFO [10]

[12]. Tính FIFO: Một cạnh (v_i, v_j) có tính FIFO

nếu và chỉ nếu $w_{i,j}(t_0) \leq t_\Delta + w_{i,j}(t_0 + t_\Delta)$

với $t_\Delta \geq 0$ hoặc $t_1 + w_{i,j}(t_1) \leq t_2 + w_{i,j}(t_2)$ với

$t_2 \geq t_1$. Tính chất này khẳng định nếu bắt đầu xuất phát ở một cạnh trước thì sẽ ra khỏi cạnh đó trước. Tính chất này phù hợp với việc lưu thông trên đường nếu mọi xe đều chạy với đúng tốc độ hiện tại trên đường đó, và cũng phù hợp với phương tiện là xe buýt, với kích thước lớn trong khi đường lại nhỏ trong địa bàn Tp.Hồ Chí Minh.

Đầu vào:

Đơn đồ thị

Điểm bắt đầu và điểm cuối s, e ; thời gian bắt đầu ts

Đầu ra: Đường đi p từ s đến e

$fs = ts$

$Q.enqueue(\{fs, s\})$, Q is a priority queue containing pairs, $\{fi, vi\}$, ordered by fi in ascending order.

While Q is not empty

$\{fi, vi\} = Q.dequeue()$

If vi is e , stop

For each neighbors vk of vi

if vk is not visited

$fk = fi + w_{i,k}(fi)$

$Q.enqueue(\{fk, vk\})$

$label(vk) = \{fk, vi\}$

elseif $\{fi + w_{i,k}(fi), vi\}$ is better label(vk)

$fk = fi + w_{i,k}(fi)$

$Q.enqueue(\{fk, vk\})$

$label(vk) = \{fk, vi\}$

end for

end while

if e is visited

$\{te, vp\} = label(e)$

$t^* = te - ts$

$p = e$

while $vp \neq s$

$p = vp.p$

$\{fi, vp\} = label(vp)$

end while

p = s.p
end if

Giải thuật trên dùng một hàng đợi chứa những cặp giá trị bao gồm một đỉnh vi và thời gian đi tới đỉnh đó f_i từ đỉnh bắt đầu s. Giải thuật kết thúc khi gặp đỉnh cuối e hoặc hàng đợi trống (không có nghiệm). Giải thuật sẽ thực hiện việc tính toán các giá trị thời gian f_k để đi tới một đỉnh v_k và đưa cặp giá trị nào vào trong hàng đợi.

3.3 Áp dụng cho bài toán tìm đường đi xe buýt theo thời gian

Như đã trình bày ở trên là về bài toán tìm đường đi ngắn nhất cho một đồ thị phụ thuộc thời gian nói chung. Còn bài toán tìm đường đi xe buýt theo thời gian mà nhóm muốn giải quyết được là tìm đường đi bằng xe buýt có thời gian di chuyển nhỏ nhất từ điểm bắt đầu đến điểm đích với thời điểm bắt đầu trên đồ thị xe buýt theo thời gian thỏa mãn các ràng buộc (số lần chuyển tuyến, số tiền cần dùng, thời gian đợi). Không mất tính tổng quát, nghiên cứu trong bài báo chỉ sử dụng ràng buộc số lần chuyển tuyến, các ràng buộc khác ở thể áp dụng tương tự.

Điểm khác biệt ở đây là đồ thị xe buýt theo thời gian có một số đặc điểm riêng của nó như: được kết nối bởi một hệ thống các tuyến xe buýt, có thể di chuyển bằng cách đi bộ giữa các trạm gần nhau, tốn chi phí khi chuyển từ chuyến này sang chuyến khác. Đó đó đồ thị xe buýt theo thời gian có một số đặc điểm khác như sau:

Với mỗi cạnh $(v_i, v_j) \in E$,

$$(v_i, v_j) = \begin{cases} d_{i,j} \\ w_{i,j}(t) \\ k_{i,j} \end{cases} \begin{cases} \text{chuyển tuyến} \\ \text{còn lại} \end{cases}$$

cho biết một cạnh đó có phải là cạnh chuyển tuyến (giá trị là 1) hay không. Có ba loại cạnh trong đơn đồ thị này:

- Loại a: Cạnh được kết nối bởi các tuyến xe, với $d_{i,j}$ là quãng đường di chuyển, $w_{i,j}(t)$ là thời gian di chuyển, $k_{i,j} = 0$.
- Loại b: Cạnh cho biết thông tin chuyển

tuyến, $d_{i,j} = 0$, $k_{i,j} = 1$, $w_{i,j}(t) = \tilde{w}(v_i)$ chính là thời gian đợi để chuyển sang tuyến mới

tại v_j .

