**TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI**

**ĐỒ ÁN TỐT NGHIỆP**

**Đề xuất mô hình và giải thuật tối ưu tần suất gửi tin trong hệ quan trắc môi trường di động.**

**NGUYỄN ĐÌNH TUẤN ANH**

Anh.ndt164767@sis.hust.edu.vn

**Ngành …**

**Chuyên ngành …**

|  |  |
| --- | --- |
| **Giảng viên hướng dẫn:** | TS. Nguyễn Phi Lê |
| **Bộ môn:** | Công nghệ phần mềm |
| **Viện:** | Công nghệ thông tin – Truyền thông |
| **HÀ NỘI, 12/2019** Lời cam kết . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  Điện thoại liên lạc: . . . . . . . . . . . . . . . . . . . . . Email: . . . . . . . . . . . . . . . . . . . . . . . .  Lớp: . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Hệ đào tạo: . . . . . . . . . . . . . . . . . . . .  Tôi – *điền tên sinh viên* – cam kết Đồ án Tốt nghiệp (ĐATN) là công trình nghiên cứu của bản thân tôi dưới sự hướng dẫn của *học hàm học vị+điền tên giáo viên hướng dẫn*. Các kết quả nêu trong ĐATN là trung thực, là thành quả của riêng tôi, không sao chép theo bất kỳ công trình nào khác. Tất cả những tham khảo trong ĐATN – bao gồm hình ảnh, bảng biểu, số liệu, và các câu từ trích dẫn – đều được ghi rõ ràng và đầy đủ nguồn gốc trong danh mục tài liệu tham khảo. Tôi xin hoàn toàn chịu trách nhiệm với dù chỉ một sao chép vi phạm quy chế của nhà trường.   |  |  | | --- | --- | |  | *Hà Nội, ngày tháng năm*  Tác giả ĐATN  *Họ và tên sinh viên* |  Lời cảm ơn Quãng thời gian học tập với vai trò là sinh viên Đại học Bách Khoa Hà Nội đã mang đến cho em nhiều cung bậc cảm xúc khác nhau, áp lực học tập có, vui buồn cũng có và cả những niềm hạnh phúc. Những ngày đầu bước chân vào cổng trường, dưới những hàng cây mát rượi của buổi chiều hè, em đã cảm thấy đây chính là nơi bản thân sẽ có 5 năm quý giá của tuổi sinh viên. Từ một cậu học sinh chưa có hiểu biết gì về công nghệ và thông tin, nay đã sắp tốt nghiệp dưới mái trường kỹ thuật hàng đầu Việt Nam, em có đôi lời nhắn gửi đến các thầy cô và bạn bè tại ngôi trường thân yêu này.  Trước hết em gửi lời cảm ơn tới tất cả thầy cô của trường Đại học Bách Khoa Hà Nội, các thầy, các cô là một phần không thể thiếu trong việc đào tạo ra thế hệ các sinh viên tài năng, và đầy bản lĩnh, sẵn sàng đóng góp xây dựng đất nước. Cảm ơn các thầy cô không chỉ ở những bài học lý thuyết hay thực hành mà cả các bài học cuộc sống mà các thầy cô đã trao đổi với em trong quá trình giảng dạy.  Em đặc biệt gửi lời cảm ơn đến cô Nguyễn Phi Lê, tuy em cũng chỉ mới học tập và làm việc với cô trong thời gian ngắn, cô đã rất tận tình giúp đỡ em trong mọi câu hỏi thắc mắc cả về khoa học lẫn cuộc sống. Trong quá trình làm đồ án tốt nghiệp, cô không chỉ đóng vai trò là người hướng dẫn mà còn là người động viên, giúp em và cả các bạn luôn yên tâm, tập trung, nỗ lực hoàn thành đồ án của mình.  Tiếp theo, em xin gửi lời cảm ơn đến tập thể lớp Kỹ sư tài năng Công nghệ thông tin khóa 61, những người bạn đã luôn đồng hành cùng em trong những năm qua.  Cuối cùng em muốn gửi lời cảm ơn đến bố và mẹ của em, đã luôn động viên em cố gắng, cung cấp cho em điều kiện thuận lợi nhất để an tâm học hành.  Trong quá trình làm đồ án, với vai trò là một sinh viên còn nhiều kiến thức hạn chế, em không tránh khỏi những sai sót, em mong nhận được sự góp ý của thầy cô và các bạn để hoàn thiện kiến thức của bản thân.  