

BỘ GIÁO DỤC VÀ ĐÀO TẠO
ĐẠI HỌC ĐÀ NẴNG

HUỲNH NGUYỄN NGỌC THẢO

**XÂY DỰNG CHƯƠNG TRÌNH TỐI ƯU HÓA
QUÁ TRÌNH ĐỊNH TUYỂN TRÊN MẠNG IP
DỰA VÀO GIẢI THUẬT DI TRUYỀN**

Chuyên ngành: KHOA HỌC MÁY TÍNH

Mã số: 60.48.01

TÓM TẮT LUẬN VĂN THẠC SĨ KỸ THUẬT

Đà Nẵng - Năm 2011

Công trình được hoàn thành tại
ĐẠI HỌC ĐÀ NẴNG

Người hướng dẫn khoa học: **PGS.TSKH. Trần Quốc Chiến**

Phản biện 1: **PGS.TS. Phan Huy Khánh**

Phản biện 2: **PGS.TS. Đoàn Văn Ban**

Luận văn được bảo vệ tại Hội đồng chấm Luận văn tốt nghiệp
thạc sĩ kỹ thuật họp tại Đại học Đà Nẵng vào ngày 11 tháng 09
năm 2011

Có thể tìm hiểu luận văn tại:

- Trung tâm Thông tin - Học liệu, Đại học Đà Nẵng
- Trung tâm Học liệu, Đại học Đà Nẵng

MỞ ĐẦU

1. Lý do chọn đề tài

Khi từ một máy tính này đến một máy tính khác và từ mạng này đến một mạng khác cũng ngày càng trở nên phức tạp. Nhằm đáp ứng các nhu cầu hết sức đa dạng và phong phú của người sử dụng Internet, các nhà c Internet ngày càng phát triển thì vấn đề định tuyến lưu lượng ung cấp dịch vụ mạng cần sử dụng một cách hiệu quả cơ sở hạ tầng mạng của mình để có thể đưa ra các dịch vụ với chất lượng cao mà không cần phải nâng cấp thiết bị phần cứng nhằm giảm thiểu chi phí, vì vậy cần phải có giải pháp nhằm khai thác một cách tốt nhất thiết bị hạ tầng mạng, mà điển hình là việc cho phép cân bằng tải trọng trên tất cả các kết nối, tránh tình trạng đường truyền này thì rảnh trong khi các đường truyền khác lại bị tắc nghẽn điều này đòi hỏi cần phải có các nghi thức truyền thông hoạt động một cách hiệu quả, đây chính là mục tiêu của việc *tối ưu hóa quá trình định tuyến trên mạng*. Đó là lý do tôi chọn đề tài: ***“Xây dựng chương trình tối ưu hóa quá trình định tuyến trên mạng IP dựa vào giải thuật di truyền”***.

2. Mục đích nghiên cứu

Việc nghiên cứu và đề xuất đề tài này để giảm thiểu chi phí và nâng cao chất lượng dịch vụ Internet trong hoàn cảnh ngày càng có nhiều dịch vụ mới ra đời chiếm dụng băng thông lớn như hiện nay.

3. Đối tượng và phạm vi nghiên cứu

Sự phát triển của Internet cũng đồng nghĩa với việc tăng trưởng về quy mô và công nghệ nhiều loại mạng LAN, WAN... và

đặc biệt là lưu lượng thông tin trên mạng tăng đáng kể. Chính điều đó đã làm cho vấn đề điều phối luồng thông tin trên mạng hay là vấn đề định tuyến trở nên quan trọng hơn bao giờ hết. Trong việc thiết kế mạng việc lựa chọn giao thức định tuyến sao cho phù hợp với chi phí, tài nguyên của tổ chức là đặc biệt quan trọng.

Internet phát triển càng mạnh, lượng người truy nhập càng tăng yêu cầu định tuyến càng phải tin cậy, tốc độ chuyển mạch nhanh và không gây ra lặp và mất dữ liệu trên mạng. Hơn nữa khi nhiều tổ chức tham gia vào mạng thì nhiều giao thức được đưa vào sử dụng dẫn đến sự phức tạp về định tuyến cũng gia tăng, và số lượng các giao thức để phục vụ cho việc định tuyến cũng có rất nhiều. Việc hiểu biết và thiết kế các mạng thông tin cỡ lớn có sử dụng các thiết bị định tuyến đang trở thành một nhu cầu vô cùng cấp thiết trong thực tế. Nó đòi hỏi người thiết kế mạng phải có sự hiểu biết sâu về giao thức sẽ sử dụng cho việc thiết kế mạng cũng như các loại giao thức định tuyến khác.

4. Phương pháp nghiên cứu

Phương pháp nghiên cứu tư liệu: Các tài liệu mà giáo viên hướng dẫn đưa, trên các trang web và các bài báo khoa học gần đây có liên quan đến đề tài, sách giáo trình, tài liệu tham khảo, tạp chí khoa học, các đề tài nghiên cứu có liên quan...

5. Ý nghĩa khoa học

Quá trình tối ưu cần đưa ra một giải pháp nhằm đáp ứng tốt nhất nhu cầu truyền thông là quá trình chọn lọc các tuyến đường nhằm tìm ra một bộ trọng số thích hợp. Đối với những mạng nhỏ chúng ta có thể giải được bài toán một cách nhanh chóng bằng

phương pháp quy hoạch tuyến tính, tuy nhiên, đối với các mạng lớn (số lượng bộ định tuyến lên đến hàng nghìn, hàng vạn) thì phương pháp quy hoạch tuyến tính hầu như không khả thi. Vì vậy, đề tài này sẽ nêu ra một phương pháp mới áp dụng giải thuật di truyền (Genetic Algorithm -GA) để tìm ra một bộ các trọng số tối ưu sẽ được gán vào cho các thiết bị định tuyến.

