

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

NGUYỄN TRUNG CHIẾN

**ỨNG DỤNG THUẬT TOÁN ĐÀN KIẾN
TRONG TÌM KIẾM ĐƯỜNG ĐI TỐI ƯU**

LUẬN VĂN THẠC SĨ: CÔNG NGHỆ THÔNG TIN

Thái Nguyên – 2014

MỞ ĐẦU

Trí tuệ nhân tạo là một trong lĩnh vực được quan tâm nhiều trong công nghệ thông tin hiện nay. Trong trí tuệ nhân tạo, chúng ta thường xuyên phải đối mặt với những bài toán tìm kiếm. Đặc biệt là những bài toán lập lịch và học máy, tìm kiếm đóng vai trò hết sức quan trọng.

Vấn đề tìm kiếm có thể hiểu là tìm ra một kết quả thỏa mãn điều kiện được đặt ra trước đó trong một tập hợp lớn các đối tượng. Có rất nhiều vấn đề có thể quy ra bài toán tìm kiếm, ví dụ như trò chơi: cờ vua, cờ caro có thể xem như bài toán tìm kiếm – từ tập hợp các nước có thể đi ta chọn ra tập hợp các nước đi ngắn nhất để trở thành người thắng. Hay như bài toán tháp Rùa – Hồ Gươm cũng có thể quy ra bài toán tìm kiếm – từ tập hợp tất cả các cách chuyển tháp từ A đến C ta chọn ra tập hợp các bước chuyển ít nhất, ...

Ngày nay, với sự đòi hỏi cao về khoa học và công nghệ, các kỹ thuật tìm kiếm cổ điển đã không còn phù hợp mà thay vào đó là cách tìm kiếm không rõ đối tượng (tìm kiếm mù), các kỹ thuật tìm kiếm kinh nghiệm (heuristic), các kỹ thuật tìm kiếm tối ưu, ...

Một trong số những thuật toán tìm kiếm dựa trên kinh nghiệm khá hiệu quả hiện nay là thuật toán tối ưu đường đi của loài kiến (do nhà khoa học người Bỉ Marco Dorigo giới thiệu trong luận án tiến sĩ của mình năm 1996). Thuật toán này sử dụng giải pháp Meta-heuristic, là một tập các khái niệm về thuật toán được sử dụng để xác định các phương thức tìm kiếm thích hợp cho một tập các vấn đề khác nhau, có thể coi là một phương thức tìm kiếm đa năng. Nó giúp tối ưu hóa phương pháp giải các bài toán NP-Khó. Hiện nay, các thuật toán kiến đã được ứng dụng vào thực tế ở nhiều lĩnh vực khác nhau như: Áp dụng vào việc kinh doanh của nhiều hãng vận tải lớn tại Mỹ, ứng

dụng trong ngành bưu chính tại Đan Mạch, tìm kiếm thông tin trên mạng internet, ...

Trong giới hạn về đề tài này, dưới sự hướng dẫn của PSG, TS Đoàn Văn Ban, em mạnh dạn chọn đề tài: “**ỨNG DỤNG THUẬT TOÁN ĐÀN KIẾN TRONG TÌM KIẾM ĐƯỜNG ĐI TỐI ƯU**” để tìm hiểu và thực hiện.

Cấu trúc của luận văn được chia làm ba chương, với nội dung chính của mỗi chương như sau:

- **Chương 1:** Giới thiệu về hệ thống đàn kiến, phương pháp giải heuristic, một số thí nghiệm liên quan và một số thuật toán đàn kiến.
- **Chương 2:** Tìm hiểu một số vấn đề liên quan đến các kỹ thuật tìm kiếm tối ưu như : thuật toán A*, thuật toán nhánh và cận, thuật toán leo đồi và một số bài toán tìm kiếm dựa trên kinh nghiệm.
- **Chương 3:** Phát biểu và mô tả bài toán tìm đường D-TSP. Xây dựng hướng giải quyết bài toán và demo.

Ngoài ra, trong luận văn còn có phần “Mở đầu” và phần “Kết luận”.

