TỔNG QUÁT VỀ ĐẶC TÍNH HÓA LUỒNG LƯU LƯỢNG IP

Dương Tuấn Anh, Hoàng Minh

Học viện Công nghệ Bưu chính Viễn thông

Đến Tòa soạn ngày 7/10/2010

1. GIỚI THIỆU

1.1. Về vấn đề nghiên cứu

Sự bùng nổ của Internet đã "cách mạng hoá" mọi khía cạnh trong cuộc sống của chúng ta. Tuy nhiên, với tốc độ phát triển nhanh chóng cả về dịch vụ, số lượng khách hàng, cả về thông tin "hữu ích" lẫn "tin rác", "tin tặc" trong khi sự phát triển hạ tầng cơ sở bị hạn chế nhìn từ phía đảm bảo chất lượng dịch vụ QoS, khía cạnh an ninh thông tin, an ninh mạng. Điều đó yêu cầu các nhà quản lí, khai thác mạng và cung cấp dịch vụ phải hiểu rõ các mẫu và đặc tính lưu lượng của IP.

Xuất phát từ thực tiễn, tính cấp thiết của những vấn đề đã nêu, "Tổng quát về đặc tính hóa luồng lưu lượng IP" nhằm khai thác đặc thù của mỗi loại hình dịch vụ (thể hiện luồng lưu lượng) để xây dựng hệ các điều kiện giới hạn phục vụ việc bảo an, phân bổ lưu lượng Internet và các giải pháp hỗ trợ, giải quyết tắc nghẽn đường truyền và đảm bảo QoS cho mọi ứng dụng phân bố trong môi trường hỗn tạp. Trong đó, chúng tôi tập trung vào trình bày một số mục tiêu quan trọng sau: (1). Nghiên cứu tổng quan về đặc tính hóa lưu lượng của Internet. (2). Nghiên cứu, xác định các tham số đại diện cho mỗi lớp dịch vụ để xác định thứ tự ưu tiên của các dịch vụ. (3). Nghiên cứu các phương pháp điều khiển ưu tiên khác nhau đối với các lớp dịch vụ để đề xuất chiến lược định tuyến theo thời gian thực một cách phù hợp với các hướng lưu lượng nhằm phục vụ nhiệm vụ chống tắc nghẽn và đảm bảo chất lượng dịch vụ. (4). Triển khai thử nghiệm tại cơ sở cung cấp dịch vụ Viễn thông Thừa Thiên Huế để đánh giá tính xác thực của kết quả thu được.

1.2. Các bước tiến hành nghiên cứu

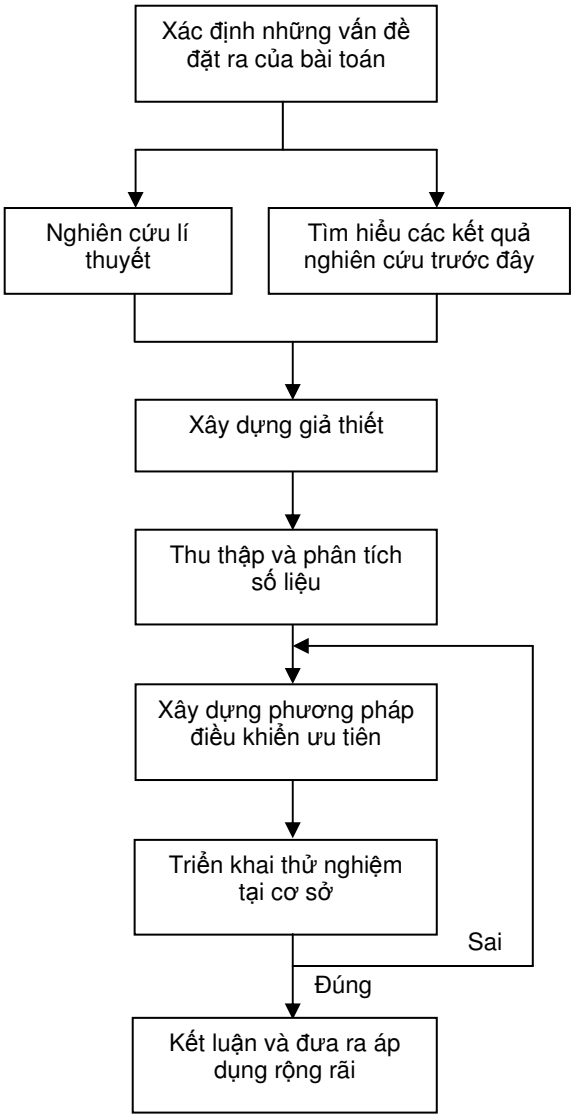
Để giải quyết những vấn đề đã nêu, chúng tôi đã đề xuất các bước triển khai nghiên cứu như lược đồ mô tả trong hình 1.1.

1.3. Một số kết quả

Trong bài báo này chúng tôi trình bày những kết quả nghiên cứu liên quan đến vấn đề đặc tính hóa lưu lượng của Internet và xác định các tham số đại diện cho mỗi lớp dịch vụ, tạo tiền đề cho viêc thu thập và phân tích số liệu.

Những kết quả nghiên cứu liên quan đến nhiệm vụ xây dựng phương pháp xác định, phân loại thông tin và thứ tự dịch vụ ưu tiên trên cơ sở các tham số đặc trưng luồng lưu lượng sẽ được

trình bày ở công trình tiếp theo. Tiếp đến là công trình báo cáo về luận cứ khoa học và đề xuất phương án điều khiển theo thứ tự ưu tiên của các dịch vụ và cuối cùng là báo cáo về kết quả thử nghiệm tại Viễn thông Thừa Thiên Huế và những xét, đánh giá, kiến nghị liên quan.



Hình 1.1. Sơ đồ tổng quát mô tả các bước triển khai nghiên cứu

2. TỔNG QUAN VỀ LƯU LƯƠNG IP

Đã có nhiều công trình nghiên cứu về đặc tính hóa lưu lượng Internet, cả về khía cạnh phân tích lẫn khía cạnh nắm bắt, đo lường chất lượng [1 - 5]. Các công trình đó chỉ ra rằng công nghệ mạng Internet phát triển nhanh hơn so với sự phát triển về phân tích và hiểu biết lí thuyết về cách xử lí mạng. Điều đó kiềm chế trong lĩnh vực dự đoán các dịch vụ, phép đo liên quan đến độ dài hàng đợi và trễ mạng, tuy những kĩ thuật sử dụng mô hình toán học truyền thống có một số thành công, nhất là lí thuyết hàng đợi. Nhiều lí do khiến việc tập hợp các đặc tính lưu lượng

Internet (cả mạng đường trục) đã không nhận được nhiều sự quan tấm đúng mức. Trong đó, nguyên nhân chính là nổ lực chỉ tập trung vào các hoạt động sản xuất, không có thời gian, nguồn lực dành cho việc thu thập, hệ thống hóa dữ liệu. Điều này dẫn đến sự tồn tại khoảng cách giữa các kết quả thí nghiệm (trong môi trường cô lập) với những đặc tính của cơ sở hạ tầng Internet rộng lớn phân tán ở mọi nơi và trên phạm vi toàn cầu [6 - 10].

Để giảm nhỏ khoảng cách giữa lí thuyết và thực nghiệm, ở đây dựa trên cơ sở lí thuyết tiến hành điều tra mạng viễn thông tại đơn vị cung cấp dịch vụ nhằm tổng hợp các số liệu lưu lượng, đóng góp vào sự hiểu biết về hành vi mạng ở quy mô lớn. Sở dĩ như vậy là vì khả năng lí thuyết về hành vi mạng có thể áp dụng, tạo điều kiện cho các nhà cung cấp dịch vụ mạng thực hiện có hiệu quả về quản lí, phát triển cơ sở hạ tầng quy mô lớn. Ngược lại, đòi hỏi thực tế về cung cấp dịch vụ làm động lực ứng dụng kết quả của lĩnh vực nghiên lí thuyết đến thế giới thực của mạng. Từ cách nhìn nhận như vậy, những kết quả nghiên cứu được trình bày ở các phần tiếp theo.

2.1. Về vấn đề lưu lượng và phân loại lưu lượng trên Internet

Có nhiều loại lưu lượng khác nhau cùng tồn tại trên mạng, tùy theo từng ứng dụng cụ thể. Có thể phân các ứng dụng thành loại theo thời gian thực và không theo thời gian thực. Các loại lưu lượng khác nhau có những đặc tính ngẫu nhiên và nhu cầu về QoS khác nhau. Có thể phân tích các yêu cầu về QoS dựa trên các mức thời gian và mức truyền dịch vụ (mức gói, cuộc gọi) khác nhau. Thường sử dụng phân tích thống kê để xác định các đặc tính lưu lượng và sử dụng các đặc tính lưu lượng đó với các yêu cầu tương ứng của loại dịch vụ làm các tham số chính để phân loại lưu lượng IP [11 - 15]. Mô hình lưu lượng đối với dịch vụ cơ bản (thoại) dựa trên lí thuyết Erlang, gồm hai tham số chính (tốc độ xuất hiện cuộc gọi và thời gian chiếm giữ trung bình) để xác định lưu lượng ra ứng với thời gian sử dụng tài nguyên.

Nếu phân bố Poisson phù hợp với thống kê xuất hiện các cuộc gọi thì đối với một hệ thống có N kênh, tổng lưu lượng A, thời gian giữ cuộc trung bình τ, độ trễ lớn nhất cho phép trước khi cuộc gọi mới bị khóa T, xác suất không có server (hệ thống bận, cuộc gọi mới bị khóa) GOS theo mô hình Erlang B và xác suất khi độ trễ lớn hơn T tuân theo mô hình Erlang C [11]:

$$\text{Erlang B: } GOS = \frac{\frac{A^N}{N!}}{\sum_{k=0}^{N}\frac{A^k}{k!}};$$

$$\text{Erlang C: } GOS = \frac{A^N}{A^N + N!\left(1-\frac{A}{N}\right)\sum_{k=0}^{N-1}\frac{A^k}{k!}} e^{\frac{(N-A)T}{\tau}} \qquad (1.1)$$

2.2. Luồng lưu lượng IP, cơ chế kết thúc và thuộc tính luồng

Hiện chưa có định nghĩa thống nhất và chuẩn hoá nào đối với luồng lưu lượng Internet (tùy mục đích nghiên cứu mà luồng lưu lượng được định nghĩa theo cách khác nhau). Mục tiêu ở đây là xây dựng mô hình nhận dạng lưu lượng của các ứng dụng trên mạng Internet trên cơ sở các kết quả phân tích đặc tính, biểu hiện hoạt động của luồng lưu lượng. Nên, nhu cầu sử dụng định nghĩa có thể cung cấp được nhiều thông số thống kê luồng lưu lượng và đồng nhất theo thời gian.