6. Cấu trúc luận văn

Với định hướng như trên, ngoài phần **Mở đầu** và phần **Kết luận và hướng phát triển**, luận văn gồm 3 chương:

Chương 1: Tổng quan về định tuyến mạng

Chương 2: Bài toán ước lượng nhu cầu truyền thông

Chương 3: Tối ưu hóa quá trình định tuyến mạng bằng giải thuật di truyền.

CHƯƠNG 1

TỔNG QUAN VỀ ĐỊNH TUYẾN MẠNG

1.1. CÁC KHÁI NIỆM CƠ BẢN

1.1.1. Hệ thống tự trị (Autonomous System - AS)

Internet toàn cầu được tổ chức theo từng khối riêng lẻ, hoạt động tương đối độc lập với nhau gọi là những hệ thống tự trị (*Autonomous System - AS*). Mỗi AS bao gồm một hoặc một số nhóm các địa chỉ IP, mỗi nhóm tượng trưng cho một mạng con, sử dụng chung một chính sách cụ thể, rõ ràng cho việc định tuyến và cũng hoạt động dưới sự kiểm soát của một tập thể các chuyên gia điều hành mạng.

Công việc thiết kế mạng thường được giới hạn trong phạm vi một hệ thống tự trị AS. Có thể nhận thấy rõ ràng, nếu các AS cũng được xây dựng và vận hành theo một phương thức thì toàn bộ hệ thống Internet sẽ hoạt động tốt hơn với chi phí thấp và hiệu suất cao, tuy nhiên thực tế lại hoàn toàn không phải như vậy, mỗi AS hay ngay cả trong từng mạng con của một AS cũng thường được xây dựng và vận hành một cách độc lập với nhau, chỉ những nơi nào cần thiết mới có những hợp đồng, thỏa thuận nhằm đảm bảo chúng kết nối được với nhau.

1.1.2. Đường đi tối ưu cơ bản

1.2. TỔNG QUAN VỀ ĐỊNH TUYẾN MẠNG

1.2.1. Khái niệm định tuyến mạng

Định tuyến (*routing*) là quá trình chọn lựa các đường đi trên một mạng máy tính để gửi dữ liệu qua đó.

Định tuyến mạng được dùng trong lĩnh vực truyền thông Internet để chỉ quá trình gồm hai thao tác: xác định đường truyền (*path determination*) và chuyển gói tin (*forwarding*) theo con đường đã chọn, hai thao tác này luôn luôn đi kèm với nhau trong thiết bị phân cứng chuyên dùng được gọi là bộ định tuyến (*router*).

Trong các mạng thông tin khác nhau, việc xác định đường truyền cũng diễn ra khác nhau. Tuy nhiên, cách xác định đường truyền nào cũng gồm hai công việc cơ bản :

Thứ nhất, là thu thập và phân phát thông tin về tình trạng của mạng (như trạng thái đường truyền, tình trạng tắc nghẽn...) và của thông tin cần truyền (như lưu lượng, băng thông, tốc độ, yêu cầu dịch vụ...). Các thông tin này sẽ được sử dụng làm cơ sở cho việc xác định đường truyền.

Thứ hai, là việc phân tích, tính toán để chọn ra đường truyền khả dụng (cũng có thể là đường truyền tối ưu) dựa trên các thông tin trạng thái trên. Đường truyền khả dụng là đường truyền thoả mãn mọi yêu cầu của thông tin cần truyền (như tốc độ, băng thông) và điều kiện của mạng (như khả năng của đường truyền). Còn đường truyền tối ưu (theo một tiêu chuẩn nào đó) là đường truyền tốt nhất trong những đường truyền khả dụng.

1.2.2. Phân loại định tuyến mạng

1.2.2.1. Định tuyến tĩnh

Định tuyến tĩnh là dạng định tuyến mà các thông tin về đường đi phải do người quản trị mạng nhập cho router. Khi cấu trúc mạng có bất kỳ thay đổi nào thì chính người quản trị mạng phải xoá hoặc thêm các thông tin về đường đi cho router. Các đường đi này là cố định nên trong hệ thống mạng lớn việc bảo trì bảng định tuyến cho

router tốn rất nhiều thời gian. Định tuyến tĩnh là cách định tuyến không linh hoạt, không có khả năng tự thích nghi khi các điều kiện truyền thông thay đổi nên thường phù hợp với hệ thống mạng nhỏ hoặc tuyến đơn ít có biến đổi về thông tin định tuyến.

1.2.2.2. Định tuyến động

Định tuyến động là dạng định tuyến mà các đường đi được tự động cập nhật bởi router, việc lựa chọn tuyến đường dựa trên thông tin trạng thái hiện thời của mạng. Thông tin trạng thái có thể đo hoặc dự đoán và tuyến đường có thể thay đổi khi topo mạng hoặc lưu lượng mạng thay đổi. Thông tin định tuyến được cập nhật tự động vào trong các bảng định tuyến của các node mạng trực tuyến, và đáp ứng tính thời gian thực nhằm tránh tắc nghẽn cũng như tối ưu hiệu năng mạng. Định tuyến động có thể thích ứng với việc thay đổi kiến trúc mạng và lưu lượng truyền thông, phù hợp đối với mạng lớn, thường biến đổi trong quá trình hoạt động.

Chính vì định tuyến tĩnh đòi hỏi người quản trị mạng phải cấu hình mọi thông tin về đường đi cho router nên nó không có được tính linh hoạt như định tuyến động. Trong những hệ thống mạng lớn, định tuyến tĩnh thường được sử dụng kết hợp với giao thức định tuyến động cho một số mục đích đặc biệt.

1.2.3. Các giao thức định tuyến mạng

1.2.3.1. Định tuyến theo đường đi ngắn nhất

Để chọn một tuyến đường đi giữa hai bộ định tuyến, thuật toán chỉ cần tìm ra đường đi ngắn nhất giữa chúng trên đồ thị.