CHƯƠNG 1 : HỆ THỐNG ĐÀN KIẾN

1.1. Tổng quan

Từ xa xưa, thiên nhiên đã là nguồn cảm hứng vô tận, nó không chỉ là nguồn cảm hứng cho những thi sĩ, nghệ sĩ - những người làm văn hóa nghệ thuật mà nó còn là nguồn cảm hứng cho những nhà khoa học. Từ những sự kiện trong tự nhiên rất đời thường khi đi vào khoa học kỹ thuật nó đều trở thành những phát minh vĩ đại: Từ một quả táo rơi, với Isaac Newton ta có định luật vạn vật hấp dẫn; Từ những cánh chim và ước mơ được bay lên của hai anh em nhà Wright để ngày nay chúng ta có những chiếc máy bay tối tân hiện đại; Và còn rất nhiều những phát minh khác có nguồn gốc từ thiên nhiên như: áo chống đạn dựa trên cách giăng tơ của loài nhện, cảm biến dựa trên bộ râu của loài gặm nhấm...

Dựa trên các yếu tố về mặt tự nhiên, các nhà khoa học mô phỏng lại, cải biến, hoàn thiện và đưa nó thành những sản phẩm nhằm phục vụ mục đích của họ. Ngoài những thiết bị vật lý mà chúng ta nhìn thấy từ sản phẩm thì nằm sâu trong chúng là những bài toán nhằm giải quyết hay mô phỏng sao cho gần với tự nhiên nhất.

Ngày nay, trí tuệ nhân tạo được sử dụng nhiều trong các ngành khoa học kỹ thuật. Phương pháp tìm kiếm bầy đàn được áp dụng khá rộng rãi, cụ thể một trong những thuật toán đó là thuật toán đàn kiến (do Marco Dorigo giới thiệu vào năm 1992). Kể từ đó tới nay, thuật toán đàn kiến đã có rất nhiều cải tiến và được ứng dụng vào nhiều lĩnh vực như: trí tuệ nhân tạo, trong bộ máy tìm kiếm, tin sinh, ... [12].

Dưới đây là một số thuật toán ACO theo trình tự về thời gian xuất hiện:

| Thuật toán ACO | Tác giả |
|-----------------------|--|
| Ant System | Dorigo, Maniezzo & Colomi (1991) |
| Elitist AS | Dorigo (1992); Dorigo, Maniezzo & Colomi (1996) |
| Ant-Q | Gambardella & Dorigo (1995); Dorigo & Gambardella (1996) |
| Ant Colony System | Dorigo & Gambardella (1996) |
| Max-Min AS | Stutzle & Hoos (1996,2000); Stutzle (1999) |
| Rank-based AS | Bullnheiner, Hartl & Strauss (1997,1999) |
| ANTS | Maniezzo (1999) |
| Hyper-cube AS | Blum, Roli & Dorigo (2001); Blum & Dorigo(2004) |

Bảng 1.1. Một số thuật toán ACO

1.2. Hành vi của đàn kiến trong tự nhiên

Trong thế giới tự nhiên, cách tìm môi của đàn kiến bắt đầu bằng việc đi lang thang ngẫu nhiên và trong quá trình tìm kiếm đó chúng lưu lại trên con đường mà chúng đi qua một lượng Pheromone. Hành vi đi ngẫu nhiên này có thể sẽ không được lặp lại với các con kiến đi sau mà thay vào đó là sự chọn lựa các vết pheromone do các con kiến đi trước tạo ra để quay trở lại tổ hoặc củng cố lại con đường đó nếu nguồn thức ăn được tìm thấy.

Tuy nhiên, theo thời gian các vết pheromone sẽ bay hơi và làm giảm sự hấp dẫn với các con kiến khác. Thời gian tiêu hao của một con kiến đi từ tổ của nó đến nguồn thức ăn và quay lại sẽ tương ứng với lượng pheromone bị bay hơi. Từ đó đường đi tối ưu được hình thành nhờ mật độ kiến đi qua nhiều nhất và lượng pheromone để lại là đậm nhất.

Sự bay hơi của pheromone là lợi thế để tránh sự hội tụ dẫn đến một giải pháp tối ưu hóa cục bộ. Nếu không có sự bay hơi của pheromone, các con đường đã được lựa chọn bởi những con kiến đi đầu sẽ không quá khác biệt và

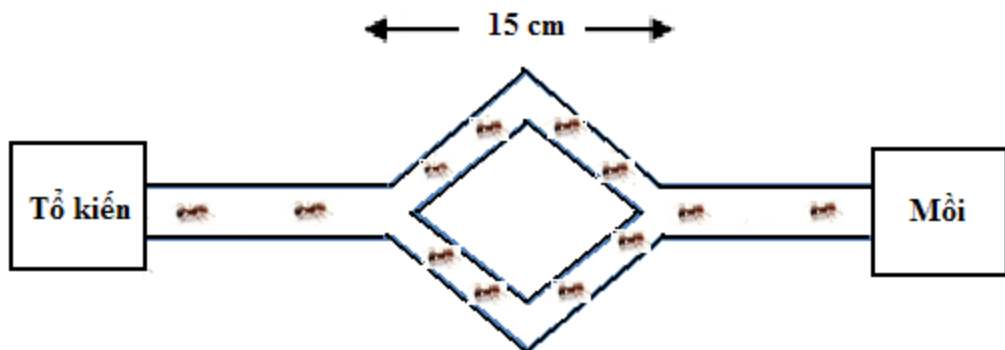
khó xảy ra hội tụ. Trong trường hợp đó các giải pháp thăm dò sẽ hết sức hạn chế.