Chúng tôi đã sử dụng phương pháp định nghĩa luồng lưu lượng gần giống với phương pháp của RTFM, trong đó thuật ngữ "*khoá"* được sử dụng để chỉ tập các thuộc tính địa chỉ trên các lớp khác nhau của mô hình TCP/IP, gồm các thông số về địa chỉ IP nguồn, địa chỉ IP đích, cổng ứng dụng nguồn, cổng ứng dụng đích và chỉ số nhận dạng giao thức TCP: *Hai gói liên tiếp được xem là thuộc cùng một luồng nếu chúng không cách nhau quá một khoảng thời gian cho trước (gọi là thời gian time-out), nếu không một luồng mới sẽ được tạo ra.*

Bảng 2-1. Các thuộc tính chung của các luồng

STT	Thuộc tính	Mô tả
1.	Khoá (định nghĩa luồng)	Tập 5 tham số: địa chỉ IP nguồn, đích; cổng ứng dụng nguồn, đích; số nhận dạng giao thức
2.	Bắt đầu	Thời điểm đến của gói đầu tiên của luồng
3.	Kết thúc	Thời điểm đến của gói cuối cùng của luồng
4.	Thời lượng	Độ dài luồng lưu lượng đo bằng μs
5.	Kích thước gói đầu	Độ dài phần tải tin của gói đầy đủ đầu tiên
6.	Hướng một số gói đầu	Xác định hướng truyền của luồng (10 giá trị rời rạc): 1: hướng từ nút khởi tạo đến nút đáp ứng -1: theo hướng ngược lại
7.	Tỉ lệ dữ liệu	Tỉ lệ giữa tổng số byte dữ liệu truyền giữa khởi tạo và đáp ứng luồng
8.	Phân bố giữa các thời điểm đến của gói	Phân bố rời rạc khoảng thời gian giữa thời điểm đến của các gói thuộc luồng, thể hiện bởi chuỗi 9 bin giá trị liên tục (tỉ lệ tương đối các gói có thời điểm đến nằm trong bin đó)
9.	Phân bố giữa các thời điểm đến của luồng	Độ dài phần đuôi của biểu đồ phân bố thời điểm đến của các luồng lưu lượng quan sát trong một cửa sổ thời gian cho trước
10.	Phân bố số luồng	Tỉ lệ % số lượng luồng quan sát được trong một cửa sổ thời gian cho trước (2 phút)
11.	Các thuộc tính thể hiện phiên làm việc kiểu đối thoại	
	αdt	Tỉ lệ giữa số lượng gói đầy đủ thuộc đoạn đối thoại trên tổng số gói đầy đủ của luồng
	Bdt	Tỉ lệ giữa số lượng gói đầy đủ thuộc đoạn đối thoại trung bình trên tổng số gói đầy đủ của luồng
	Γdt	Tỉ lệ các gói đối thoại truyền bởi phía khởi tạo phiên làm việc
12.	Các thuộc tính thể hiện phiên làm việc kiểu giao dịch	
	αgd	Dấu hiệu cho biết mức độ thường xuyên của các trao đổi gói kiểu ping-pong quan sát được trong luồng lưu lượng

Bảng 2-2. Các thuộc tính trên mỗi hướng của luồng

STT	Thuộc tính	Mô tả	Kiểu giá trị
1	Giữa thời điểm đến (I.A.T) của các gói	Phân bố rời rạc: Các khoảng thời gian giữa các I.A.T của các gói trên hướng đang xét. Giá trị trong mỗi bin nằm giữa 0 và 1, thể hiện tỉ lệ tương đối các gói trong bin.	9 giá trị liên tục
2	Phân bố tải tin	Phân bố rời rạc: Chiều dài phần tải tin của gói (bin giá trị về chiều dài phần tải tin của mỗi gói, giá trị trong mỗi bin nằm giữa 0 và 1)	23 giá trị liên tục
3	Số byte	Tổng số byte truyền (cả các byte tiêu đề lớp mạng và lớp vận chuyển) trên hướng xét	Liên tục
4	Số byte dữ liệu	Tổng số byte dữ liệu truyền trên hướng xét	Liên tục
5	Số gói	Tổng số gói truyền trên hướng xét	Liên tục
6	Số gói dữ liệu	Tổng số gói đầy đủ truyền trên hướng xét	Liên tục
7	Các thuộc tính thể hiện phiên làm việc kiểu tương tác từ bàn phím		
	Ap	Thay đổi thời lượng giữa các lần phát gói tương tác (nhận từ bàn phím)	Liên tục
	Bp	Dấu hiệu tương tác dựa trên tỉ lệ các gói nhỏ	Liên tục
	γp	Dấu hiệu về sự xuất hiện liên tục của các gói nhỏ	Liên tục
	Δp	Dấu hiệu về kiểu truyền kết hợp	Liên tục
	Ep	Sự không bình thường (đồng bộ) của thời lượng giữa các gói của các gói nhỏ liên tiếp	Liên tục
8	Các thuộc tính thể hiện phiên làm việc kiểu tương tác dòng lệnh		
	Adl	Sự thay đổi về thời lượng giữa các lần phát gói (đối với các gói dòng lệnh)	Liên tục
	Bdl	Dấu hiệu tương tác dựa trên tỉ lệ các gói nhỏ	Liên tục
	Γdl	Dấu hiệu về sự xuất hiện liên tục của các gói nhỏ	Liên tục
	Δdl	Dấu hiệu về kiểu truyền kết hợp	Liên tục
	Edl	Sự bất thường (đồng bộ) của thời lượng giữa các gói của các gói nhỏ liên tiếp	Liên tục
9	Các thuộc tính thể hiện phiên làm việc kiểu truyền file		
	Adl	Sự thay đổi về thời lượng giữa các lần phát gói trong quá trình truyền file	Liên tục
	Bdl	Dấu hiệu về kiểu truyền dữ liệu dựa trên tỉ lệ gói lớn	Liên tục
	Γdl	Dấu hiệu về sự xuất hiện liên tục các gói truyền kết hợp	Liên tục

10	Bit trung bình	Tốc độ bit trung bình đo trong các khoảng 5s	Liên tục
11	Abtb	Dấu hiệu thể hiện mức độ gần với trị trung bình của tốc độ bit đo trong các khoảng 5s (tính không đổi tốc độ bit)	Liên tục
12	Gói trung bình	Tốc độ gói trung bình đo trong những khoảng 5s	Liên tục
13	Agtb	Dấu hiệu thể hiện mức độ gần với trị trung bình của tốc độ gói đo trong các khoảng 5s (tính không đổi tốc độ gói)	Liên tục
14	Giá trị trung tâm tải tin	Trung tâm của phân bố chiều dài phần tải tin của các gói	Liên tục
15	Aptb	Thể hiện mức độ gần với trị trung tâm của độ dài phần tải tin của các gói (tính không đổi độ dài phần tải tin)	Liên tục

Trong quá trình phân tích luồng, ngoài định nghĩa khoá, phải xác định các cơ chế bắt đầu, kết thúc của một luồng lưu lượng. Có ba phương pháp cơ bản thường dùng để xác định sự kết thúc của một luồng lưu lượng; dựa trên cơ chế hoạt động của giao thức, thời gian time-out cố định và thời gian Time-out thích ứng. Trên thực tế, số thuộc tính của luồng lưu lượng IP thu được là rất lớn, nhưng chỉ một số đặc tính của luồng được sử dụng cho việc nhận dạng, phân loại luồng lưu lượng [5] như ở trong bảng 2-1 và bảng 2-2.

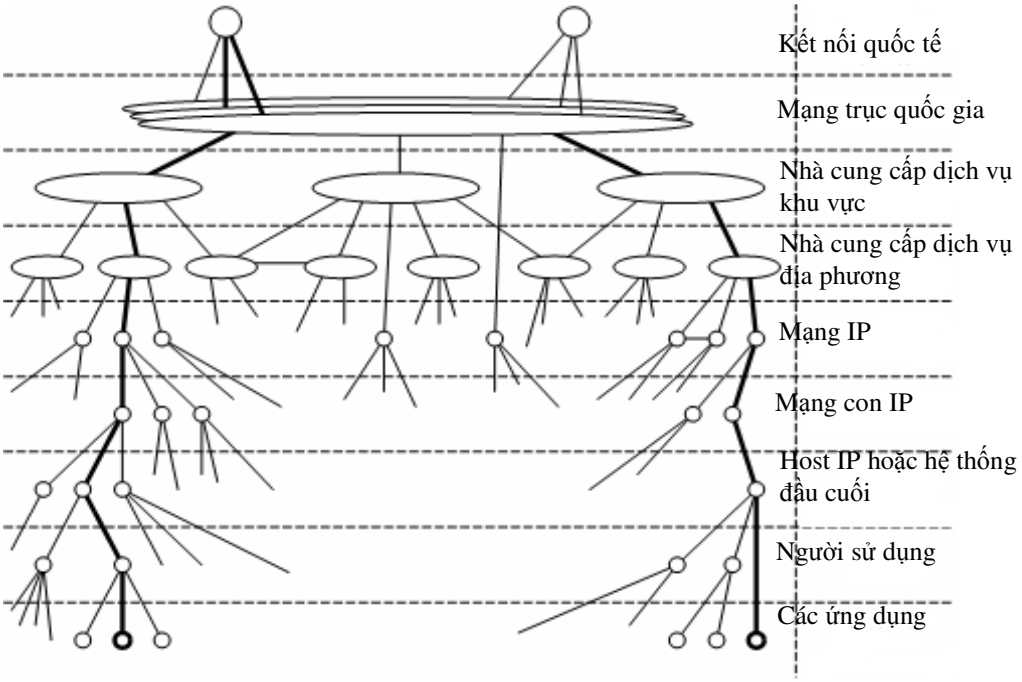
2.3. Các đặc tính lưu lượng

Để hiểu cơ chế hoạt động phức tạp của mạng yêu cầu đặc tính hoá lưu lượng như một hàm nhiều tham số. Phương pháp thống kê lưu lượng phục vụ mục đích đặc tính hoá lưu lượng mạng được chúng tôi sử dụng ở đây là phương pháp gộp chi tiết (Aggregation granularity, AG); phân tích hệ thống trên cơ sở chia hệ thành các phân hệ khác nhau (chi tiết hơn) và tổng hợp đặc tính của các phân hệ để có những đặc tính của toàn bộ hệ đó. Hình 2-1 biểu diễn một cấu trúc điển hình và phân lớp tương ứng của môi trường mạng Internet.