- Loại c: Cạnh thể hiện việc đi bộ, với

$d_{i,j}$ là quãng đường đi bộ, $w_{i,j}(t) = w_0$ là thời

gian đi bộ, $k_{i,j} = 0$.

Để giải quyết bài toán tìm đường đi xe buýt theo thời gian với ràng buộc về số lần chuyển tuyến, nhóm đề xuất giải thuật (gọi tắt là D3) dựa trên phương pháp gán đa nhãn (multi-labeling method) sau:

Giải thuật D3 có vài điểm đáng chú ý là tại mỗi đỉnh có thể gán nhiều nhãn, trong giải thuật có hai thao tác là chọn nhãn có giá tốt hơn và nhãn có thời gian tốt nhất. Với 2 nhãn $n_i = \{f_i, c_i, v_i\}$, $n_j = \{f_j, c_j, v_j\}$, nhãn n_i tốt hơn nhãn n_j khi và chỉ khi $f_i < f_j$ (tốt hơn về mặt thời gian) hoặc $c_i < c_j$ (tốt hơn về mặt chuyển tuyến). Còn nhãn có thời gian tốt nhất là nhãn có giá trị f nhỏ nhất trong các nhãn còn lại, thao tác này dùng để chọn ra được đường đi có thời gian nhỏ nhất.

4. Kết quả

Để hiện thực chương trình mô phỏng, nhóm dùng công nghệ WebGL trên nền web. WebGL có thể chạy trên đối tượng đồ họa 3 chiều, trực tiếp trên trình duyệt, WebGL sử dụng ngôn ngữ lập trình Javascript. Công nghệ này có thể truy cập vào card đồ họa sử dụng JavaScript. WebGL có thể chi phối giao diện DOM, dựa trên công nghệ OpenGL ES 2.0, có thể vận hành trên nhiều thiết bị khác nhau gồm có máy tính cá nhân, điện thoại di động... WebGL có thể chạy trực tiếp trên card đồ họa. Có một vài thư viện sử dụng cho WebGL tạo ra để user có thể dễ dàng sử dụng. Tuy nhiên sẽ tốt hơn nếu hiểu rõ cốt lõi của WebGL từ đó tiếp cận công nghệ này. WebGL có thể chạy thời gian thực với đối tượng đồ họa 3D và tất nhiên WebGL có thể áp dụng cho nhiều lĩnh vực thú vị với những ai muốn tiếp cận công nghệ này.

Ngoài ra nhóm nghiên cứu lựa chọn nền bản đồ Open Street Map. Đây là dự án cho phép mọi người cùng xây dựng bản đồ thế giới mở với một

Đầu vào:

Đơn đồ thị

Điểm bắt đầu và điểm cuối s, e ; Thời gian bắt đầu ts

Số lần đổi tuyến cho phép m

Đầu ra: Đường đi p từ s đến e

$fs = ts, c = 0, parent = nil$

$Q.enqueue(\{fs, c, s, nil\})$, Q is a priority queue containing vectors, $\{fi, c, vi, parent\}$, ordered by fi in ascending order, c is transit times, $c \leq m$.