Khái niệm *đường đi ngắn nhất* ở đây có một số vấn đề cần phải làm rõ. Một cách để đo độ dài đường đi là số lượng các nút mà nó đi qua (còn được gọi là số bước nhảy). Bên cạnh các độ đo số

bước nhảy và khoảng cách địa lý, một số độ đo khác cũng có thể được sử dụng. Ví dụ, mỗi cung có thể được gán nhãn tương ứng với độ trễ trung bình của các gói tin do phải xếp hàng đợi trước khi được truyền đi. Với nhãn được gán kiểu này thì đường dẫn ngắn nhất chính là đường truyền nhanh nhất thay vì là đường ít bước nhảy nhất hay đường ngắn nhất xét về khoảng cách vật lý.

Trong trường hợp tổng quát, các nhãn trên mỗi cung được tính như một hàm theo khoảng cách, băng thông, lưu lượng trung bình, chi phí truyền tin, thời gian chờ trung bình, độ trễ,... và những yếu tố khác. Bằng cách thay đổi hàm trọng số, thuật toán sẽ tính lại được đường đi “ngắn nhất” tùy theo số lượng các yếu tố được chọn.

Cho đến nay, đã có khá nhiều giải thuật được đưa ra nhằm giải thuyết việc tính toán đường đi ngắn nhất trên đồ thị, nhưng giải thuật Dijkstra (1959) vẫn được xem là hiệu quả nhất và được sử dụng phổ biến trong các nghi thức định tuyến theo đường dẫn ngắn nhất hiện nay.

1.2.3.2. Định tuyến ngập lụt (Flooding)

1.2.3.3. Định tuyến theo vectơ khoảng cách

Hai giải thuật động thông dụng nhất hiện nay là *định tuyến theo vectơ khoảng cách* và *định tuyến theo tình trạng kết nối*. Giải thuật định tuyến theo vectơ khoảng cách hoạt động dựa vào việc duy trì một hàng (được gọi là một *vector*) bên trong mỗi bộ định tuyến cho biết *khoảng cách được biết* tốt nhất tới mỗi nút đích và cổng nào sẽ được sử dụng để đi đến đó, hàng này được cập nhật thường xuyên nhờ việc trao đổi thông tin với các bộ định tuyến kế cận.

- 1.2.3.4. Định tuyến theo tình trạng kết nối**
- 1.2.3.5. Định tuyến phân cấp**
- 1.2.3.6. Định tuyến quảng bá (broadcast)**
- 1.2.3.7. Định tuyến theo nhóm (multicast)**
- 1.2.3.8. Định tuyến cho mạng có các trạm di động**
- 1.2.3.9. Định tuyến cho mạng di động hỗn tạp (Ad Hoc)**
- 1.2.3.10. Tìm kiếm các nút trong mạng điểm nối điểm (Peer-to-Peer)**

1.3. KẾT LUẬN CUỐI CHƯƠNG

Để một mạng máy tính có thể tồn tại theo thời gian thì đòi hỏi phải có các thao tác vận hành thích hợp. Có rất nhiều các yếu tố khác nhau ảnh hưởng đến suốt quá trình hoạt động của mạng, đôi khi trong bản thân những yếu tố này lại chứa đựng mâu thuẫn lẫn nhau. Vì vậy sẽ không thể có một giải pháp chung đem lại hiệu quả tối ưu cho tất cả các mô hình mạng, mà ngược lại, nhà quản trị cần phải xem xét kỹ lưỡng từng hoàn cảnh cụ thể để xác định được những yếu tố nào cần phải được ưu tiên trước, những yếu tố nào có thể dung hòa được lẫn nhau, để từ đó đưa ra quyết định sáng suốt trong việc lựa chọn phương án thích hợp cho hệ thống mạng của mình.

CHƯƠNG 2

BÀI TOÁN ƯỚC LƯỢNG

NHU CẦU TRUYỀN THÔNG

2.1. PHÁT BIỂU BÀI TOÁN

2.1.1. Vai trò và mục tiêu đặt ra của bài toán

Nhu cầu truyền thông giữa các cặp nút nguồn - đích trên mạng thường được biểu diễn dưới dạng ma trận gọi là *ma trận nhu cầu truyền thông*, cùng với *sơ đồ cấu trúc mạng* và *nghi thức định tuyến*, chúng làm nên những yếu tố cơ bản cho việc giám sát, phân tích và tối ưu hóa quá trình truyền dữ liệu trên mạng. Hơn nữa, việc thu thập nhu cầu truyền thông trong một khoảng thời gian dài sẽ rất hữu ích cho việc thiết kế mạng, lập kế hoạch phát triển mở rộng mạng và từ đó đưa ra được các chiến lược kinh doanh thích hợp. Thông thường với những mạng nhỏ, lưu lượng truyền không cao, thì các ma trận này có thể được thu thập một cách trực tiếp thông qua chức năng lưu vết (log) của các bộ định tuyến. Tuy nhiên, cơ chế này sẽ không còn khả thi đối với các mạng trục lớn, ở đó lưu lượng truyền rất cao, bộ xử lý và bộ nhớ trong các thiết bị định tuyến hầu như phải tập trung hết cho việc luân chuyển các gói tin, nếu bật chế độ lưu vết để thu thập thông tin nhằm xây dựng ma trận nhu cầu truyền thông thì chúng sẽ trở nên trì trệ không thể chấp nhận được. Ngược lại, thông tin về lưu lượng truyền - ở đây chúng ta gọi là *tải trọng* - trên từng cổng (interface) của các thiết bị định tuyến thì hầu như luôn có sẵn và có thể được thu thập một cách dễ dàng thông qua nghi thức quản trị mạng đơn giản (SNMP).

Đến đây, chúng ta đã có thể thấy rõ vai trò không thể thiếu của ma trận nhu cầu truyền thông, tuy nhiên việc xác định chính xác chúng là điều hầu như không thể thực hiện được, vì vậy đề tài sẽ sử dụng một phương pháp nhằm *ước lượng ma trận nhu cầu truyền thông* dựa vào *thông tin tải trọng trên từng kết nối, sơ đồ cấu trúc mạng* và *nghi thức truyền thông* đang sử dụng.