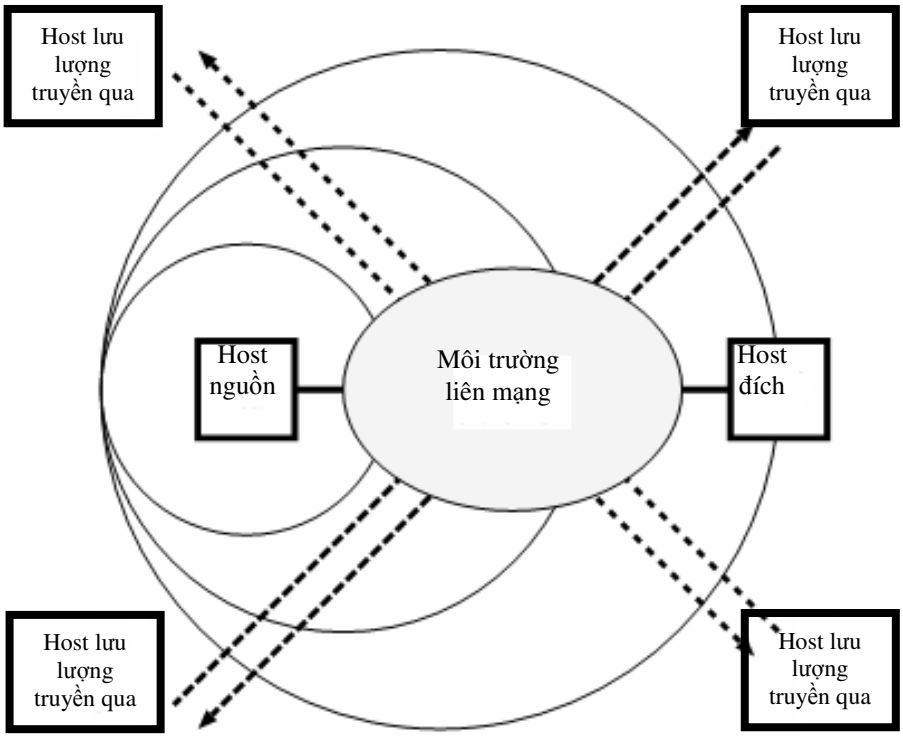
Một trong những cách thu thập các thống kê liên quan đến việc đặc tính hoá và mô hình hoá lưu lượng mạng là rời rạc hoá các thành phần; theo nghĩa "Rời rạc các thành phần để lấy thông tin từ mỗi thành phần và rời rạc các thành phần để mỗi thành phần biểu diễn được thông tin".

Việc phân tích lưu lượng thường được thực hiện theo hai cách: Phân tích hướng đầu cuối sử dụng (chia tải theo loại đầu cuối – đầu cuối), mô tả trong hình 2-2 và phân tích từ phía mạng (phân tích, tổng hợp các luồng lưu lượng của một môi trường mạng cho trước nào đó) như mô tả trong hình 2-3.

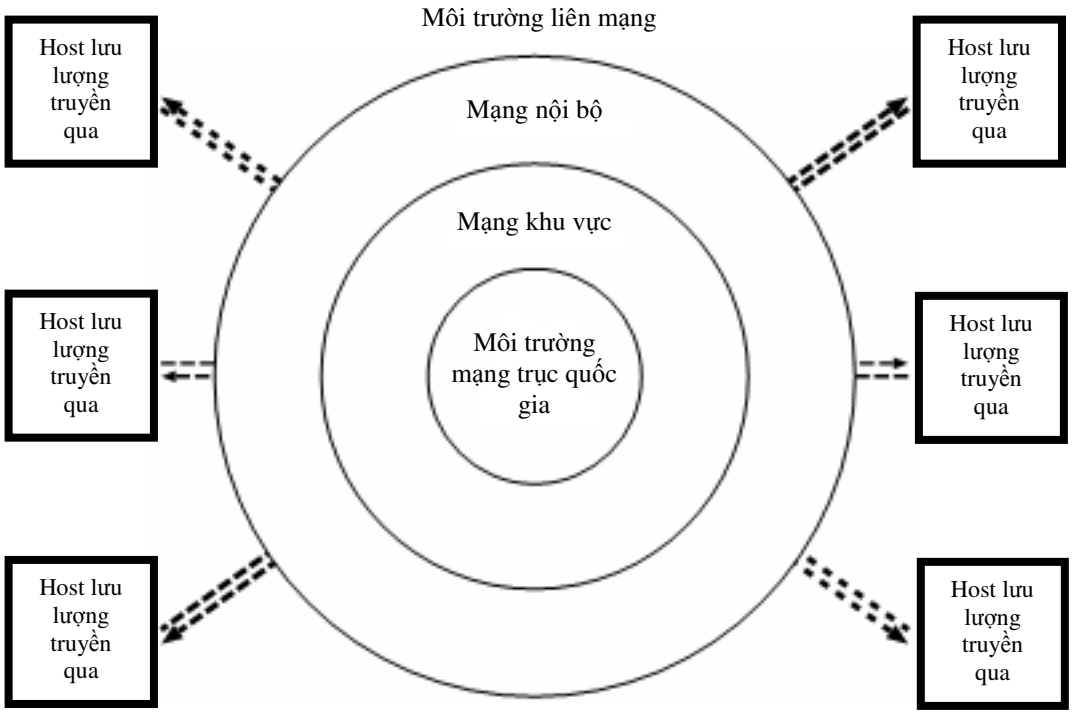
Để đánh giá toàn diện, trước hết cần phân tích khả năng đáp ứng của một điểm đầu cuối của mạng, sau đó phân tích các đặc tính lưu lượng của mạng. Việc đánh giá theo hướng đầu cuối này được chia làm nhiều loại tham số khác nhau như: độ trễ và trượt thời gian, độ mất mát gói tin và thông lượng.



Hình 2-1. Biểu diễn các phân lớp của môi trường mạng Internet



Hình 2-2. Phân tích các luồng lưu lượng ở phía đầu cuối



Hình 2-3. Phân tích các luồng lưu lượng của một môi trường mạng

2.4. Phân tích lưu lượng mạng

Sau khi thu được những kết quả cần thiết cho việc phân tích lưu lượng như trình bày ở trên, chúng tôi đã nghiên cứu và đưa ra các yếu tố điều kiện cho một phép phân tích lưu lượng mạng.

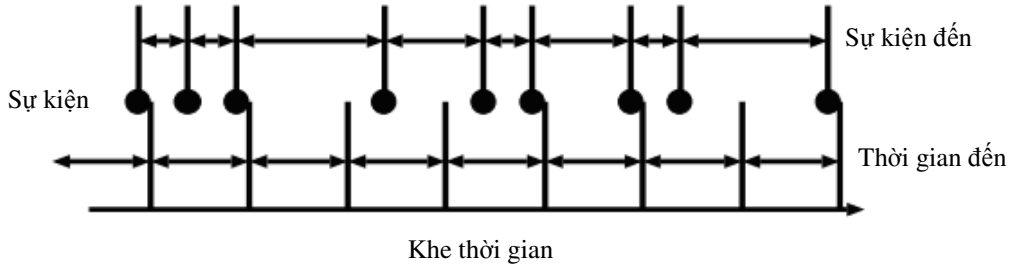
- *Sự sử dụng:* Mức độ sử dụng có thể phản ánh trong mỗi phép đo các đặc tính phân chia của mạng. Các số liệu thống kê về việc phân bố dữ liệu (giá trị trung bình, độ sai khác, tỉ lệ thống kê) chỉ ra khuynh hướng sử dụng trong khoảng thời gian nào đó. Liên quan đến khái niệm về sự sử dụng là vấn đề tắc nghẽn mạng, hạn chế về tài nguyên băng thông và năng lực chuyển mạch.

- *Khả năng tiếp cận:* Phục vụ cho việc duy trì bảng định tuyến, việc tìm kiếm trong bảng định tuyến để chuyển gói tin trong các node chuyển mạch. Các số liệu (kích cỡ của các bảng định tuyến, số mạng IP trong mạng Internet có thể định tuyến để chuyển lưu lượng) là các chỉ số phản ánh khả năng tiếp cận mạng.

- *Tính địa phương*: Liên quan đến khả năng tiếp cận của mạng chính là lưu lượng tại một điểm mạng cụ thể, phản ánh sự không đồng nhất trong việc phân bố lưu lượng về mặt địa lí. Các nhà thiết kế mạng đã rất quan tâm đến việc sử dụng bộ nhớ thực qua việc sử dụng bộ nhớ ảo cũng như bộ đệm [25]. Ví dụ, Jain đã nghiên cứu so sánh việc sử dụng bộ nhớ địa chỉ tại điểm đó (theo không gian hoặc thời gian) [26]; nhóm của Gulati cung cấp 4 đặc trưng của đặc tính địa phương hoá [27]; Estrin và Mitzel đã khám phá tính cục bộ địa phương bằng cách xét mào đầu trong các thiết bị định tuyến [31]...

- *Tính truyền (Burstiness):* Trên quan điểm phân tích mạng, liên quan đến việc phân bố chung về độ trễ, độ trượt là tính truyền của toàn bộ lưu lượng tải làm việc. Tính truyền mô tả

như trong hình 2-4, phải biểu diễn được tiến trình dữ liệu giữa các thời điểm đến (thời gian giữa các gói đến), tiến trình dữ liệu các gói đến (số gói đến trong mỗi khoảng thời gian) và những tham số bổ sung về nơi đến (các gói, các byte, các luồng, các phiên lớp giao vận).



Hình 2-4. Mô tả một quá trình điều khiển sự kiện, thời gian truyền đến lẫn nhau giữa các sự kiện trong khoảng một chu kì thời gian nhất định

- *Tải có ích (Payload):* Được hiểu là số lượng thông tin chứa trong một gói. Như vậy, tải có ích của một gói IP là nội dung của gói theo sau thông tin mào đầu IP, nhưng tải có ích của một gói *telnet* lại không tính các mào đầu TCP và IP. Đôi khi tải có ích cũng được định nghĩa một cách không rõ ràng, bao gồm toàn bộ các mào đầu gói. Dùng khái niệm gói tải có ích để chỉ hiệu quả của giao thức mặc dù phân tích chính xác về tính hiệu quả cũng phản ánh cách xử lí cuối đến cuối, bao gồm cả việc thừa nhận, sự phát lại và các chiến lược cập nhật các giao thức khác.

- *Vùng giao lưu lượng:* Nhiều đặc tính tải khác nhau khi áp dụng vào các ứng dụng khác nhau. Các ứng dụng và các giao thức vận chuyển tầng dưới khác với dạng hoạt động của nó (TCP khác với UDP, *ftp* khác từ *telnet*). Vì vậy, ngoài toàn bộ phần mô tả lưu lượng, việc xét theo từng giao thức cũng là một phần là quan trọng trong mô tả mạng.