Em xin chân thành cảm ơn ! Tóm tắt Nội dung  [Lời cam kết 2](#_Toc72066326)  [Lời cảm ơn 3](#_Toc72066327)  [Tóm tắt 4](#_Toc72066328)  [Chương 1: Giới thiệu đề tài 7](#_Toc72066329)  [1.1 Đặt vấn đề 7](#_Toc72066330)  [1.2 Mục tiêu và định hướng giải pháp 8](#_Toc72066331)  [1.4 Đóng góp của đồ án 8](#_Toc72066332)  [1.5 Bố cục của đồ án 8](#_Toc72066333)  [Chương 2: Các kiến thức nền tảng 9](#_Toc72066334)  [2.1 Phân phối poisson và phân phối poisson đa chiều. 9](#_Toc72066335)  [2.2 Các phương pháp tính gần đúng tích phân. 9](#_Toc72066336)  [1.3 Tổng quan các giải thuật tối ưu 13](#_Toc72066337)  [1.3.1 Giải thuật bầy ong nhân tạo( artificial bee colony). 13](#_Toc72066338)  [2.3.2 Giải thuật PSO . 14](#_Toc72066339)  [Chương 3: Giải thuật tối ưu cho bài toán đo chất lượng không khí bằng cảm biến tích hợp trên xe. 14](#_Toc72066340)  [3.1Mô tả bài toán 14](#_Toc72066341)  [3.1.1 Mô hình. 14](#_Toc72066342)  [3.1.2 Xác định mục tiêu và hàm tối ưu. 14](#_Toc72066343)  [3.2 Tổng quan các hương tiếp cận giải quyết bài toán. 14](#_Toc72066344)  [3.3 Phương pháp tối ưu sử dụng giải thuật bầy ong nhân tạo 14](#_Toc72066345)  [3.4 Phương pháp tối ưu sử dụng thuật toán PSO. 14](#_Toc72066346)  [Chương 4: Đánh giá thực nghiệm. 14](#_Toc72066347)  [4.1 Các tham số đánh giá. 14](#_Toc72066348)  [4.2 Phương pháp thí nghiệm. 14](#_Toc72066349)  [4.3 Kết quả thí nghiệm. 14](#_Toc72066350)  [4.3.1 Tham số tối ưu cho phương pháp tính gần đúng tích phân, so sánh với phương pháp monte - carlo. 15](#_Toc72066351)  [4.3.2 So sánh kết quả thuật toán ABC với PSO. 15](#_Toc72066352)  [4.3.3 Ảnh hưởng của số làn đường và khoảng cách giữa các làn đến kết quả tối ưu. 15](#_Toc72066353)  [Chương 5: Kết luận 15](#_Toc72066354)  [5.1 Kết luận 15](#_Toc72066355)  [5.2 Định hướng phát triển của đề tài 15](#_Toc72066356) Chương 1: Giới thiệu đề tàiĐặt vấn đề Việt Nam với vai trò là một nước đang phát triển, sự phát triển kinh tế luôn đi kèm với quá trình công nghiệp hóa và hiện đại hóa, bên cạnh những lợi ích về mặt cơ sở vật chất, đời sống nhân dân được cải thiện, thì những ảnh hưởng của quá trình ô nhiễm môi trường lên đời sống con người cũng ngày càng mạnh mẽ. Một trong số đó là vấn đề ô nhiễm không khí, gây tổn hại không nhỏ đến sức khỏe con người.  Đặc biệt, các chuyên gia cho biết cần quan tâm tới hai chỉ số bụi mịn: PM10 và PM2.5 – các hạt bụi có đường kính nhỏ hơn 10 µm và nhỏ hơn 2.5 µm. Những năm gần đây, nước ta có sự xuất hiện của bụi siêu mịn PM1.0 (dưới 1µm) và bụi nano PM0.1 (dưới 0.1 µm). Nhờ kích cỡ rất nhỏ của mình, bụi siêu mịn PM2.5 có thể qua đường hô hấp để len lỏi vào sâu trong phổi, gây ra các bệnh hô hấp. Tình hình chất lượng không khí hiện nay của Việt Nam đang trong tình trạng xấu, không chỉ do yếu tố không khí và địa lý mà còn do hành động thiếu trách nhiệm của con người, các công trình và thiếu ý thức chung trong việc giữ gìn môi trường xung quanh. Khi người dân thấy được mức độ ảnh hưởng của các chỉ số bụi mịn này tới sức khỏe của bản thân, mọi người sẽ có ý thức để cùng chung tay bảo vệ môi trường.  