2.1.2. Mô hình hóa bài toán

2.2. CÁC PHƯƠNG PHÁP ƯỚC LƯỢNG NHU CẦU TRUYỀN THÔNG

2.2.1. Phương pháp suy diễn dựa vào xác suất

Có bốn yếu tố làm đầu vào cho các phương pháp dựa vào lý thuyết xác suất: những giả định về xác suất phân bố của nhu cầu truyền thông sẽ quyết định chiến lược để tiến hành các bước kế tiếp và đó cũng chính là yếu tố để phân biệt sự khác nhau giữa các phương pháp. Yếu tố đầu vào thứ hai chính là ma trận nhu cầu truyền thông trước đó, được dùng làm khởi điểm cho thuật toán. Tham số kế tiếp là trọng số của các kết nối được sử dụng để tính đường đi ngắn nhất nhằm xây dựng nên ma trận định tuyến A. Tham số cuối cùng là tải trọng của các kết nối được thu thập trực tiếp từ các thiết bị định tuyến thông qua nghi thức SNMP dùng để áp đặt các ràng buộc trên ma trận nhu cầu truyền thông sẽ ước lượng.

2.2.2. Phương pháp quy hoạch tuyến tính (Linear Programming)

2.2.3. Phương pháp suy diễn Tomography

Phương pháp tomography đã được áp dụng vào bài toán ước lượng nhu cầu truyền thông, ở đây ta đo được tổng lưu lượng thông

tin truyền trên mỗi kết nối mạng bằng nghi thức SNMP và biết được cấu trúc bên trong của mạng thông qua chính sách định tuyến, vấn đề của chúng ta là phải tính ngược lại được nhu cầu truyền thông giữa từng cặp kết nối đó.

2.2.4. Phương pháp suy diễn dựa vào mô hình trọng lực (Gravity)

Một trong những phương pháp ước lượng ma trận nhu cầu truyền thông đơn giản nhất là dựa vào mô hình lực hấp dẫn. Theo luật hấp dẫn Newton thì lực hút giữa hai vật thể tỉ lệ thuận với tích khối lượng của hai vật thể chia cho bình phương khoảng cách giữa chúng.

2.3. THUẬT TOÁN TOMO-GRAVITY

2.3.1. Giới thiệu thuật toán

Như đã đề cập ở trên, trong cả hai phương pháp ước lượng nhu cầu truyền thông dựa vào mô hình gravity và tomography đều có những ưu và khuyết điểm riêng của nó, nhưng nhìn chung cả hai phương pháp đều không cho kết quả tốt như mong đợi. Vì vậy ở đây chúng ta sẽ nêu ra một giải pháp kết hợp ưu điểm của cả hai phương pháp trên và đặt tên là TomoGravity, giải pháp này tận dụng được cả nguyên lý hấp dẫn thông tin giữa các nút và những ràng buộc về tải trọng của các kết nối trên mô hình tomography.

2.3.2. Sơ đồ thuật toán

2.3.3. Chi tiết thuật toán

➤ **Bước 1:** Giải bài toán ước lượng nhu cầu truyền thông theo mô hình trọng lực cải tiến

➤ **Bước 2:** Tìm trong số những lời giải thuộc không gian vector bị ràng buộc bởi các điều kiện thu thập được qua phương pháp tomography một lời giải gần nhất so với lời giải thu được ở bước thứ nhất

2.4. CÁC GIẢI THUẬT TÌM ĐƯỜNG ĐI TỐI ƯU

Đường đi tối ưu từ A đến B là đường đi “ngắn nhất” trong số các đường đi có thể. Tuy nhiên khái niệm “ngắn nhất” có thể được hiểu theo nhiều ý nghĩa khác nhau tùy thuộc vào đơn vị dùng để đo chiều dài đường đi. Đối với các router, các đại lượng sau có thể được sử dụng để đo độ dài đường đi:

- Số lượng các router trung gian phải đi qua (HOP)
- Độ trì hoãn trung bình của các gói tin
- Chi phí truyền tin

Mỗi giải thuật chọn đường trước tiên phải chọn cho mình đơn vị đo chiều dài đường đi.

Để xác định được đường đi tối ưu, các giải thuật chọn đường sử dụng phương pháp đồ thị để tính toán. Trước tiên, nó mô hình hóa tình trạng mạng thành một đồ thị có các đặc điểm như sau:

- Nút là các router.
- Cạnh nối liền 2 nút là đường truyền nối hai router.
- Trên mỗi cạnh có giá đó là chiều dài đường đi giữa 2 router thông qua đường truyền nối hai router .

➤ Chiều dài đường đi từ nút A đến nút B là tổng tất cả các giá của các cạnh nằm trên đường đi. Nếu không có đường đi giữa 2 router thì xem như giá là vô cùng. Trên đồ thị này sẽ thực hiện việc tính toán tìm đường đi ngắn nhất.

2.4.1. Giải thuật tìm đường đi ngắn nhất Dijkstra

Mục đích là để tìm đường đi ngắn nhất từ một nút cho trước trên đồ thị đến các nút còn lại trên mạng.

2.4.2. Giải thuật chọn đường tối ưu Ford-Fulkerson

Mục đích của giải thuật này là để tìm đường đi ngắn nhất từ tất cả các nút đến một nút đích cho trước trên mạng.

2.4.3. Giải pháp vạch đường Vector khoảng cách (Distance Vector)

Ý tưởng của Distance Vector như sau: mỗi nút thiết lập một mảng một chiều (vector) chứa khoảng cách (chi phí) từ nó đến tất cả các nút còn lại và sau đó phát vector này đến tất cả các nút láng giềng của nó. Giá thiết đầu tiên trong Distance Vector là: mỗi nút phải biết được chi phí của các đường nối từ nó đến tất cả các nút láng giềng có đường nối kết trực tiếp với nó. Một nối kết bị đứt (down) sẽ được gán cho chi phí có giá trị vô cùng.