- *Các metric lưu lượng riêng lẻ*: Nhiều nghiên cứu đã chỉ ra số liệu đo ở các khoảng thời gian, các loại mạng khác nhau ảnh hưởng đến vùng giao của lưu lượng mạng diện rộng [6, 30], [58]. Các số liệu đo tham chiếu (metric) đóng vai trò quan trọng không chỉ đối với nhiệm vụ đánh giá về vùng giao mà còn đối với việc đánh giá sự đóng góp của từng thành phần riêng của vùng giao vào lưu lượng tải tổng hợp. Khảo sát cách xử lí các ứng dụng cụ thể (chuyển tập tin trao đổi hay đăng nhập màn hình từ xa) và các giao thức truyền sử dụng ở lớp dưới (TCP, UDP, IP) cấu thành thành phần vào thực hiện nhiệm vụ thiết lập định hình và dạng (profile) lưu lượng. Nhiều nghiên, cứu lập profile lưu lượng TCP đã xét đến lưu lượng internet thông qua các điểm chuyển tiếp đến mạng diện rộng như nhóm của Caceres định hình các trao đổi Internet bằng cách đặc tính hoá các trao đổi TCP giữa các điểm đầu cuối với nhau [3]; Danzig và Jamin đã dùng kết quả của nhóm Caceres để tạo thư viện về đặc tính tải của TCP mang tính thực nghiệm sử dụng vào mô phỏng mạng [32, 60]; Paxson xem xét các kết nối TCP trong mạng diện rộng bắt nguồn từ ba cấp độ của Internet (khuôn viên trường đại học, các tổ chức nghiên cứu, vùng và môi trường cổng quốc tế) [30, 59, 60]; nhóm của Danzig tập trung lập profile của một ứng dụng đơn lẻ để phân tích lưu lượng dữ liệu phân giải tên miền (DNS) trong mạng diện rộng [8]…

- *Tổng hợp các metric lưu lượng:* Vì các mạng tích hợp đòi hỏi những thông tin về luồng đối với một số luồng được định hướng duy trì kết nối suốt thời gian sống, nên nhiều metric phục vụ cho việc tổng hợp lưu lượng đóng vai trò quan trọng cho các nhà phát triển thiết bị mạng như số lượng các luồng hoạt động trên giây, số lượng các luồng mới yêu cầu thiết lập trên mỗi giây và số lượng các luồng nhàn rỗi cần xóa.

- *Các đặc tính phụ thuộc môi trường mạng:* Ngoài cấu trúc phân cấp của các giao thức còn một thành phần phân cấp nữa của đặc tính lưu lượng là kiến trúc phân cấp kết nối Internet như trong hình 1.1. Ở đó chỉ cần phù hợp với kiến trúc phân cấp, ngay cả khi profile tải làm việc của mỗi giao thức riêng khác nhau qua các thành phần mạng khác nhau. Một số kết quả được tham khảo ở phần này gồm nghiên cứu của nhóm Acharya về đặc tính hoá có phân cấp lưu lượng tại một mạng campus bằng việc phân tích quan sát lưu lượng từ ba vùng campus theo nội mạng LAN, liên mạng LAN và lưu lượng tương ứng LAN đến WAN [33]; Wakeman và đồng nghiệp và Asaba với nghiên cứu đặc tính lưu lượng trên cơ sở hạ tầng mạng quốc tế, tương ứng với lưu lượng xuyên Đại Tây Dương [63] và Thái Bình Dương [64].

3. ĐẶC TÍNH, PHÉP ĐO VÀ THỐNG KÊ LƯU LƯỢNG

Trong phần trên, chúng tôi đã trình bày tổng quát về mạng Internet, lưu lượng Internet, một số nghiên cứu về vấn đề này cũng như tính cần thiết phải phân tích và nhận dạng luồng lưu lượng ứng dụng Internet. Các tham số hoạt động và tham số cần thiết liên quan đến chất lượng dịch vụ cũng được phân tích. Một trong những vấn đề quan trọng đối với phát triển dịch vụ là phải giám sát được và biết chính xác khả năng đáp ứng QoS của mỗi ứng dụng IP.

Trong phần này, chúng tôi trình bày về đặc tính lưu lượng IP đối với các môi trường mạng khác nhau, phương pháp thống kê và đo đạc. Tuy nhiên, các dịch vụ ứng dụng công nghệ IP được phát triển nhanh chóng nên trong phạm vi lĩnh vực nghiên cứu chúng tôi không trình bày các phân nhánh dịch vụ của công nghệ IP mà tập trung vào việc phân tích thống kê, đo đạc lưu lượng và phân tích đặc tính lưu lượng IP và các dịch vụ IP hiện thời nhằm vào việc phát triển mạng IP vô tuyến và hữu tuyến. Sau đây là một số kết quả chúng tôi đạt được.

3.1. Đặc tính lưu lượng của những mô hình dịch vụ và môi trường mạng khác nhau

3.1.1. Tương tác đa phương tiện và lưu lượng thời gian thực

Những kết quả về lưu lượng thời gian thực và đa phương tiện thu được từ các mã hóa-giả mã tốc độ Bit như lưu lượng âm thanh PCM và hình ảnh MPEG-1 đã được mô tả một cách bao quát bởi hàng loạt kĩ thuật, mô hình như các phân tích cấu trúc tự tương quan của bộ đếm cơ bản và các quá trình điểm. Với sự phát sinh của những dịch vụ tương tác mới trong môi trường đa phương tiện và Internet di động dẫn đến sự hiện diện của các phương pháp mô tả mới. Ví dụ, kĩ thuật mã hóa-giải mã thích ứng như bộ mã hóa-giải mã đa tốc độ thích ứng UMTS (AMR) và MPEG-4.

Trước hết, các đòi hỏi về chất lượng dịch vụ của những dịch vụ nổi trội nhất (hội nghị truyền hình và phương tiện di động) được trình bày. Thông qua nghiên cứu, thu thập số liệu chúng tôi đã đưa ra được khuyến nghị và khả năng cho phép đối với QoS của các dịch vụ Audio và Video.

- *Những đòi hỏi về QoS của Audio và Video*: Chất lượng QoS thể hiện ở các thông số (băng thông, trễ, Jitter và tổn thất) đều có thể đo, vẽ được lên bản đồ ứng với những đòi hỏi của ứng dụng. Những ứng dụng trên IP có thể là ứng dụng Streaming một chiều hoặc tương tác hai chiều (Audio tương tác có thể dùng để sắp xếp thành trò chơi âm nhạc, hội nghị truyền hình hay dịch vụ di động đa phương tiện) [23, 24]. Các yêu cầu đối với độ rộng băng thông điển hình cho định dạng Audio và Video khác nhau được chúng tôi phân tích và tổng hợp chi tiết trong bảng 3-1 và bảng 3-2.

Bảng 3-1. Yêu cầu về băng thông điển hình đối với những định dạng Audio khác nhau

Định dạng Audio	**Băng thông đòi hỏi**
Chuẩn mã hóa nén HE-AAC MPEG-4, âm thanh nổi, được xem là có 'chất lượng tốt' đối với hầu hết người nghe	40 - 48 kbps
Chuẩn mã hóa nén AAC MPEG-4, âm thanh nổi, được xem là rõ ràng 'không thể phân biệt được từ chất lượng CD'	128 kbps
Chuẩn nén Dolby Digital cho người xem phim ở nhà bằng định dạng DVI-Video (chuẩn 5.1)	384 kbps
Chuẩn chất lượng CD không nén, âm thanh nổi	1,4 Mbps
Chuẩn âm thanh chất lượng cao đa kênh không nén bằng định dạng DVD-Audio	9,6 Mbps
Chuẩn nén âm thanh vòm mp3, chuẩn 5.1, chất lượng đạt	128 kbps
Các chuẩn mã hóa lời nói	4 - 64 kbps

Bảng 3-2. Yêu cầu về băng thông đối với những định dạng Video khác nhau

Định dạng Video	**Băng thông đòi hỏi**
Chuẩn HDTV không nén	1,5 GBps
Chuẩn HDTV tạm thời	360 Mbps
Chuẩn SMPTE, theo tiêu chuẩn TV	270 Mbps
Chuẩn truyền hình chất lượng HDTV (MPEG-2)	19,4 Mbps
MPEG-2 (SDTV)	6 Mbps
MPEG-1	1,5 Mbps
MPEG-4	5 kbps - 4 Mbps
H.323	28 kbps - 1 Mbps

- *Đặc tính lưu lượng Audio và Video:* Đối với một nguồn Audio, việc mã hóa tạo ra những mẫu cách đều nhau, một nguồn Video thì tạo ra một cách đều đặn các khung hình trong một nhóm những hình ảnh mô tả sự phân bố các khung hình và sự phụ thuộc của chúng. Những mẫu và khung hình cần được đóng gói trước khi gửi chúng trên mạng IP và quá trình này cần được biết để luận ra đặc tính của luồng gói tin. Thêm vào đó bản thân luồng gói tin có thể được xem xét không cần quan tâm đến việc đóng gói và có thể dùng để khám phá xem sự phân cấp theo những ưu tiên mức cao ảnh hưởng như thế nào đến các gói được tạo ra tại nguồn.

Rất ít thông tin tìm được về đặc tính của lưu lượng Audio chất lượng cao vì hầu hết các nghiên cứu đều thực hiện trên VoIP. Thêm vào đó, phần lớn chỉ tập trung mô tả lưu lượng nguồn đơn từ các định dạng Audio, Video theo các phương thức khác nhau của các tuyến truyền và mã hóa quá trình đóng gói nhằm mục tiêu đánh giá hiệu suất End-to-End hoặc tại các Node vào trong tài khoản [26].

Đã có nhiều nghiên cứu mô tả lưu lượng MPEG-1/2 trong những năm 1990, nhưng không thấy nhiều nghiên cứu về mô tả lưu lượng trên chuẩn MPEG-4. Công trình của Fitzek và Reisslein cung cấp những dấu vết về kích thước khung hình tạo ra bởi MPEG-4, H.263 và cung cấp một số thông tin tham khảo quan trọng khác [25].

- *Về các tiêu chuẩn, bộ mã hóa thông dụng (codec):* Các chuẩn quan trọng nhất trong việc mã hóa Video, Audio và Voice cũng như mô tả vắn tắt chương trình Video mới nhất đã xuất hiện.

Đối với Video: Chuẩn MPEG-4 (EP02), H.264/AVC, Windows Media 9 và những kĩ thuật không chuẩn hóa khác. Một cách cụ thể:

+ MPEG-4 (EP02): Chuẩn này xuất phát từ khái niệm của kĩ thuật truyền thống (mã hóa Video hình chữ nhật, lấy mục tiêu tương tác giữa thành phần đối tượng cơ sở và mã hóa) và quá trình mã hóa thực tế là khối cơ sở, tương tự như các khối của chuẩn MPEG-1/2. Tuy nhiên, trong số các luồng khác nhau trên IP, một số lỗi và các trường hợp truyền bị trễ ảnh hưởng được chọn mục tiêu trong chuẩn.

+ H.264/AVC (H.264 hoặc bộ mã hóa Video nâng cao MPEG-4AVC): Sáng kiến chung của ITU-T và ISO, tốt hơn tất cả các chuẩn mã hóa Video trước đó về khía cạnh hiệu quả mã hóa. H.264 gồm hai lớp riêng biệt, lớp mã hóa Video (VCL) và lớp thích ứng mạng (NAL). Lớp VCL thực hiện tất cả nhiệm vụ xử lí tín hiệu, trong khi đó lớp NAL có các công cụ làm cho dữ liệu đã mã hóa phù hợp với mạng mục tiêu một cách thỏa đáng. Lớp NAL thực hiện việc đóng gói RTP theo một số các thiết kế hạn chế, gồm: (i). Phân biệt các gói tin quan trọng với các gói ít quan trọng mà không cần giải mã nội dung của từng gói; (ii). Phát hiện các gói có thể đã bị hỏng trong quá trình truyền mà không cần thiết phải giải mã dòng Bit trong gói (ví dụ như "đánh dấu gói" có thể được thực hiển bởi các Gateway); và (iii). Chi phí thấp.

