

Một số vấn đề về đồ họa máy tính

Đại học Khoa học Tự nhiên

Khoa Toán - Cơ - Tin học

Khoa học dữ liệu K4

Tháng 12 năm 2022

Chương 7 - Kỹ thuật đồ họa cho dữ liệu hướng thời gian

Nguyễn Mạnh Linh, Nguyễn Đức Thịnh

Trong chương này chúng tôi giải thích chi tiết các kỹ thuật đồ họa cho dữ liệu hướng thời gian. Chúng tôi tổ chức và cấu trúc chương theo cuốn sách ... Do đó, trước tiên chúng tôi nhấn mạnh tầm quan trọng của việc xử lý các yếu tố thời gian qua ví dụ. Thứ hai, chúng tôi đưa ra các khái niệm cần thiết và các khía cạnh của thời gian và dữ liệu hướng thời gian. Các tập dữ liệu cơ bản đã có trong chương 2. Ở chương này, chúng tôi tập trung vào các đặc điểm của thời gian. Thứ 3, chúng tôi cung cấp một cái nhìn tổng quan về các kỹ thuật trực quan hóa khác nhau. Thứ tư, chúng tôi giải thích ngắn gọn về *TimeBench*, là mô hình và một thư viện phần mềm phân tích dữ liệu hướng thời gian. Chúng tôi kết luận với một đánh giá về phân. Chúng tôi kết luận với đánh giá phân loại theo các kỹ thuật trực quan hóa được trình bày và giới thiệu *TimeViz Browser*, là một kho lưu trữ để hỗ trợ các nhà nghiên cứu và học viên trong việc tìm kiếm các kỹ thuật trực quan hóa tích hợp, tiếp theo là các bài đọc, bài tập và đề án thực hành.

1 Giới thiệu

Bản thân thời gian vốn là dữ liệu một chiều, là trung tâm của việc phát hiện xu hướng, xác định các mẫu lặp lại và mối các mối quan hệ trong dữ liệu. Thời gian và dữ liệu hướng thời gian có những đặc điểm riêng biệt đáng để chúng ta nghiên cứu và xử lý như một kiểu dữ liệu riêng. Do tầm quan trọng của dữ liệu hướng thời gian, cấu trúc của nó đã được nghiên cứu trong nhiều ấn phẩm khoa học.

Các ví dụ sau minh họa tầm quan trọng của thời gian. Hình (td) thể hiện ba biểu diễn trực quan khác nhau của cùng một tập dữ liệu hướng thời gian, chứa số ca mắc cúm hàng ngày xảy ra ở miền bắc nước Đức trong khoảng thời gian ba năm. Dữ liệu thể hiện một mô hình tuần hoàn mạnh. Hình ngoài cùng bên trái sử dụng một biểu đồ đường đơn giản để trực quan hóa dữ liệu. Mặc dù thời gian cao điểm có thể được nhận ra dễ dàng, tuy nhiên khi kiểm tra biểu diễn này, hành vi theo chu kỳ của dữ liệu chỉ có thể được đoán và thật khó để phân biệt các chu kỳ mà trên thực tế có tồn tại. Ngược lại, hình ở giữa và hình bên phải biểu diễn một vòng tròn nhấn mạnh các đặc tính tuần hoàn của dữ liệu bằng cách sử dụng trục thời gian hình xoắn ốc. Với hình xoắn ốc bên trái, mô hình tuần hoàn không được phát hiện ra. Điều này là do độ dài chu kỳ được đặt là 24 ngày, không khớp với sự tuần hoàn của dữ liệu. Biểu diễn xoắn ốc bên phải được cài đặt với chu kỳ 28 ngày ngay lập tức tiết lộ sự tuần hoàn trong dữ liệu.