2.5. KẾT LUẬN CUỐI CHƯƠNG

Bài toán ước lượng nhu cầu truyền thông có ứng dụng rất lớn không chỉ cho việc tối ưu hóa quá trình định tuyến mà còn phục vụ cho việc duy trì và nâng cấp mạng nhằm luôn đảm bảo được chất lượng dịch vụ trước yêu cầu tài nguyên mạng ngày càng gia tăng. Về lý thuyết, bài toán có thể được giải một cách đơn giản bằng việc lưu lại vết của các luồng dữ liệu khi đi qua tất cả các bộ định tuyến, tuy nhiên trong điều kiện công nghệ như hiện nay, giải pháp này không thể thực hiện được nếu không muốn các thiết bị định tuyến bị trì trệ. Vì vậy giải pháp khả thi nhất vẫn là ước lượng nhu cầu truyền thông dựa vào việc thu thập một số thông tin khác với chi phí chấp nhận được.

CHƯƠNG 3

TỐI ƯU HÓA QUÁ TRÌNH ĐỊNH TUYẾN MẠNG BẰNG GIẢI THUẬT DI TRUYỀN

Trong phần này, đề tài sẽ tập trung trình bày giải pháp tối ưu hóa quá trình định tuyến áp dụng cho kỹ thuật truyền thông trên mạng IP, dựa vào nghi thức định tuyến hướng đích (destination - based routing protocol) nói chung và nghi thức định tuyến ưu tiên đường dẫn ngắn nhất (Open Shortest Path First - OSPF) nói riêng. Ngoài ra, đề tài cũng giới thiệu các khái niệm khác nhau trong việc tối ưu hóa quá trình định tuyến và ứng dụng của chúng trong việc triển khai cho kỹ thuật truyền thông. Vấn đề cần quan tâm đặc biệt là kỹ thuật định tuyến dựa vào nhiều độ đo - cụ thể là độ đo mức trễ (delay) và băng thông (bandwidth) - để có thể đưa ra đường dẫn ngắn nhất tới mỗi nút trên mạng. Một giải thuật mới dựa trên **giải thuật di truyền** (genetic algorithm) sẽ được trình bày cho phép tính toán tập tối ưu các trọng số của những kết nối, kết hợp với các nghi thức truyền thông dựa trên độ đo đơn cũng như kép. Cuối cùng là phần trình bày những ưu điểm của nghi thức dựa trên độ đo kép so với độ đo đơn.

3.1. TỔNG QUAN BÀI TOÁN TỐI ƯU HÓA QUÁ TRÌNH ĐỊNH TUYẾN MẠNG

Tối ưu hóa quá trình định tuyến là một phương pháp rất quan trọng trong kỹ thuật truyền thông trên Internet, được áp dụng cho các luồng thông tin (ví dụ như quyết định hợp hay tách các luồng thông tin vào/ra giữa các nút nhất định) sau những khoảng thời gian cụ thể

cho trước, thường tối thiểu là vài giờ đồng hồ. Đối với những nơi cần gia tăng thêm lưu lượng truyền thông hoặc những biến đổi lưu lượng tạm thời gây ra tắc nghẽn đường truyền một cách cục bộ, tối ưu hóa quá trình định tuyến có thể giải quyết, hoặc ít nhất cũng giảm bớt được những vấn đề tiềm tàng ảnh hưởng đến hiệu suất mạng. Ý tưởng cơ bản là *điều chỉnh lộ trình của những luồng tải hiện thời* và do đó nâng cao tài nguyên mạng hiện có, dẫn tới việc chất lượng dịch vụ (QoS - Quality of Service) được cải thiện.

3.1.1. Phân biệt định tuyến hướng đích với định tuyến hướng nguồn hay lộ trình.

Hai khái niệm định tuyến khác nhau cơ bản hiện đang tồn tại, chúng ảnh hưởng mạnh mẽ đến các phương thức tối ưu hóa cũng như kết quả đạt được, đó là: *định tuyến hướng đích* và *định tuyến hướng nguồn* hay lộ trình. Các nghi thức định tuyến truyền thống đều thuộc lĩnh vực định tuyến hướng đích như OSPF (Open Shortest Path First), EIGRP (Enhanced Interior Gateway Routing Protocol), hay IS-IS (Intermediate System-to-Intermediate System) do chúng sẽ chỉ ra bước kế tiếp dựa trên thông tin đích đến.

3.1.2. Phân biệt giữa định tuyến đơn độ đo và đa độ đo

Với nghi thức định tuyến hướng đích, các bộ định tuyến sẽ xác định cổng đi ra (outgoing-interface) dựa trên những giá trị độ đo mà có thể lượng hóa bằng khoảng cách tới nút đích. Thông thường nhất, các độ đo đơn sẽ được gán vào mỗi kết nối và thuật toán tìm đường đi ngắn nhất được dùng để xác định đường dẫn thích hợp từ một nút tới các nút khác trên mạng, gọi là *định tuyến đơn độ đo*. Trong khi các độ đo kết nối thường có một ý nghĩa vật lý nhất định chẳng hạn như là *độ trễ* hoặc *chi phí*, chúng cũng có thể được dùng

theo cách thông thường thuận túy chỉ để tối ưu hóa việc định tuyến. Bằng cách thiết lập các độ đo phù hợp một cách gián tiếp chúng ta có thể tối ưu hóa được phương thức định tuyến.

Ngoài những nghi thức với đơn độ đo, cũng đã có các phương thức định tuyến cho phép sử dụng nhiều hơn một độ đo khi tính toán chiều dài của các đường dẫn tới mỗi nút đích, gọi là *định tuyến đa độ đo*.