+ Windows Media 9: Những Codec Video mới nhất của Microsoft, Windows Media Video 9 (WMV9), đại diện cho sự cải tiến so với sản phẩm Windows Media trước trong việc hiệu quả mã. Do tính độc quyền nên ít được biết đến các kĩ thuật của Codec này, nhưng đã được chứng minh rằng kết quả tương đương với chuẩn H.264/AVC. Phiên bản nâng cao của WMV9 cung cấp khả năng "định dạng vận chuyển độc lập", cho phép các Bitstream dựa trên chuẩn WMV9 vận chuyển và xử lí bởi những nền tảng không dựa trên Windows như MPEG-2 Transport Streams, RTP và DVD [htt]. Tuy nhiên, chi tiết của những công cụ này vẫn chưa được biết. Cùng với H.264/AVC, một biến thể của WMV9 hiện đang được xem xét cho việc chuẩn hóa với SMPTE[2].

+ Những kĩ thuật không chuẩn hóa: Gần đây, khả năng mở rộng mã hóa Video thông qua việc phân tích mã hóa theo ba chiều [27]. Đó là, các lược đồ mã hóa cho ra một Bitstream được "nhúng" và các lớp tùy ý hoặc các biến thể được trích xuất thông qua khả năng mở rộng theo cả hai SNR (phân giải không gian và thời gian). Những lược đồ mã hóa này không có cấu trúc khung thông thường I-P-B[3] (được biết qua mã hóa MPEG/H), dẫn đến tỉ lệ biến động thời gian thấp hơn và cơ hội tốt hơn để kiểm soát tốc độ điều khiển và thích nghi với các Video đã mã hóa với những điều kiện mạng khác nhau và khả năng khác nhau của thiết bị đầu cuối.

Đối với Audio và Voice: Chuẩn AMR-WB, MPEG, trễ thấp và chuẩn vòm. Cụ thể,

+ AMR-WB: Chuẩn mã hóa tiếng nói đa tốc thích nghi, băng rộng đã được chọn bởi cả 3GPP và ITU-T. Đây là chuẩn mã hóa tiếng nói đầu tiên chung cho cả các dịch vụ không và có dây được chờ đợi cho hàng loạt ứng dụng mã hóa tiếng nói, từ tiếng nói trên GSM và UMTS đến Voice trên IP. Chuẩn thực hiện 9 chế độ mã hóa tiếng nói với tốc độ bit tương ứng trên khoảng 6,6 kbps - 23,85 kbps và bộ mã hóa dựa trên mã đại số dự đoán tuyến tính kích thích (ACELP). Sự thích ứng của bộ mã hóa có ứng dụng nổi bật trong hai khía cạnh là giảm chất lượng tiếng nói/sự biến dạng đối với dung lượng mạng và truyền không liên tục thông qua việc phát hiện hoạt động giọng nói.

+ MPEG mã hóa âm thanh và vấn đề liên quan: Nén âm thanh bằng mã hóa nhận thức đạt hiệu quả hơn do sử dụng "dải phổ tái tạo" (SBR)[4], phần thừa giữa các thành phần tần số thấp và cao của phổ. Những tần số thấp được xử lí bởi bộ mã hóa nhận thức thông thường, phổ tần số cao được khôi phục lại tại phía nhận, từ phổ tần số thấp và thông tin phía bitrate thấp được ghép vào trong luồng âm thanh nén. SBR với mp3 thành mp3PRO, và với AAC thành aacPlus, đã được chuẩn hóa bởi MPEG bằng với MPEG-4 (AAC hiệu suất cao HE-AAC). HE-AAC có thể cung cấp chuẩn âm thanh 5.1 vòm tại tốc độ 128 kbps, chuẩn âm thanh Stereo chất lượng cao tại 48 kbps, thậm chí có thể sử dụng tại tốc độ thấp hơn bitrate. AAC có thể bao gồm khả năng mở rộng Bitrate và mở rộng băng thông để thích ứng với những yêu cầu truyền động khác nhau. Để đối phó với lỗi truyền dẫn, có thể sử dụng những kĩ thuật phát hiện lỗi (sử dụng chuỗi CRC), dấu lỗi (tổng hợp những phần bị mất của tín hiệu), bảo vệ lỗi và khả năng phục hồi lỗi.

+ Mã hóa trễ thấp (được thực hiện với khía cạnh trễ Codec): Chuẩn MPEG-4 bao gồm bộ mã hóa trễ thấp (bộ mã hóa có trễ xử lí một chiều trong khoảng từ 20 - 30 ms, AAC-LD). Tại 64 kbps, AAC-LD có thể cho độ rộng băng thông ở tần số 15 kHz, chất lượng tốt hơn mp3. Tại 32 kbps, chất lượng của AAC-LD có thể so sánh với AAC tại 20 kbps (có thể sử dụng xuống tốc độ 24 kbps). Hiện nay tại Fraunhofer đang thực hiện mã hóa trễ cực thấp (ULD), với các trễ một chiều thấp tới 6ms tại các tần số mẫu trong khoảng từ 32 kHz đến 48 kHz.

+ Chuẩn âm thanh vòm (Surround): Sử dụng "Mã hóa báo hiệu kép" phát triển tại Agere Systems, để mở rộng định dạng mp3 sang chuẩn âm thanh vòm (kênh 5.1), như Fraunhofer chứng minh đối với định dạng mp3 Surround. Không gian thông tin từ các kênh bổ sung được mô tả như một nhánh thông tin tốc độ thấp, sử dụng để phục hồi lại tại phía thu. Đây là loại định dạng tương thích ngược của các bộ giải mã âm thanh nổi mp3. Tốc độ dữ liệu của âm thanh vòm mp3 có thể đạt tới tốc độ của mp3 thông thường.

3.1.2. Đối với lưu lượng Web và Client-Server

Những nghiên cứu trước đã tập trung phân tích lưu lượng Client-Server trong các mạng gói tốc độ cao, đặc tính lưu lượng phân biệt 3 phân đoạn của hành vi người dùng, loại dịch vụ và cấu trúc truy cập. Truy cập Internet và lưu lượng mạng lõi được mô tả xấp xỉ bởi mô hình phân cấp để xác định những quy mô thời gian, mức của các luồng lưu lượng phát sinh từ những ứng dụng khác nhau (nơi lưu lượng Web với mô hình Client-Server cơ bản tạo thành một tỉ lệ chủ yếu). Trong mô hình phân cấp xấp xỉ, sự chồng lên nhau của Web, kết quả lưu lượng Internet của giao tiếp Client-Server cơ bản được xác định bởi các thuộc tính tương ứng với các tính năng lưu lượng thu được tại các tỉ lệ thời gian và các tầng điều khiển lưu lượng: (i). Mức người dùng, (ii). Các mức cuộc gọi và phiên làm việc, (iii). Cấp độ của trang, (iv). Mức độ kết hợp của các luồng gồm: Luồng các gói IP của các kết nối phổ biến giữa các cặp yêu cầu - đáp ứng; Sự bùng nổ của các gói trong phạm vi các luồng; và Cấu trúc gói trong phạm vi quy mô bùng nổ. Những thuộc tính vừa nêu phải được tái đánh giá và nghiên cứu đối với những dịch vụ theo mô hình Client-Server.

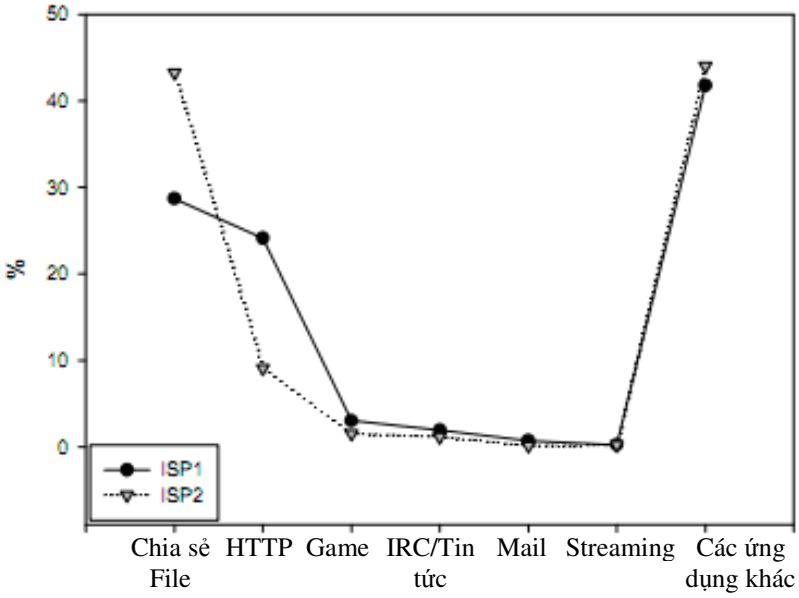
Do sự xuất hiện liên tục các ứng dụng, dịch vụ mới, kể cả công nghệ truy cập mạng băng rộng như ADSL và CATV nên mô tả lưu lượng truy cập Internet trở thành vấn đề quan trọng đối với cả nhà cung cấp dịch vụ Internet (ISPs) và khai thác mạng truy cập. Sự thay đổi nhanh về các đặc tính lưu lượng đòi hỏi các phép đo lưu lượng một cách thường xuyên và phải đáp ứng tiêu chí kĩ thuật; các phép đo lưu lượng thực hiện ở các mức phân giải khác nhau phù hợp với nhiệm vụ đã chọn làm mục tiêu và các phép đo được thực hiện trên nhiều loại khác nhau của đối tượng.

Sau khi phân tích hành vi người dùng thì nhóm lại bởi phương tiện của cụm các phương pháp phân tích (cụm các phương pháp phân tích gồm loại phân cấp và không phân cấp) [28] và đã được tổ chức như sau. Đầu tiên chúng tôi đưa ra tổng quan về các dấu vết lưu lượng đã phân tích. Sau

đó, sử dụng các thành phần chủ yếu như dữ liệu phân tích sơ bộ đến cụm phân tích. Và, xác thực kết quả cụm phân tích bằng cách sử dụng biệt số phân tích cùng với việc mô tả các cụm thu được dựa trên các số liệu thống kê. Cuối cùng là những kết luận.

Chúng tôi thực hiện các phép đo tại hai ISP khác biệt, ISP1 và ISP2. Sử dụng một mạng CATV làm ISP1, ADSL làm ISP2. Cả hai đều cung cấp một số loại hình dịch vụ đặc trưng bởi tốc độ truyền tải tối đa cho phép trong các hướng Downstream/Upstream. Đối với ISP1, các dịch vụ tương ứng là 128/64, 256/128 và 512/256 (kbs); còn đối với ISP2 là 512/128 và 1024/256. Bộ dữ liệu của ISP1 bao gồm 3432 người sử dụng và một bộ dữ liệu của ISP2 bao gồm 874 người sử dụng.