Vì vậy việc theo dõi chất lượng không khí trong các thành phố lớn là hết sức cần thiết, từ đó cho phép người dân biết được chất lượng không khí xung quanh nơi ở, nơi làm việc, giúp cho họ có các biện pháp bảo vệ sức khỏe của bản thân và cả những người xung quanh. Hơn nữa hiện nay các đề tài nghiên cứu khoa học liên quan đến chất lượng không khí cũng đang thu hút được rất nhiều sự quan tâm của các nhà nghiên cứu trong nhiều lĩnh vực, vì vậy đặt ra nhu cầu thu thập dữ liệu về chất lượng không khí không chỉ trong thành phố để phục vụ cho con người mà còn thu thập dữ liệu trên một vùng bất kỳ để làm phương tiện cho các nghiên cứu khoa học liên quan. Mục tiêu và định hướng giải phápĐóng góp của đồ án Đồ án này có 3 đóng góp chính như sau :   * Đề xuất một mô hình hệ quan trắc di động nhắm mục đích thu thập dữ liệu về chất lượng không khí trên một vùng cần theo dõi. * Đề xuất một mô hình toán học khả thi trong hệ quan trắc đó, xác định hàm chi phí và hàm lợi ích của mô hình. * Đề xuất các tham số phù hợp cho việc tính toán và giải thuật tối ưu cho hệ quan trắc di động đã nêu trên.  Bố cục của đồ án Kết cấu của đồ án của em trong phần còn lại như sau:  Chương 2 em sẽ trình bày về các nghiên cứu liên quan đến vấn đề trên, những ý tưởng chính của mô hình và giải pháp cũng như phân tích các ưu nhược điểm của các nghiên cứu đó. Từ đó đề xuất ra mô hình và giải pháp mới.  Chương 3 em sẽ trình bày cơ sở lý thuyết cần thiết để xây dựng mô hình và giải quyết vấn đề. Bao gồm: phân phối Poisson và phân phối Poisson đa chiều được sử dụng để mô hình lý thuyết của bài toán, các phương pháp tính tích phân gần đúng để thực hiện khâu tính toán công thức, và các giải thuật tiến hóa tối ưu là giải thuật bầy ong nhân tạo và giải thuật bầy đàn.  Chương 4 sẽ phân tích mô hình của hệ quan trắc môi trường di động, đồng thời nêu ra các hướng tiếp cận giải quyết bài toán và cuối cùng là các phương pháp tối ưu hiệu quả trong việc giải bài toán đề ra.  Chương 5 em sẽ phân tích các tham số đầu vào của bài toán và nêu ra kết quả thực nghiệm.  Chương 6 từ kết quả thực nghiệm ở chương 5, em rút ra kết luận cho mô hình bài toán đang xét và đề xuất một số hướng nghiên cứu liên quan để mở rộng bài toán giải quyết các vấn đề còn tồn đọng. Chương 2: Các kiến thức nền tảng2.1 Phân phối poisson và phân phối poisson đa chiều.2.2 Các phương pháp tính gần đúng tích phân. Trong nhiều bài toán trên thực tế và lý thuyết, ta cần phải tính toán tích phân xác định của một hàm toán học nào đó. Hàm toán học đó có thể rất đơn giản, cũng có thể rất phức tạp, có thể là hàm đơn biến hoặc một hàm đa biến. Trong phần này em xin tập trung vào việc trình bày phương pháp tính gần đúng tích phân của hàm đơn biến, việc mở rộng sang hàm đa biến trong nhiều trường hợp là hoàn toàn tương tự.  Giả sử ta cần tính  Nếu hàm f(x) liên tục trên [a,b] và có nguyên hàm F(x), theo công thức Newton-Lepnit ta có:  Trong trường hợp f(x) cho ở dạng bảng hoặc f(x) đã biết nhưng không xác định được nguyên hàm F(x) ví dụ như hàm f(x) = , thì thay vì việc tính chính xác ta chỉ có thể tính gần đúng tích phân xác định của f(x).  Trước hết ta phân hoạch đoạn [a,b] thành n phần nhỏ hữu hạn với (i = 1,2,…,n) bởi những điểm a = < < ….. < .  Ta có tổng Riemann S của hàm f trên đoạn [a,b] với phép phân chia trên là :  S = với .  Khi ta chia đoạn [a,b] thành các đoạn con đủ nhỏ thì tổng S trên được coi là kết quả gần đúng của phép tính tích phân xác định  2.2.1 Phương pháp sử dụng tổng Reimann  Việc chọn là hoàn toàn tùy ý trong khoảng trên. Tuy nhiên có một số trường hợp đặc biệt như sau :   * thì S được gọi là tổng Riemann trái. * thì S được gọi là tổng Riemann phải. * thì S được gọi là tổng Riemann giữa.     