Vì vậy, khi một con kiến tìm thấy đường đi tốt từ tổ của chúng đến nguồn thức ăn, những con kiến khác có nhiều khả năng cũng theo con đường đó và có những phản hồi tích cực dẫn đến việc cả đàn kiến sẽ theo một con đường duy nhất. Ý tưởng của thuật toán đàn kiến là mô phỏng lại quá trình tìm đường của đàn kiến ngoài tự nhiên để duyệt đồ thị đại diện cho các bài toán cần giải quyết.

1.2.1. Thí nghiệm chiếc cầu đôi

Một thí nghiệm được nhắc đến trong nhiều tài liệu về đường đi của đàn kiến là thí nghiệm trên một chiếc cầu đôi của Deneubourg và các đồng sự [5]. Ông đã sử dụng một cây cầu có 2 nhánh nối tổ kiến tới chỗ thức ăn. Hai nhánh của cây cầu gọi là nhánh ngắn và nhánh dài, tỉ lệ $r = l_l/l_s$ trong đó l_l là độ dài của nhánh dài và l_s là độ dài nhánh ngắn. Tỉ lệ r được thay đổi nhiều lần trong các thí nghiệm nhằm quan sát sự thay đổi trong hành vi của các con kiến.

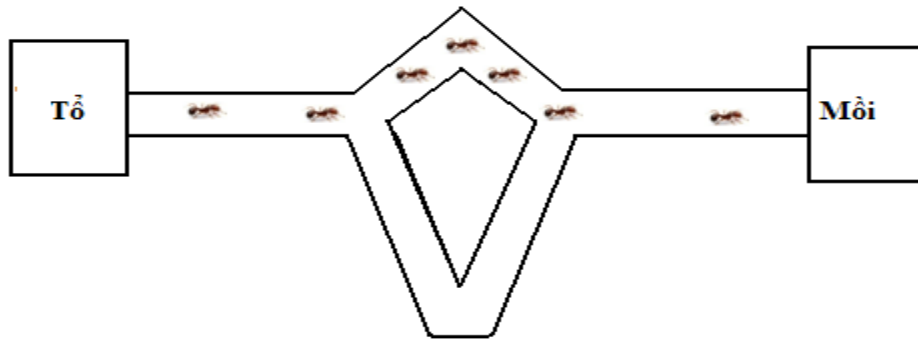
Đầu tiên đàn kiến được thử nghiệm trên chiếc cầu với $r = 1$. Kết quả sau một thời gian quan sát : đàn kiến vẫn chọn chung một con đường mặc dù có suất phát là ngẫu nhiên.



Hình 1.1. Thí nghiệm chiếc cầu đôi với 2 nhánh có kích thước bằng nhau

Khi thí nghiệm được bắt đầu, trên chiếc cầu hoàn toàn không hề có vết pheromone, hai nhánh cầu bằng nhau nhưng đàn kiến vẫn chọn một con đường. Điều này được lý giải là do sự ngẫu nhiên của xác suất (xấp xỉ 50% trên 1 nhánh cầu) dẫn đến lượng pheromone ở nhánh này nhiều hơn nhánh kia. Tức là xác suất mà con kiến sau chọn đi vào đường có nồng độ pheromone cao hơn là lớn hơn.

Để minh chứng rõ hơn, trong lần thí nghiệm sau đó, cây cầu được thiết kế với $r = 2$ tương đương với $l_l = 2l_s$.