Việc tối ưu hóa quá trình định tuyến dựa trên các nghi thức với độ đo kép có thể cho kết quả thực thi tốt hơn so với độ đo đơn. Bởi vì độ đo thứ hai có thể được sử dụng một cách linh hoạt cho phép sản sinh ra nhiều mô hình định tuyến phong phú. Mặt khác, ta luôn luôn có thể sử dụng nghi thức định tuyến với độ đo kép để sinh ra một mô hình định tuyến với độ đo đơn.

3.1.3. Nhiều đường dẫn với chi phí bằng nhau (Equal-Cost Multi-Path ECMP)

Một đặc tính khác có thể có của các nghi thức định tuyến gây ảnh hưởng rất lớn đến quá trình tối ưu hóa là việc chia sẻ tải trọng. Trong các nghi thức định tuyến hướng đích, khả năng này thường được triển khai dưới hình thức của khái niệm “nhiều đường dẫn với chi phí bằng nhau” (ECMP).

3.1.4. Những lớp tối ưu hóa quá trình định tuyến

Dựa trên các khái niệm đã được trình bày trong phần trên, nhiều loại giải pháp định tuyến trên mạng IP có thể được triển khai. Các nghi thức định tuyến hướng nguồn hay lộ trình (như MPLS) chắc chắn sẽ đem lại những tiềm năng to lớn trong việc tối ưu hóa quá trình định tuyến. Giả sử rằng không có bất cứ hạn chế nào liên quan đến quá trình định tuyến trên toàn bộ mạng, chúng ta có thể tìm

ra một giải pháp tối ưu toàn cục bằng cách áp dụng quy hoạch tuyến tính (LP-linear programming) hoặc áp dụng heuristic thích hợp. Tuy nhiên, ở đây chúng ta chỉ chú trọng đến các nghi thức định tuyến hướng đích, do đó chỉ cần tính được chặn dưới của vấn đề tối ưu hóa quá trình định tuyến bằng quy hoạch tuyến tính.

Loại giải pháp tối ưu hóa quá trình định tuyến hướng đích lại tiếp tục được chia ra thành mô hình với độ đo đơn (điển hình là OSPF) và độ đo kép (điển hình là EIGRP), mỗi loại có một phiên bản kết hợp với việc chia sẻ tải trọng (ECMP) và không chia sẻ. Hầu hết các tài liệu về loại tối ưu hóa quá trình định tuyến như thế này đều xem OSPF như là nghi thức nền tảng bên dưới. Tuy nhiên, không thể áp dụng quy hoạch tuyến tính cho các mạng với kích cỡ thuộc loại từ trung bình trở lên được, khi đó cần phải kết hợp thêm heuristic.

3.2. NGHI THỨC ĐỊNH TUYẾN MỞ ƯU TIÊN ĐƯỜNG DẪN NGẮN NHẤT (OSPF)

3.2.1. Vài nét về quá trình hình thành và phát triển

3.2.2. Sơ lược nguyên tắc hoạt động

3.3. ỨNG DỤNG GIẢI THUẬT DI TRUYỀN CHO NGHI THỨC TRUYỀN THÔNG OSPF

3.3.1. Tổng quan về giải thuật di truyền

3.3.1.1. Khái niệm giải thuật di truyền

Giải thuật di truyền là kỹ thuật chung giúp giải quyết vấn đề, bài toán bằng cách mô phỏng sự tiến hóa của con người hay của sinh vật nói chung (dựa trên thuyết tiến hóa muôn loài của Darwin) trong điều kiện qui định sẵn của môi trường, tìm kiếm dựa trên cơ chế chọn lọc và di truyền tự nhiên. Thuật giải này sử dụng các nguyên lý di truyền về sự thích nghi và sự sống các cá thể thích nghi nhất trong

tự nhiên. GA là một thuật giải và mục tiêu của GA không nhằm đưa ra lời giải chính xác tối ưu mà là đưa ra lời giải tương đối tối ưu.

Tập hợp tất cả các lời giải trong không gian tìm kiếm được gọi là *kiểu hình*. Các kiểu hình này khi mã hoá gọi là *kiểu gen*. Toán tử di truyền sẽ được thực thi trên đối tượng này. Một ánh xạ từ kiểu hình sang kiểu gen gọi là *quá trình mã hoá*. Mỗi cá thể trong kiểu gen có nhiều nhiễm sắc thể. Trong mỗi nhiễm sắc thể có chứa nhiều gen. Mỗi đặc trưng di truyền cụ thể được quy định bởi giá trị và vị trí của gen trong nhiễm sắc thể. Độ thích nghi là thước đo khả năng sống sót và phát triển của cá thể trong môi trường.

Toán tử xác định cá thể trong thế hệ hiện tại được giữ lại trong thế hệ kế tiếp được gọi là chọn lọc. Toán tử kết hợp ngẫu nhiên hai cá thể được chọn gọi là lai ghép. Toán tử thay đổi ngẫu nhiên cấu trúc cá thể, tức làm thay đổi giá trị của gen gọi là đột biến.

Thuật toán di truyền gồm có bốn quy luật cơ bản là lai ghép, đột biến, sinh sản và chọn lọc tự nhiên.

Cấu trúc thuật giải di truyền tổng quát:

Bắt đầu

$t = 0;$

Khởi tạo $P(t)$

Tính độ thích nghi cho các cá thể thuộc $P(t);$

Khi (điều kiện dừng chưa thỏa) **lặp**

$t = t + 1;$

Chọn lọc $P(t);$

Lai $P(t);$

Đột biến $P(t);$

Hết lặp

Kết thúc

3.3.1.2. Các tính chất quan trọng của giải thuật di truyền

➤ GA lập luận mang tính chất ngẫu nhiên để tìm giải pháp tối ưu cho những vấn đề phức tạp, thay vì xác định như toán học giải tích. Tuy nhiên đây là hình thức ngẫu nhiên có hướng dẫn bởi trị số thích nghi. Chính hàm số thích nghi là giúp GA tìm giải pháp tối ưu trong rất nhiều giải pháp có thể có.