Trong ISP1 (hình 3-1), không có khác biệt lớn giữa sử dụng chia sẻ tập tin và HTTP, vì là một trường hợp trong ISP2. Có vẻ như trong ISP1, có tỉ lệ phần trăm cao hơn về người dùng là do thực hiện truyền, chia sẻ tập tin thông qua HTTP. Điều này cũng giải thích thực tế là trong ISP1, HTTP có khoảng thời gian hoạt động cao hơn chia sẻ tập tin (hình 3-2).

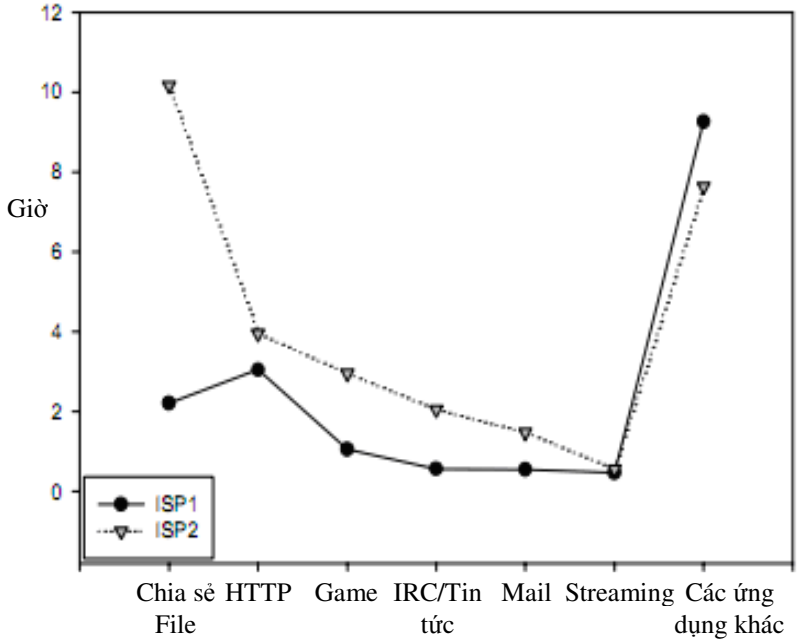


Hình 3-1. Mức độ sử dụng các ứng dụng liên quan

"Các nhóm khác" bao gồm một tỉ lệ phần trăm lưu lượng đáng kể, mặc dù thực tế cá nhân số lượng cổng chỉ định cho nhóm này đã tạo ra một vài lưu lượng (tiêu chí của chúng tôi là ít hơn 0,1% các byte tải về trong ISP1 và 0,05% trong ISP2). Với những giá trị cao của các khoảng thời gian hoạt động và tốc độ truyền, dẫn chúng tôi đến nghi ngờ rằng hầu hết những cổng này được sử dụng bởi việc chia sẻ tập tin và các ứng dụng Video, mà cuối cùng là việc phân phối lưu lượng của nó bởi một số lượng các cổng (không theo chuẩn).

Chúng tôi sử dụng phép phân tích thành phần chủ yếu (PCA) để tìm tối đa sự khác biệt giữa những tổ hợp tuyến tính của các biến (các thành phần chủ yếu) trên cơ sở dữ liệu thu được sau khi chuyển đổi: $Y_j = \ln(1 + X_j)$, với tốc độ truyền X_j (kbs) trong khoảng nữa giờ thứ j, j = 1, 2, ..., 48, đối với hai ISP (dữ liệu chuyển đổi được sử dụng trong các phân tích sau, bắt đầu với phân tích cụm để xác định các nhóm người dùng lưu lượng tương đương). Việc phân cụm đối với cả hai ISP được thực hiện bởi phương pháp Ward (phân cấp) và phương pháp Medoids (phân vùng xung

quanh, không phân cấp). Có thể giải thích những cụm đã thu được bởi cùng một cách cho cả hai ISP như mô tả trong bảng 3-1.



Hình 3-2. Trung bình khoảng thời gian hoạt động

Bảng 3-1. Giải thích các cụm của ISP1 và ISP2

Cụm	**Ý nghĩa**
C1	Tốc độ truyền cao trong tất cả các khoảng thời gian
C2	Tốc độ truyền cao/thấp trong buổi sáng/chiều
C3	Tốc độ truyền thấp trong tất cả các khoảng thời gian

Cấu trúc cụm được đánh giá bằng cách sử dụng biệt số phân tích [29] (kĩ thuật đa biến, tách các tập của các đối tượng và phân đối tượng mới vào một trong các nhóm xác định trước đó). Ở đây, biệt số phân tích được sử dụng để xác thực các cụm thu được. Trong thực tế, chúng ta có thể thu được các hàm biệt số để đưa ra những quy tắc phân loại ước lượng hay các tỉ lệ phân loại sai nhằm xác nhận tính hợp lệ của cấu trúc cụm. Bên cạnh đó, một công việc không kém phân quan trọng mà chúng tôi đề cập đến chính là việc đánh giá các cụm thu được.

Người sử dụng điển hình của mỗi cụm (C1, C2 và C3) được mô tả về ý nghĩa như sau: tỉ lệ sử dụng cao trong tất cả các khoảng thời gian trong ngày đối với C1; tỉ lệ sử dụng thấp trong nửa đầu, sử dụng cao trong nửa sau của ngày đối với C2; và tỉ lệ sử dụng thấp trong tất cả các khoảng thời gian trong ngày đối với C3. Ba cụm cũng được đánh giá trong khuôn khổ về một số đặc tính lưu lượng người dùng không sử dụng trong phân tích cụm (như số lượng lưu lượng tải về, thời lượng hoạt động và tốc độ truyền). Các kết quả cho thấy rằng cụm có các đặc tính riêng như C1 và C2 sử

dụng nhiều ứng dụng File Sharing, tại một tốc độ truyền cao hơn và trong những khoảng thời gian hoạt động dài hơn những người dùng C3, với xu hướng những người dùng C1 có cường độ mạnh hơn C2; và những người dùng C3 có lưu lượng tải về, thời lượng hoạt động trong HTTP thấp hơn những người dùng C1 và C2.

3.1.3. Đối với di động trong những môi trường mạng không dây

Những nhiệm vụ mới của dịch vụ di động và hệ thống nhúng đòi hỏi mô tả đặc tính di động, hành vi phụ thuộc vị trí của người dùng di động. Trong phần này, chúng tôi trình bày về phép đo, mô hình và phân tích các luồng trong những mạng GPRS**.** Trong đó, "Các phép biến đổi Radon và công cụ tương tự" liên quan đến những đặc tính Up-link (mã hóa kênh, số lượng các kênh dữ liệu gói Up-link (PDCHs) sử dụng trong kết nối GSM/GPRS) và "chương trình Tstat" dùng trong phân tích các luồng TCP/IP/GPRS từ dấu vết luồng TCP.

- *Về các phép đo, mô hình và phân tích luồng trong mạng GPRS*: Hiện nay, hầu như các mạng GPRS đều cung cấp khả năng đo, phân tích các dấu hiệu lưu lượng người dùng thay vì sử dụng mô phỏng với ý đồ nhắm vào đặc tính đặc biệt của lưu lượng dữ liệu di động hiện tại và tương lai. Sự phát triển này đầu tiên được thực hiện cho các công cụ thống kê (off-line) và sau đó đã được biến đổi thành những công cụ kĩ thuật (online hoặc thời gian thực).

- *Bối cảnh:* Bất kì dấu vết lưu lượng đo sử dụng trong phân tích lưu lượng thống kê phải đáp ứng một số tiêu chí về chất lượng. Chúng tôi nhấn mạnh ba tiêu chí quan trọng được đưa vào trong tài khoản sẵn sàng khi lập kế hoạch đối với bất kì loại phép đo lưu lượng nào. Trong đó, hai tiêu chí đầu đã được sử dụng trong các phép đo thử phần mềm (một dấu vết dựa trên TCPdump/Linux chẳng hạn [6]): (+) Tiêu chí đầu tiên: Mất mát gói tin do việc cài đặt phép đo phải bằng 0 (với các luồng TCP/IP/GPRS, thì mọi gói đơn thuộc về các luồng đơn mà các luồng TCP/IP/GPRS lại rất nhỏ nên những số liệu thống kê bị mất chất lượng một cách trầm trọng từ các gói bị mất). (+) Tiêu chí thứ hai: Chất lượng của các nhãn thời gian cao (các nhãn thời gian được gọi một cách chính xác nếu ngẫu nhiên hóa chúng, ví dụ phương sai nhỏ chẳng hạn). (+) Tiêu chí thứ ba: Dấu vết trình bày tương xứng với các nhu cầu thống kê từ phân tích dữ liệu dấu vết (tiêu chí này đạt được khi điểm đo và chiều dài của quá trình thu thập dữ liệu được lựa chọn một cách thích đáng).

Phương pháp thống kê (biến đổi Radon) đã được sử dụng trong việc phân tích tỉ lệ giới hạn liên kết ATM [5] và được áp dụng để thu nhận các thông tin về các đặc tính Up-link như chương trình mã hóa kênh, số lượng kênh dữ liệu gói Up-link (PDCHs) trong kết nối GSM/GPRS [18].

- *Những biến đổi Radon và công cụ tương tự*: Dùng cách thức theo nghĩa thống kê của biến đổi Radon mà [5] và [18] đã sử dụng để phát triển công cụ phân tích các quá trình đánh dấu điểm (các nhãn thời gian) và dấu hiệu các gói (kích thước các gói) từ không gian chung nhằm thu thông tin của những điểm (thời gian của gói quá trình giữa hai sự kiện tới) với tính chính xác cao về các nhãn thời gian và khả năng giữ lại chủ yếu mọi gói đơn.

Áp dụng biến đổi Radon dựa trên giải pháp kĩ thuật giải thích trong [18] đòi hỏi việc tạo bộ đệm trong đường đi của các gói (ví dụ, tắc nghẽn, không quá nặng) và đòi hỏi phải tồn tại các gói kích thước khác nhau. Biến đổi Radom là một công cụ sử lí Offline nhưng bổ sung các thông tin cần thiết liên quan đến mạng lưới thì có thể sử dụng như một công cụ xử lí Online [33].

Nếu có thể tìm được các đặc tính liên kết mạng từ thông tin mức gói ở dạng đồ thị hàm số Lomb, phương pháp hữu hiệu mới về thống kê đối với bài toán ở trên cũng như các phương pháp xử lí tín hiệu khác trong bài toán này có thể áp dụng [23, 31].

Cuối cùng, những đặc tính liên kết vô tuyến có thể chỉ được thấy tại GGSN đối với trường hợp Up-link. Tuy nhiên, việc mã hóa kênh sử dụng trong giao diện vô tuyến phụ thuộc chủ yếu vào

khoảng cách giữa Host di động và trạm cơ sở. Do đó, có thể giả định được rằng trong Down-link mã hóa kênh giống nhau hơn so với trong trường hợp Up-link và có thể đôi khi thu được những thuộc tính khác của Down-link từ các thuộc tính phản hồi của TCP.