Từ tổng Reimann ta có phương pháp gần đúng để tính tích phân xác định như sau:  Chia đoạn [a,b] thành các đoạn con bằng nhau và bằng .  Khi đó ta có tổng Reimann trái .  Tổng Reimann phải .  Tổng Reimann giữa .  Để tính được các tổng Reimann trên bằng thuật toán thì đơn giản ta chỉ cần sử dụng một vòng for là đủ.  Nếu hàm f(x) có đạo hàm bậc 1 xác định trên đoạn [a,b], người ta chứng minh được rằng:  Với  Do đó với n càng lớn thì phép tính tích phân bằng tổng Reimann càng cho độ chính xác cao.  Tương tự ta cũng có một bất đẳng thức sai số của phép lấy tổng giữa trong trường hợp f(x) có đạo hàm bậc 2 trên đoạn [a,b]:  Với  Với việc đánh giá được chặn trên của sai số, hiện nay có một số phương pháp tính tích phân gần đúng với cơ chế “thích nghi” (adaptive), với những đoạn có ước lượng sai số lớn thì người ta chia làm nhiều khoảng hơn, với những đoạn có chặn trên sai số nhỏ người ta dung ít khoảng hơn để xấp xỉ. Việc làm như vậy có ưu điểm vẫn đảm bảo được sai số nhỏ nhưng độ phức tạp thuật toán có sự cải tiến, tuy nhiên giải thuật này chỉ áp dụng được với hàm có thể tính được đạo hàm hay đạo hàm bậc cao một cách dễ dàng, hơn nữa các chặn trên của sai số có thể có sai khác tương đối với sai số thực tế.  2.2.2 Phương pháp sử dụng luật hình thang.  Trong khi các tổng Reimann xấp xỉ tích phân trên mỗi đoạn con bằng diện tích của hình chữ nhật thì trong phương pháp sử dụng luật hình thang, người ta sử dụng diện tích hình thang để ước lượng xấp xỉ giá trị của tích phân xác định đó.    Tích phân trên mỗi đoạn được tính bởi công thức: .  Lấy tổng trên tất cả các đoạn ta có:  2.2.3 Phương pháp Simpson  Mở rộng từ phương pháp sử dụng luật hình thang. Thực chất việc sử dụng luật hình thang chính là việc xấp xỉ một hàm bậc 1 đi qua 2 điểm có tọa độ và . Mở rộng ý tưởng này với 3 điểm hay nhiều điểm hơn ta được phương pháp Simpson.  Xét trường hợp 3 điểm, chỉ tồn tại duy nhất một đường parabol đi qua 3 điểm ,. Giả sử phương trình của parabol là . Giải phương trình này người ta xác định được  Lấy tổng trên tất cả các đoạn ta có:  Tổng kết lại, hiện nay có rất nhiều phương pháp khác nhau để sử dụng trong việc tính xấp xỉ tích phân xác định của một hàm số. Mỗi phương pháp đều có ưu, nhược điểm khác nhau và các thông số cần tinh chỉnh sao cho thời gian chạy thuật toán nhanh, sai số nhỏ. Trên thực tế việc chọn các thông số và thuật toán phù hợp phụ thuộc vào bài toán cần giải và kiểm nghiệm thực tế của người lập trình. Tổng quan các giải thuật tối ưu Chủ đề tối ưu hóa luôn luôn thu hút được sự quan tâm, chú ý của không chỉ các nhà khoa học mà cả các nhà quản lý, nhà kinh tế,… Đặc biệt trong bối cảnh cạnh tranh gay gắn trên mọi mặt của cuộc sống, cá nhân hay tổ chức nào đưa ra được mô hình tối ưu hơn, tiết kiệm chi phí hơn sẽ chiếm lợi thế lớn trên lĩnh vực đó. Chính vì vậy người ta phải quan tâm tìm ra phương án tối ưu nhất để đạt được mục tiêu mong muốn, trong những điều kiện rằng buộc nhất định. Đó gọi là bài toán tối ưu.  Ngày nay với sự phát triển của khoa học, kỹ thuật và công nghệ. Các phương pháp tối ưu ngày càng phát triển và được ứng dụng rộng rãi trong mọi lĩnh vực.  Thông thường các bài toán tối ưu được phát biểu như sau:  Cực tiểu hóa một hàm là một quy tắc ứng mỗi điểm  Ở đây trong hầu hết các trường hợp , X được gọi là miền xác định của x.  