➤ GA không đề ý đến chi tiết vấn đề, trái lại chỉ chú ý đến giải pháp cho vấn đề, hay tìm điều kiện tối ưu cho việc điều hành, và phân nhóm những giải pháp có được.

➤ GA được sử dụng đặc biệt cho những bài toán yêu cầu tìm kiếm tối ưu toàn cục với không gian tìm kiếm lớn và không thể kiểm soát nhờ khả năng duyệt qua không gian tìm kiếm đại diện mà không thực sự đi qua từng điểm của toàn bộ không gian.

3.3.1.3. Cơ chế thực hiện của giải thuật di truyền

- **Mã hóa**
- **Chọn lọc cá thể**
- **Lai ghép**
- **Đột biến**
- **Hàm thích nghi**
- **Xử lý các ràng buộc**

GA thích hợp cho các bài toán tìm kiếm tối ưu với điều kiện không ràng buộc. Tuy nhiên thực tế bài toán có thể chứa một hoặc nhiều ràng buộc phải thỏa. Lời giải nhận được từ chiến lược tìm kiếm tối ưu nhất thiết phải nằm trong vùng khả thi, tức phải thỏa tất cả các ràng buộc. Thông thường có thể xử lý các ràng buộc bằng hàm phạt.

- **Điều kiện kết thúc lặp của GA**

Để kết thúc vòng lặp GA, thường có thể chỉ định trước số thế hệ cần tạo ra sau đó kiểm tra lại độ thích nghi những phần tử tốt nhất bằng cách so sánh với bài toán ban đầu.

3.3.2. Tìm kiếm cục bộ với heuristic

Để cải thiện hiệu quả và tốc độ của giải thuật di truyền, một bước tìm kiếm cục bộ sẽ được áp dụng trước khi ước lượng kết quả. Trước khi ước lượng mỗi giải pháp định tuyến, các heuristic đơn giản được sử dụng để phân tán bớt những luồng dữ liệu tại kết nối đang bị chiếm dụng nhiều nhất. Điều này được lặp đi lặp lại cho đến khi không còn cải thiện được kết quả nữa. Do việc tìm kiếm với heuristic cho kết quả xác định nên với cùng một chuỗi độ đo luôn luôn cho cùng một giải pháp định tuyến. Vì vậy nó thích hợp để áp dụng vào các kết cấu di truyền.

➤ Độ đo đơn với heuristic

Trong trường hợp định tuyến dựa theo đường dẫn ngắn nhất, các luồng thông tin có thể phân tán khỏi một kết nối bằng cách gia tăng độ đo của nó. Do đó, chúng ta sẽ tăng độ đo của những kết nối đang bị chiếm dụng cao nhất và định tuyến lại tất cả các luồng thông tin. Nếu bước này làm gia tăng giá trị chiếm dụng đường truyền tối đa thì nó sẽ được phục hồi lại và quá trình tối ưu hóa cục bộ với heuristic sẽ kết thúc. Ngược lại, hành động này sẽ được lặp lại cho kết nối mà bây giờ bị chiếm dụng cao nhất.

➤ Độ đo kép với heuristic

Đối với các nghi thức định tuyến với hai độ đo như: nghịch đảo của băng thông và độ trễ, có nhiều khả năng hơn cho việc dịch chuyển các luồng thông tin khỏi những kết nối riêng lẻ, đơn giản chỉ cần gia tăng độ đo trễ hay nghịch đảo của băng thông. Ngoài ra, nếu kết hợp cả hai cũng có thể cho kết quả làm thay đổi cấu trúc đường

đẫn: độ đo trễ có thể được gia tăng trong khi độ đo nghịch đảo của băng thông lại giảm đi, hoặc ngược lại. Do đó, quá trình tìm kiếm với heuristic lặp đi lặp lại sẽ áp dụng những thay đổi độ đo này vào kết nối bị chiếm dụng nhiều nhất. Mỗi thay đổi độ đo được chấp nhận nếu nó không làm gia tăng giá trị chiếm dụng đường truyền tối đa. Nếu không còn tối ưu được nữa thì quá trình tìm kiếm heuristic với độ đo kép sẽ dừng lại.

3.3.3. Kết luận

Trong phần này của đề tài, chúng ta đã thảo luận các giải pháp khác nhau nhằm tối ưu hóa quá trình định tuyến trong kỹ thuật truyền thông trên mạng IP. Đối với vấn đề thiết lập các độ đo nhằm tối ưu hóa, chúng ta đã trình bày một giải thuật di truyền, nó được xem xét trên các nghi thức định tuyến hướng đích với độ đo đơn và cả độ đo kép. Cùng với các hàm cơ bản của giải thuật di truyền, phiên bản đề nghị còn được tích hợp thêm thuật giải tìm kiếm đơn giản với heuristic nhằm nâng cao tốc độ quá trình tối ưu hóa.

Kết quả thu được từ những kịch bản mạng khác nhau cho thấy quá trình tối ưu hóa việc định tuyến, nói chung luôn cải thiện được hiệu suất làm việc của mạng.