- *Chương trình Tstat* [26, 37]: Một công cụ kĩ thuật xây dựng trên cơ sở thống kê áp dụng đối với dữ liệu thỏa mãn ba tiêu chí đã đề cập phía trên (các số liệu thống kê mức luồng TCP từ một dấu vết mức gói) có thể sử dụng để phân tích các luồng TCP/IP/GPRS Online.

Khi xử lí dữ liệu gói, Tstat duy trì một danh sách của tất cả các kết nối và khi phân tích gói TCP/IP tiếp theo, xây dựng lại tình trạng kết nối TCP đáp ứng:

(+) Nếu việc mở kết nối bắt tay ba chiều và việc đóng kết nối được tiến hành một cách đúng đắn bởi các thống kê kết nối, thì 92 thuộc tính khác nhau được tính toán và ghi vào một dòng trong tập tin "Log_Complete".

(+) Ngược lại, dòng này được viết vào trong tập tin "Log_noComplete", nhưng vẫn chứa những thông tin hợp lệ về địa chỉ IP, các cổng TCP và số lượng các gói SYN.

Với sự hỗ trợ về việc mô tả các mạng con, đối với mỗi số liệu thống kê luồng thu thập không chỉ phân biệt được giữa các Client và các Server mà còn có thể xác định giữa các Host di động và Internet. Tuy nhiên, do cũng có lưu lượng P2P trong các kết nối GPRS nên mô hình Client-Server có thể trở thành lỗi thời.

3.2. Mô hình cấu trúc mạng và phương pháp cắt lớp mạng

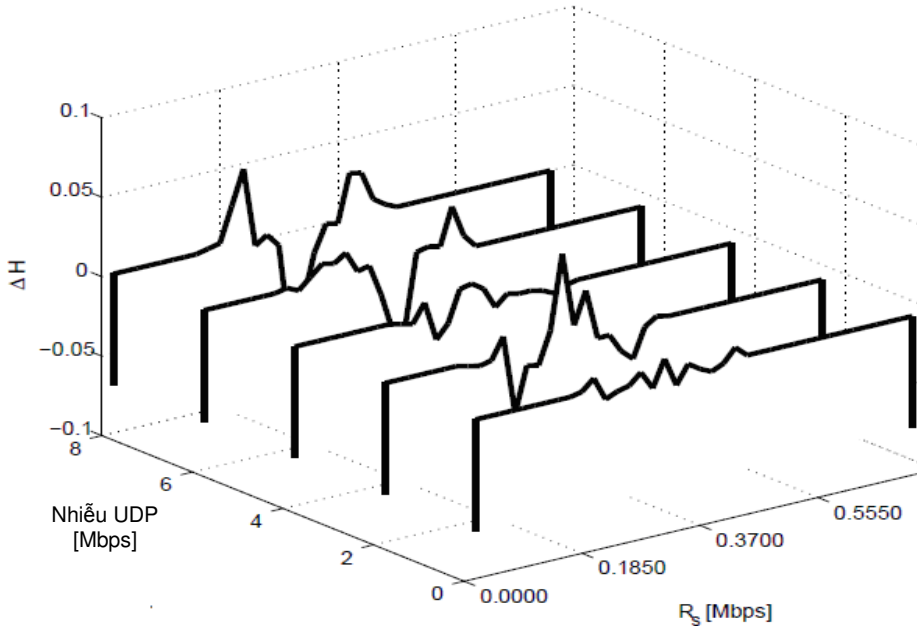
Được trình bày trong phần này là các vấn đề liên quan đến những kết luận về tình trạng mạng lưới dựa vào các phép đo (chẳng hạn, kết luận về băng thông tắc nghẽn, băng thông sẵn có, về lưu lượng chéo nhau, về các ma trận lưu lượng, v.v.) bằng cách sử dụng số liệu thống kê thông lượng (xác định suy giảm QoS của End-to-End). Để xác định hành vi thay đổi, "Những đồ thị khác nhau của biểu đồ thông lượng" được trình bày trước tiên làm cơ sở để bàn về một phương pháp xác định sự thay đổi hành vi của các tắc nghẽn trên cơ sở các phép đo về thống kê thông lượng tại cả phía nhận và phía thu của mạng.

3.2.1. Những đồ thị khác nhau của biểu đồ thông lượng

Tham số hiệu suất truyền qua ở mức gói tin và nhận thức của người dùng được chọn làm cơ sở để xác định sự thay đổi hành vi. Những khái niệm thông lượng truyền thống (thời gian trung bình để nhận một luồng đầy đủ các gói tin) được mở rộng thành biến thể ngắn hạn đo được trong khoảng thời gian trung bình tương đối nhỏ ΔT (giữa 100 ms và 1 s) trong khoảng một cửa thời gian hoặc khoảng cách quan sát ΔW (bậc phút).

Do đó, n = ΔW/ΔT giá trị của một chuỗi thời gian thông lượng $\{R_S\}_{S=1}^{n}$. Đối với luồng gói tin, thực hiện các phép đo thông lượng tại đầu vào, đầu ra của mạng thì thu được chuỗi thời gian thông lượng $\{R_S^{in}\}_{S=1}^{n}$ và $\{R_S^{out}\}_{S=1}^{n}$ từ những dấu vết gói. So sánh chuỗi $\{R_S^{in}\}_{S=1}^{n}$ và $\{R_S^{out}\}_{S=1}^{n}$, có thể thấy ảnh hưởng của tắc nghẽn trên thông lượng đã quan sát. Tuy nhiên, trong thực tế, dạng biểu đồ thông lượng $H\left(\{R_S^{in}\}_{S=1}^{n}, \Delta R, \Delta T, \Delta W\right)$ và $H\left(\{R_S^{out}\}_{S=1}^{n}, \Delta R, \Delta T, \Delta W\right)$ thường được dùng với độ phân giải thông lượng ΔR. Thực nghiệm chỉ ra rằng việc so sánh hoạt động tốt cho những khoảng thời gian $\lceil R_{max}/\Delta R + 1 \rceil \simeq 20$ đối với $n \geq 600$. Hình 3-3 cho thấy những biểu đồ

thông lượng từ các phép đo đối với ứng dụng truyền hình hội nghị theo những mức nhiễu loạn lưu lượng truyền qua khác nhau và các hiện tượng tắc nghẽn cục bộ khác nhau.



Hình 3-3. Thông lượng của Video theo các mức nhiễu khác nhau với tắc nghẽn cục bộ

3.2.2. Định hình và phân chia tắc nghẽn

- Tại một mức nhiễu thấp (khoảng 2 Mbps lưu lượng truyền qua), chúng tôi quan sát những giá trị dương của ΔR cho cả tốc độ thấp và tốc độ cao, nhưng đối với những giá trị dương của ΔH thì chỉ cho những tốc độ điển hình. Vì vậy, lưu lượng gói tin tại đầu ra truyền một cách đều đặn hơn so với tại đầu vào (hiện tượng tắc nghẽn hoạt động giống như phép nén đối với mỗi tốc độ điển hình). Điều chú ý ở đây là một vài đồ thị khác của biểu đồ thông lượng có dạng gần giống chữ "W" và trong trường hợp cụ thể này, người dùng cảm thấy không ảnh hưởng đến chất lượng của Video.

- Tại một mức cao của nhiễu (6 - 8 Mbps), quan sát thấy dạng đảo ngược của những đồ thị khác nhau của biểu đồ thông lượng so với trước gần giống như chữ "M" với các giá trị dương của ΔH đối với cả những tốc độ thấp và cao và các giá trị âm của ΔH cho những tốc độ điển hình. Sở dĩ như vậy là vì tăng các luồng tốc độ thấp do việc hình thành một hàng đợi và tăng các luồng tốc độ cao phản ánh việc giải phóng hàng đợi tại đầu ra của tắc nghẽn (người dùng tại 8 Mbps không cảm nhận thấy giảm về chất lượng tại nhiễu 6 Mbps). Mức độ suy giảm chất lượng nghiêm trọng được phản ánh tại những điểm cực trị (tiểu, đại) thể hiện trên đồ thị với ứng dụng tương ứng và trong trường hợp đó thì không thể đối phó với những vấn đề QoS của mạng lưới nữa.

- Khi trung bình đầu vào vượt quá trung bình dung lượng (lưu lượng khó được đệm vào nữa) thì tắc nghẽn quá tải xuất hiện, đồ thị của biểu đồ thông lượng đáp ứng khác nhau có dạng chữ "N"; tại các giá trị âm của ΔH thấy tốc độ tắc nghẽn không thể đối phó và tại những giá trị dương của ΔH thấy tốc độ điển hình được hỗ trợ bởi tắc nghẽn. Do vẫn tồn tại quá tải, bộ đệm giới hạn,

luồng dữ liệu mất mát đáng kể và có lẽ do thực tế này, hội nghị truyền hình "chết" tại nhiễu 10 Mbps.

- Tắc nghẽn không xác định được**:** Trái với những trường hợp đã trình bày, không thấy được loại tắc nghẽn một cách dễ dàng trên đồ thị của những trường hợp còn lại (nhiễu của 0 và 4 Mbps). Ở trường hợp 0 Mbps, hiển thị sự tồn tại của những thay đổi; tuy khá nhỏ so với những tình huống đã mô tả ở trên. Trường hợp 4 Mbps, thấy những thay đổi lớn hơn trong dạng đồ thị của biểu đồ thông lượng (kiểu như hỗn hợp của chữ "W" và "M"), không đột ngột như dạng thay đổi tắc nghẽn khi tăng nhiễu từ 2 Mbps lên 6 Mbps.

- Gần đây, viện công nghệ Blekinge và đại học Würzburg cùng nghiên cứu qua các phép đo về "tắc nghẽn thực" (liên kết giữa hai bộ định tuyến với tốc độ bit điều hướng). Thêm vào đó, tác động trung bình cửa sổ ΔT, độ phân giải thông lượng ΔR cũng như kịch bản tương tự trong môi trường mô phỏng OPNET cũng được nghiên cứu.

4. NHỮNG BÀN LUẬN VÀ ĐỊNH HƯỚNG NGHIÊN CỨU TIẾP

Trong công trình này gồm các kết quả nghiên cứu về đặc tính hóa luồng lưu lượng IP mạng diện rộng trên cơ sở áp dụng phương pháp ngẫu nhiên hiện hành đối với dữ liệu thống kê tại đơn vị quản trị, cung cấp dịch vụ viễn thông. Vấn đề liên quan đến tương tác đa phương tiên, xu thể tất yếu của các dịch vụ mới cũng được chúng tôi trình bày.