Thông thường miền xác định( hay miền đang xét) X được xác định bằng một tập hợp các hàm rằng buộc.  Hầu hết mỗi phương pháp tối ưu chỉ phù hợp với một hoặc một lớp các bài toán nhất định. Vì vậy người ta chia các phương pháp tối ưu thành 2 nhóm chính là quy hoạch tuyến tính và quy hoạch phi tuyến. Trong phần này em tập trung giới thiệu nhóm các phương pháp quy hoạch phi tuyến.  Đơn giản nhất là phương pháp chia lưới trên hệ trục tọa độ (grid search method), ở phương pháp này miền X được chia thành các lưới ô vuông và dùng giải thuật vét cạn trên các giao điểm của các lưới ô vuông đó. Ưu điểm của giải thuật này là dễ hiểu, dễ cài đặt tuy nhiên độ chính xác của kết quả phụ thuộc vào việc chia miền X thành các vùng càng nhỏ càng tốt. Khi số chiều n của X tăng lên cao thì giải thuật có độ phức tạp thời gian lớn nên chạy chậm. Tương tự với phương pháp này là phương pháp tìm kiếm ngẫu nhiên, thay vì việc chia lưới như trên thì các điểm được tính toán là tuân theo một phân phối xác suất ngẫu nhiên nào đó.  Bên cạnh hai phương pháp đơn giản nêu trên, thì các thuật toán sử dụng gradient cũng tương đối phổ biển, ví dụ như phương pháp gradient-descent hay phương pháp Newton. Điểm yếu của các phương pháp dựa vào gradient là yêu cầu hàm tối ưu phải có đạo hàm hay thậm chí là có đạo hàm bậc k. Trong thực tế nhiều bài toán dẫn đến hàm tối ưu không có đạo hàm, thậm chí là không liên tục, hoặc việc tính đạo hàm quá phức tạp.  Khắc phục được các nhược điểm của các thuật toán nêu trên, thuật toán tối ưu bầy đàn là lớp các giải thuật dựa trên hành vi, tập tính di cư hay tìm kiếm nguồn thức ăn của các bầy đàn trong tự nhiên. Các thuật toán này được gọi là metaheuristic, với khả năng tìm kiếm tối ưu cho hàm đa biến, có rằng buộc, không yêu cầu các hàm này phải liên tục hay có đạo hàm. Tất nhiên, vì là thuật toán dạng heuristic nên các phương pháp tối ưu bầy đàn không đảm bảo tìm kiếm được lời giải tối ưu. Việc áp dụng thuật toán phù hợp với mỗi bài toán đòi hỏi quá trình thử nghiệm và kinh nghiệm của mỗi người. Với bài toán của mình, được sự gợi ý của giảng viên hướng dẫn, cùng với quá trình tự tìm tòi, học hỏi, em xin trình bày hai giải thuật được em sử dụng để thử nghiệm trong mô hình của mình. Giải thuật bầy ong nhân tạo( artificial bee colony). Trí thông minh bầy đàn là một nhánh nghiên cứu mô phỏng một tập hợp các cá thể có khả năng tự khám phá. Sự tương tác giữa các cá thể của một bầy kiến, hay một đàn chim hay của hệ miễn dịch trong cơ thể đều là ví dụ tiêu biểu cho một hệ cá thể thông minh. 2.3.2 Giải thuật PSO .Chương 3: Giải thuật tối ưu cho bài toán đo chất lượng không khí bằng cảm biến tích hợp trên xe.3.1Mô tả bài toán3.1.1 Mô hình.3.1.2 Xác định mục tiêu và hàm tối ưu.3.2 Tổng quan các hương tiếp cận giải quyết bài toán.3.3 Phương pháp tối ưu sử dụng giải thuật bầy ong nhân tạo3.4 Phương pháp tối ưu sử dụng thuật toán PSO.Chương 4: Đánh giá thực nghiệm.4.1 Các tham số đánh giá.4.2 Phương pháp thí nghiệm.4.3 Kết quả thí nghiệm.4.3.1 Tham số tối ưu cho phương pháp tính gần đúng tích phân, so sánh với phương pháp monte - carlo.4.3.2 So sánh kết quả thuật toán ABC với PSO.4.3.3 Ảnh hưởng của số làn đường và khoảng cách giữa các làn đến kết quả tối ưu.Chương 5: Kết luận5.1 Kết luận5.2 Định hướng phát triển của đề tài | | |
|  | | |