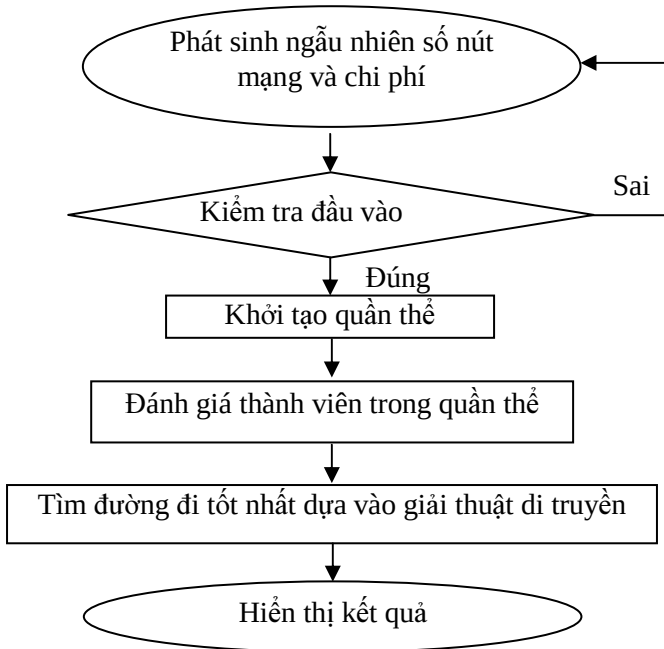
3.4. XÂY DỰNG CHƯƠNG TRÌNH TỐI ƯU HÓA QUÁ TRÌNH ĐỊNH TUYẾN MẠNG

3.4.1. Bài toán định tuyến

Cho một tập các nút mạng tương ứng với các khoảng cách hình học khác nhau được mô tả bằng một ma trận, ma trận này đại diện cho chi phí khi đi qua mỗi nút mạng. Tìm một tuyến đường đi qua tất cả các nút sao cho chi phí nhỏ nhất, mỗi nút chỉ được qua một lần sử dụng giải thuật di truyền.

3.4.2. Thiết kế ứng dụng

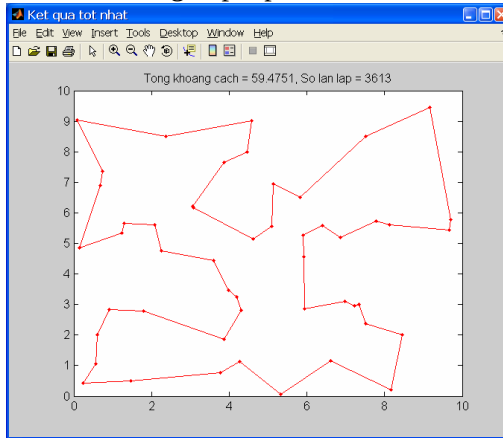
Để minh họa cho việc ứng dụng GA nhằm tìm kiếm đường đi tối ưu trong quá trình định tuyến trên mạng IP, dưới đây sẽ xây dựng chương trình mô phỏng cho bài toán định tuyến ở trên với số nút mạng và chi phí được giả định là ngẫu nhiên (phù hợp với bài toán định tuyến động trên mạng). Chương trình sử dụng công cụ lập trình Matlab sẽ tìm ra được một chu trình sao cho tổng chi phí là nhỏ nhất.



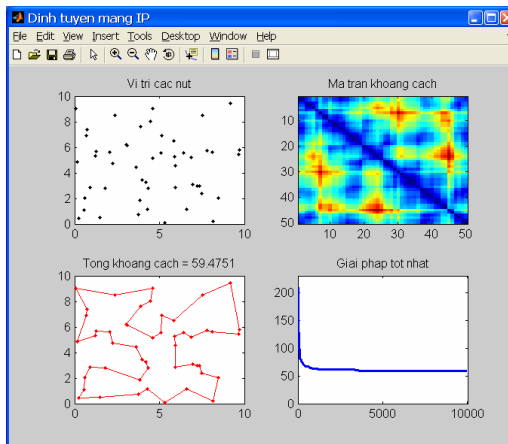
Hình 3-10. Sơ đồ giải thuật

➤ **Kết quả chương trình:** với số nút mạng và chi phí được phát sinh ngẫu nhiên, chương trình cho kết quả là chu trình với tổng khoảng cách là ngắn nhất sau nhiều lần lặp dựa vào giải thuật di

truyền. Ngoài ra còn cho hiển thị vị trí các nút đã được phát sinh, ma trận khoảng cách và tìm ra giải pháp tốt nhất cho bài toán.



Hình 3-11. Kết quả tốt nhất của bài toán tìm chu trình với khoảng cách ngắn nhất



Hình 3-12. Kết quả chạy mô phỏng tìm chu trình với khoảng cách ngắn nhất

KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Có thể nói giải thuật di truyền là một cơ chế tìm kiếm song song ẩn với các đầu vào song song chính là các cá thể của quần thể khởi tạo. Các đầu vào song song có độ thích nghi lớn sẽ đưa vào thế hệ mới, đồng thời nó cũng là cơ sở cho việc tạo thêm các cá thể mới (chính là sự thêm vào các điểm trong không gian tìm kiếm) nhờ các phép lai tạo và đột biến. Quá trình này lặp đi lặp lại nhiều thế hệ và dần dần không gian tìm kiếm sẽ hội tụ về điểm tối ưu toàn cục. Bài toán tối ưu là bài toán chung của tất cả các ngành khoa học kỹ thuật cũng như xã hội. Chính vì vậy giải thuật di truyền có ứng dụng rộng rãi. Giải thuật di truyền có thể được áp dụng cho quá trình định tuyến mạng IP vì bản chất của định tuyến là tìm ra con đường có tổng chi phí là cực tiểu. Thay vì sử dụng các giải thuật khác, giải thuật di truyền cho phép tìm ra được đúng giá trị tối ưu toàn cục.

Qua quá trình nghiên cứu, luận văn đã giải quyết được các vấn đề gồm: Phân tích giải thuật di truyền, Tìm hiểu các khái niệm về định tuyến mạng IP, Tối ưu hóa quá trình định tuyến mạng dựa vào giải thuật di truyền.

Kết quả chương trình hiện tại là tìm ra đường đi tối ưu với số nút mạng và chi phí được giả định là ngẫu nhiên. Hướng phát triển của đề tài là phát triển một phần mềm ứng dụng cho phép người dùng tương tác với chương trình và thay đổi các thông số trạng thái kết nối dựa vào kết quả nghiên cứu này.