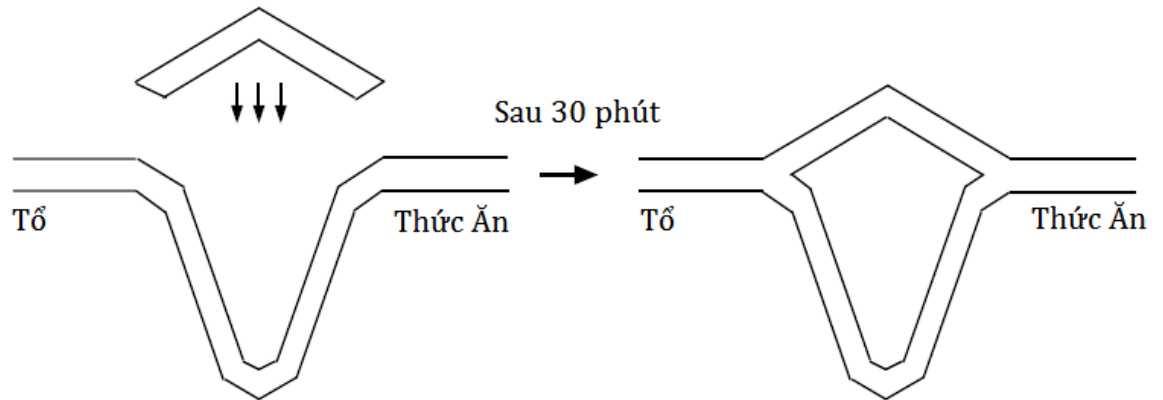


Hình 1.2. Thí nghiệm chiếc cầu đôi với 2 nhánh có kích thước không bằng nhau

Sau nhiều lần thử nghiệm, đàn kiến vẫn chọn nhánh cầu ngắn hơn. Một con kiến xuất phát từ tổ đến mồi và khi tha mồi về tổ nó vẫn phải chọn con đường có pheromone cao hơn. Chính lý do này nồng độ pheromone trên con đường ngắn hơn ngày càng cao hơn và những con kiến sau sẽ có xác suất chọn vào đường đó cao hơn. Tuy nhiên, cũng không hẳn toàn bộ đàn kiến đều đi vào đường ngắn mà vẫn có những con kiến đi vào đường khác, các nhà khoa học cho rằng đó có thể là một dạng khám phá đường đi mới.

Khi đàn kiến hội tụ vào một đường, tức là chúng đã tạo ra một kết nối giữa tổ và mồi thông qua vết pheromone. Một thí nghiệm mở rộng được đưa ra để kiểm nghiệm việc tạo ra vết pheromone và sự bay hơi của chúng: Người ta đưa một cây cầu chỉ có nhánh dài và nhánh ngắn được tháo ra, cây cầu được

nổi từ tổ kiến đến mối. Sau khi đàn kiến đã di chuyển qua một khoảng thời gian, người ta lắp thêm nhánh cầu ngắn nhưng việc chọn lựa con đường này của kiến là rất thấp.



Hình 1.3. Thí nghiệm chiếu cầu đôi mở rộng

Khi khởi tạo thí nghiệm, chiếc cầu chỉ có 1 nhánh dài, sau 30 phút, một tình huống mới xuất hiện, một nhánh ngắn hơn được nối vào trên chiếc cầu.

Điều này có thể được giải thích bởi nồng độ chất pheromones ở nhánh dài cao và làm chậm tốc độ bay hơi của chất pheromones. Thực tế, đa số các con kiến đều chọn nhánh dài vì tại nhánh này nồng độ pheromones cao, và hành vi “tự xúc tác” của chúng tiếp tục được củng cố trên nhánh dài, ngay cả khi có sự xuất hiện của nhánh ngắn.

Quá trình bốc hơi của pheromones rất có lợi cho quá trình thăm dò, khám phá ra những con đường mới, tuy nhiên quá trình bốc hơi này diễn ra khá chậm: thời gian tồn tại của pheromones được so sánh với thời gian tiến hành cuộc thử nghiệm, có nghĩa là chất pheromones bốc hơi quá chậm để cho các con kiến có thể “quên” đi con đường không tối ưu mà chúng đã hội tụ để có thể khám phá ra con đường mới ngắn hơn và có thể “học được”.

1.2.2. Mô hình ngẫu nhiên

Deneubourg và các đồng nghiệp (Deneubourg et al., 1990, Goss et al., 1989) đã đưa ra một mô hình xác suất thống kê mô tả chức năng của bầy kiến mà họ đã quan sát được qua thí nghiệm chiếc cầu đôi [1][5][7]. Trong mô hình này, cứ 1 giây sẽ có ψ con kiến băng qua cầu (mỗi hướng có ψ con kiến) với tốc độ không đổi là v cm/s và mỗi con sẽ tiết ra một lượng *mùi* (pheromone) là 1 đơn vị. Cho biết trước độ dài của nhánh dài l_l và nhánh ngắn l_s (đơn vị là cm) của cây cầu. một con kiến chọn nhánh ngắn sẽ di chuyển trên nhánh này với thời gian $t_s = l_l/v$ (giây) trong đó một con kiến đi nhánh dài sẽ tốn lượng thời gian là $r * t_s$, với $r = l_l/l_s$.