While Q is not empty

$\{fi, c, vi, vp\} = Q.dequeue()$

If vi is e , stop

For each neighbors vk of vi

if $c + ki, j \leq m$

if vk is not visited

$fk = fi + wi, k(fi)$

$Q.enqueue(\{fk, c + ki, j, vk, vi\})$

$labels(vk) = \{fk, c + ki, j, vi\}$

set vk is visited

elseif $\{fi + wi, k(fi), c + ki, j, vi\}$ is better than $labels(vi)$

$fk = fi + wi, k(fi)$

$Q.enqueue(\{fk, c + ki, j, vk, vi\})$

$labels(vk).add(\{fk, c + ki, j, vi\})$

end if

end if

end for

end while

if e is visited

$\{te, c, vp\}$ is label has time is best of $labels(e)$

$t* = te - ts, p = e$

while $vp \neq s$

$p = vp.p$

$\{fi, vp\}$ is label has time is best of $labels(vp)$

end while

cộng đồng người dùng đông đảo tham gia hỗ trợ. Ở Việt Nam, cũng có một dự án con nhằm xây dựng bản đồ chi tiết và đầy đủ hơn. Đây là nguồn bản đồ mở khá chi tiết để phát triển những ứng dụng về GIS. Bên cạnh đó, nhóm còn sử dụng một số thư viện 3D như ReadyMap – thư viện hỗ trợ mô phỏng trái đất, và ThreeJS cho phép tạo những đối tượng 3D một cách tiện lợi.

Hiện tại nhóm đã hiện thực chương trình trực quan hóa thông tin xe buýt bao gồm hai nội dung cần trực quan hoá. Nội dung trực quan hoá thứ

nhất đó là phân tích bản đồ xe buýt theo thời gian. Trong đó chiều cao thể hiện thông tin về thời gian, với những đoạn dốc cho biết đi qua đoạn đó mất nhiều thời gian hơn, ứng với vận tốc chậm và ngược lại. Nội dung trực quan hoá thứ hai đó là biểu diễn kết quả giải thuật tìm đường đi ngắn nhất lộ trình xe buýt theo thời gian thực. Hình 4a, 4b và hình 5 cho thấy việc hiển thị kết quả tìm đường đi xe buýt ngắn nhất theo thời gian trên nền bản đồ 3D giúp người dùng cuối dễ dàng quan sát được tốc độ di chuyển của xe

theo thời gian thực hơn. Màu sắc của từng đoạn đường được thể hiện theo Bảng 1.

Màu sắc	Giá trị vận tốc
Đỏ	$\leq 5\text{km/h}$
Vàng	$\leq 10\text{km/h}$
Xanh dương	$\leq 20\text{km/h}$
Xanh lá cây	$> 20\text{km/h}$

Bảng 1: Ghi chú màu thể hiện đoạn đường đi xe buýt trên nền bản đồ 3D.

Như hình 4a và 4b, vận tốc chậm ở những nơi như điểm bắt đầu xuất phát, điểm kết thúc lộ trình, và những chốt lưu thông. Ngoài ra, quan sát thấy được vận tốc của các đoạn đường khi xuất phát vào lúc 8h (mật độ lưu thông cao) thấp hơn lúc 15h (mật độ lưu thông thấp hơn).

5. Kết luận

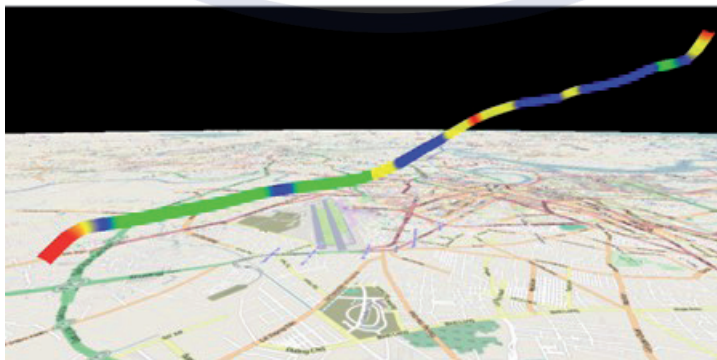
Trong bài báo, nhóm nghiên cứu đã trình bày các kết quả nghiên cứu nổi bật liên quan đến bài



Hình 4a: Trực quan hóa một lộ trình của xe buýt xuất phát từ bến xe An Sương đến Chợ Lớn bắt đầu từ lúc 8h trên bản đồ 3D, có góc nhìn từ trên xuống