Ta định nghĩa xác suất $p_{ia}(t)$ là xác suất khi con kiến tới điểm cần phải đưa ra quyết định $i \in \{1, 2\}$ nó sẽ chọn nhánh $a \in \{s, l\}$, trong đó s và l là chỉ nhánh ngắn và nhánh dài trên cầu, tại thời điểm t . Xác suất này là hàm của của tổng lượng *mùi* (pheromone) $\varphi_{ia}(t)$, mà hàm này lại tỉ lệ với số kiến từng đi qua nhánh a này cho tới thời điểm t .

Ví dụ, ta có thể xây dựng 1 hàm $p_{is}(t)$ như sau:

$$p_{is}(t) = \frac{(t_s + \varphi_{is}(t))^\alpha}{(t_s + \varphi_{is}(t))^\alpha + (t_s + \varphi_{il}(t))^\alpha} \quad (1.1)$$

Hàm này cùng với giá trị $\alpha = 2$ được rút ra từ thí nghiệm lần đầu (Deneubourg et al., 1990), tương tự với $p_{il}(t)$, ta có $p_{is}(t) + p_{il}(t) = 1$.

Mô hình này dựa trên giả thiết rằng lượng *mùi* (pheromone) trên một nhánh tỉ lệ với số lượng kiến đã từng đi qua nhánh đó trong quá khứ. Nói một cách khác, ở đây không có sự xảy ra sự bay hơi *mùi* (điều này tương ứng với các quan sát từ thí nghiệm rằng thời gian để cho các con kiến hội tụ về đường đi ngắn nhất cũng tương đương với thời gian bay hơi của *mùi* (Goss et al., 1989; Beckers, Deneubourg, & Goss, 1993)) [15].

Các phương trình vi phân mô tả quá trình phát triển của mô hình xác suất thống kê này là như sau:

$$\frac{d\varphi_{is}}{dt} = \psi p_{js}(t - t_s) + \psi p_{is}(t), (i = 1, j = 2; i = 2, j = 1) \quad (1.2)$$

$$\frac{d\varphi_{il}}{dt} = \psi p_{jl}(t - r \cdot t_s) + \psi p_{il}(t), (i = 1, j = 2; i = 2, j = 1) \quad (1.3)$$

Phương trình (1.2) có thể được giải thích như sau: tại thời điểm t , sự thay đổi lượng mùi (pheromone) trên nhánh s tại điểm quyết định i được tính bởi lượng kiến đi qua ψ (giả sử là hằng số), nhân với xác suất chọn nhánh ngắn hơn tại điểm quyết định j tại thời điểm $t - t_s$ và cộng với ψ nhân với xác suất lựa chọn nhánh ngắn hơn tại điểm quyết định i tại thời điểm t . Hằng số t_s thể hiện khoảng thời gian trễ cần thiết để một con kiến đi qua nhánh ngắn.

Phương trình (1.3) diễn giải điều tương tự với nhánh dài, ngoại trừ thời gian trễ là $r \cdot t_s$.

Hệ động xác định nhờ vào các phương trình trên được mô phỏng nhờ phương pháp Monte Carlo (Liu, 2001) [16]. Trong hình 8 là kết quả của 2 thí nghiệm với 1000 lần thử cho mỗi thí nghiệm và trong đó tỉ lệ độ dài giữa 2 nhánh được đặt bằng $r = 1$ và $r = 2$. Chúng ta có thể thấy khi 2 nhánh có cùng độ dài ($r = 1$) các con kiến sẽ hội tụ về việc chỉ sử dụng một trong 2 nhánh với xác suất là tương đương nhau qua cả 1000 thí nghiệm. Ngược lại, khi một nhánh có độ dài gấp 2 lần nhánh còn lại thì trong hầu hết thí nghiệm, tất cả các con kiến sẽ chọn nhánh ngắn hơn (Goss et al., 1989).

Trong mô hình này các con kiến tiết ra mùi cả trong lúc đi tìm đường lẫn lúc quay về tổ. Điều này đã cho thấy đây là một hành vi cần thiết để đàn kiến có thể hội tụ về đường đi ngắn hơn. Trên thực tế, nếu ta xem xét mô hình mà trong đó kiến chỉ tiết mùi lúc tìm đường hoặc trên đường quay về tổ thì kết quả sẽ là đàn kiến không thể chọn được ra được đường đi ngắn nhất.