Hình 4b: Trực quan hóa một lộ trình của xe buýt xuất phát từ bến xe An Sương đến Chợ Lớn bắt đầu từ lúc 15h trên bản đồ 3D, có góc nhìn từ trên xuống



Hình 5: Kết quả giống như hình 4a, có góc nhìn ngang thể hiện thời gian di chuyển

toán giao thông Việt Nam nói chung và bài toán xe buýt tại Tp. Hồ Chí Minh nói riêng. Nhóm đã thu thập được dữ liệu giao thông xe buýt tại Tp. Hồ Chí Minh và dựa vào dữ liệu này đề xuất mô hình lưu trữ dữ liệu tuyến xe buýt theo thời gian. Nhóm cũng đã phát triển giải thuật tìm đường đi xe buýt ngắn nhất theo thời gian thực, đó là giải thuật được phát triển dựa trên phương pháp gán đa nhãn. Để hiển thị tốc độ di chuyển của xe buýt theo từng tuyến, từng thời điểm cũng như

kết quả tìm đường đi xe buýt ngắn nhất theo thời gian thực một cách trực quan trên nền bản đồ 3D, nhóm đề tài lựa chọn công nghệ WebGL và nền bản đồ Open Street Map. Kết quả bước đầu rất khả quan và là động lực cho nhóm đề tài tiếp tục nghiên cứu, tối ưu các nội dung đã hoàn thành. Trong thời gian tới, nhóm sẽ tiến hành nghiên cứu thêm các phương cách trực quan hoá dữ liệu giao thông để người dùng có thể dễ dàng quan sát đặc biệt vào những giờ cao điểm.

TÀI LIỆU THAM KHẢO

- [1] “*Wide-area, Four-Dimensional, Real-time, Interactive Transportation System Visualization*”, Michael L. Pack, Phillip Weisberg và Sujal Bista, Transportation Research Record: Journal of the Transportation Research Board, pp. 97-108, 2007.
- [2] “*Data Visualization: Principles and Practice*”, Alexandru C. Telea, A K Peters/CRC Press, 2007.
- [3] “*Visualization Analysis of Multivariate Spatial – Temporal Data of the Red Army Long March in China*”, Ding Ma, Zhimim Ma, Lumin Meng và Xia Li, International Symposium on Spatial Analysis, Spatial-Temporal Data Modeling, and Data Mining, Proceedings of SPIE Vol. 7492, 74920X, 2009.
- [4] “*Simulation of Mixed Traffic Flow within Intersection*”, V.H. Thanh và V. H. Tran, Proceedings of KSE2010 - 2nd International Conference on Knowledge and Systems Engineering, IEEE Computer Society, pp. 131-140, 2010.
- [5] “*Direct Multiple Shooting Method for Solving Approximate Shortest Paths Problem*”, P.T. An, N.N. Hai và T.V. Hoai, Journal of Computational and Applied Mathematics 244, pp. 67-76, 2013.
- [6] “*Dynamic Environmental Visualization within a Virtual Environment*”, Huang B., presented at 83rd Annual Meeting of Transportation Research Board, Washington D.C., 2004.
- [7] “*The Integration of ArcView/3D Analyst and 3 Dimensional Visualization Technologies for Interactive Visualization of Urban Environments*”, Fletcher, M.B., O'Toole, B. E. và Banks, R. G., Proc. Twentieth Annual ESRI User Conference, 2000.
- [8] “*Visualizing Urban Features: Geomatics Decision Support for Canadian Cities*”, Eric J. Miller và các cộng sự, Dự án TSII-201, 2011.
- [9] “*Milestones in the history of thematic cartography, statistical graphics, and data visualization*”, Michael Friendly, Trường Đại Học York Canada, 2009.
- [10] “*Finding time-dependent shortest paths over large graphs*”, B. Ding, J. X. Yu và L. Qin, Proceedings of the 11th International Conference on Extending Database Technology: Advances in Database Technology, pages 205–216. ACM, 2008.
- [11] “*An appraisal of some shortest-path algorithms*”, S. E. Dreyfus, In Operations Research Vol. 17, No. 3, 1969.
- [12] “*Shortest route with time dependent length of edges and limited delay in nodes*”. J. Halpern, In MMO, 1969.