Rõ ràng, luồng lưu lượng IP truyền thông trên môi trường phân tán, không đồng nhất ở đây là đối tượng nghiên cứu nằm trong khuôn khổ của bài toán phân tích tín hiệu. Có thể lí giải bản chất vật lí của "luồng lưu lượng IP" đó như luồng tín hiệu được điều chế bởi một quá trình điều chế đa hình thể, biến dạng theo môi trường, không theo bất kể quy luật nào. Vì vậy, để nắm bắt, đặc tính hóa luồng lưu lượng, các nhà khoa học đi trước đã buộc sự biến dạng của luồng một cách ngẫu nhiên, dùng công cụ đối với các sự kiện ngẫu nhiên là phương pháp khoa học đúng đắn, tất yếu. Nhưng, các kết quả đó chỉ là định tính theo nghĩa "từng thời điểm cụ thể, sự kiện nào đó" vì chưa có phân bố ngẫu nhiên nào được đông đảo xác nhận là phân bố chung đối với mọi sự kiện, thời điểm này. Có hai câu hỏi liên quan đến vấn đề đặc tính đường bao điều chế luồng tín hiệu để xem liệu những công cụ sẵn có đã đáp ứng nhu cầu về tính chính xác đòi hỏi của công việc (?) và để định lượng chính xác của quá trình đặc tính hóa thì cần những giải pháp gì nữa (?).

Hai câu hỏi trên rất tường minh nếu tách vai trò cung cấp dịch vụ với quản trị mạng. Sở dĩ như vậy là vì cung cấp dịch vụ trong cấu trúc phức tạp, liên mạng cũng chỉ đòi hỏi chiến lược ưu tiên, đảm bảo QoS đối với dịch vụ và khai thác tối đa tiềm năng mạng. Nhưng, khi gắn với nhiệm vụ của nhà quản trị mạng, ngoài đảm bảo QoS, khai thác tiềm năng thì nhiệm vụ an ninh mạng, an ninh thông tin đòi hỏi về các phát hiện dạng lưu lượng có đặc tính khác (đặc tính "tấn công" của "tin tặc" và "tin rác"). Điều này đòi hỏi việc phân tích tín hiệu theo không gian nhiều chiều, áp dụng phân bố ngẫu nhiên nhiều chiều tương ứng để phát hiện thêm các tham số trong quá trình đặc tính hóa luồng tín hiệu IP [1].

Nhìn từ phương diện đảm bảo QoS đối với dịch vụ IP tương ứng trong bối cảnh tài nguyên hữu hạn, bài toán bảo lưu chiến lược ưu tiên được đề xuất làm cơ sở lưu giữ những động học có đóng góp tương ứng với dịch vụ IP ưu tiên vào quá trình xác lập quan hệ vào ra của hệ thống. Có thể xem bài toán bảo lưu chiến lược ưu tiên vừa nói theo hai hướng. Đó là, áp dụng bài toán điều khiển bền vững [2] để bảo lưu chiến lược ưu tiên trong quá trình phân bổ tài nguyên trước tác động của các dịch vụ khác, nhất là trước sự tấn công của các dạng tín hiệu khác thường. Và, áp dụng bài toán giảm bậc bộ điều khiển [3] để bảo lưu dịch vụ theo mức xác định ưu tiên trong

khi cắt bỏ những lưu lượng đặc trưng cho các loại dịch vụ khác và thực hiện quy trình "bền vững hóa" nhằm chống lại sự tác động của những luồng lưu lượng bất thường.

Các vấn đề bàn luận phía trên đã được chúng tôi đưa vào nội dung nghiên cứu cả trước mắt và dài hạn. Những kết quả về xác định mức ưu tiên của dịch vụ thông qua đặc tính hóa lưu lượng là nội dung trong công trình tiếp theo.

TÀI LIỆU THAM KHẢO

1. Nguyễn Ngọc San, Hoàng Ứng Huyền, Hoàng Minh, Nguyễn Gia Thái - Các giải pháp khoa học và công nghệ với mạng viễn thông Việt Nam, NXB Bưu điện, Hà Nội, 2008.
2. Nguyễn Thúy Anh, Hoàng Minh, Nguyễn Ngọc San - Ước lượng tham số mô hình hệ động học, NXB Khoa học và Kỹ thuật, Hà Nội, 2008.
3. Nguyễn Ngọc San - Nhận dạng các hệ thống tuyến tính liên tục, NXB Khoa học và Kỹ thuật, Hà Nội, 2006.
4. D. R. Boggs, J. C. Mogul, and C. A. Kent - Measured capacity of an Ethernet: Myths and reality, Proceed. ACM SIGCOMM '88, 1988, pp. 222-234.
5. R. Gusella - A measurement study of diskless workstation trac on an Ethernet, IEEE Trans. Communications **38** (1990) 1557-1568.
6. R. Caceres, P. Danzig, S. Jamin, and D. Mitzel - Characteristics of wide-area TCP/IP conversations, Proceed. ACM SIGCOMM '91, 1991, pp. 101-112.
7. A. Schmidt and R. Campbell - Internet protocol trac analysis with applications for ATM switch design, Computer Communiations Review **23** (1993) 39-52.
8. P. B. Danzig, K. Obraczka, and A. Kumar - An analysis of wide-area name server trac, Proceed. ACM SIGCOMM '92, 1992.
9. R. Wilder, K. Thomson - Wide-area internet traffic patterns and characteristics, IEEE Trans. Network **43** (1997) 1538-1543.
10. R. Chow, S. F. Hussaini - Performance analysis and traffic characterization of an ethernet campus network to identify and develop possible SMDS applications and scenarios, Proceed. IEEE Southeastcon '92, 1992, pp. 14-21.
11. S. Floyd and V. Jacobson - On trac phase eects in packet-switched gateways, Internet working Research and Experience **3** (1992) 115-156.
12. S. Floyd and V. Jacobson - The synchronization of periodic routing messages, Proceed. ACM SIGCOMM '93, 1993, pp. 33-44.
13. L. Zhang, S. Shenker, and D. D. Clark - Observations on the dynamics of a congestion control algorithm: The eects of two-way trac, Proceed. ACM SIGCOMM '91, 1991, pp. 133-148.
14. J. Mogul - Observing TCP dynamics in real networks, Proceed. ACM SIGCOMM '92, 1992, pp. 305-317, Aug.
15. Z. Wang and J. Crowcroft - Eliminating periodic packet losses in the 4.3 Tahoe BSD TCP congestion control algorithm, Computer Communications Review, 1992.
16. D. Sanghi, A. Agrawala, O. Gudmundsson, and B. Jain - Experimental assessment of end-to-end be-havior on the Internet, Proceed. IEEE INFOCOM'93, 1993, pp. 867-874.

17. A. K. Agrawala and D. Sanghi - Network dynamics: an experimental study of the Internet, Proceed. GLOBECOM '92, 1992.
18. V. Rutenburg and R. G. Ogier - Fair charging policies and minimum-expected-cost routing in internets with packet loss, Proceed. IEEE INFOCOM'91, 1991, pp. 279-288.
19. B. Kumar - Eect of packet losses on end-user cost in internetworks with usage based charging, Computer Communiations Review, 1993.
20. L. Zhang - Virtual clock: a new trac control algorithm for packet switching networks, Proceed. ACM SIGCOMM '90, 1990, pp. 19-29.
21. ITU-T Recommendation G.107 - The E-model, a computational model for use in transmission planning, ITU-T, 2003.
22. ITU-T Recommendation G.114 - One-way transmission time, ITU-T, 2003.
23. F. H. P Fitzek and M. Reisslein - MPEG-4 and H.263 video traces for network perfor-mance evaluation. Network, IEEE Trans. Network **35** (6) (2001) 640-654.
24. A. Nevin, Y. Jiang, B. Libæk, S. Johansen, O. I. Hillestad, V. Nicola, P. Svensson, and P. J. Emstad - Traffic management in a multi-provider context: State of the Art, deliverable no: D.JRA.2.2.1, 3.5: Service differentiation for audio and video streams, EuroNGI-T, 2004.
25. S-T. Hsiang and J. W. Woods - Embedded video coding using invertible motion compensated 3-D subband/wavelet filter bank, Signal Processing: Image communication **16** (3) (2001) 538-543.
26. G. J. McLachlan - Discriminant Analysis and Statistical Pattern Recognition: wide area TCP conversations, IEEE Trans Network **29** (8) (1995) 821-824.
27. D. Estrin and D. Mitzel - An assessment of state and lookup overhead in routers, Proceed. IEEE INFOCOM'92, Florence, Italy, 4-8 May, 1992, pp. 2332-2342.
28. P. B. Danzig, S. Jamin, R. Caceres, D. J. Mitzel, and D. Estrin - An empirical workload model for driving wide-area TCP/IP network simulation, Internetworking: Research and Experience **3** (1) (1991) 38-45.
29. M. Acharya, R. Newman-Wolfe, H. Latchman, R. Chow, and B. Bhalla - Real-time hierarchical traffic Modelling Techniques and Tools for Computer Performance Evaluation characterization of a campus area network, Proceed Sixth Inter. Conf., 1992.
30. M. Acharya and B. Bhalla - A ow model for computer network trac using real-time measurements, Proceed. Second Inter. Conf. Telecom. Systems, Modeling and Analysis, Nashville, TN, 1994, pp. 24-27.
31. J. Mogul - Network locality at the scale of processes, Proceed. ACM SIGCOMM '91,Zurich, Switzerland, 1991, pp. 273-285.
32. Y. Rekhter and B. Chinoy - Injecting inter-autonomous system routes into intro-autonomous system routing: A performance analysis, Internetworking Research and Experience **3** (1992) 198-202.
33. Y. Rekhter - Forwarding database overhead for inter-domain routing, ACM Computer Communiations Review **23** (1993) 66-81.
34. D. Estrin, Y. Rekhter, and S. Hotz - Scalable inter-domain routing architecture, Proceed. ACM SIGCOMM '92, 1992, pp. 40-52.

35. D. Estrin, J. Mogul, G. Tsudik, and K. Anand - Visa protocols for controlling inter-organizational datagram ow, IEEE Trans. Selected Areas Communications **7** (4) (1989) 486-98.
36. M. Steenstrup - An architecture for inter-domain policy routing, Internet Request for Comments Series RFC 1478, 1993.
37. M. Steenstrup - Inter-domain policy routing protocol specification and usage: version 1, Internet Request for Comments Series RFC 1479, 1993.
38. C. Topolcic - Experimental Internet stream protocol: version 2 (ST-II), Internet Request for Comments Series RFC 1190, 1990.
39. D. Ferrari and D. C. Verma - A scheme for real-time channel establishment in wide-area networks, IEEE Trans. Selected Areas Communications **8** (3) (1990) 368-379.
40. D. Verma, H. Zhang, and D. Ferrari - Guaranteeing delay jitter bounds in packet switching networks, Proceed. IEEE Conf. Communications Software: Communications for Distributed Applications and Systems, 1991, pp. 35-43.
41. A. Lazar and G. Pacici - Control of resources in broadband networks with quality of service guarantees, IEEE Communications Magazine **29** (10) (1991) 66-73.
42. D. D. Clark, S. Shenker, and L. Zhang - Supporting real-time applications in an integrated services packet network: Architecture and mechanism, Proceed ACSIGCOMM ' 92 , 1992, pp. 14-26.
43. L. Zhang, S. Deering, D. Estrin, S. Shenker, and D. Zappala - RSVP: a new resource reservation protocol, IEEE Trans Network **17** ( ) 8-18.
44. D. J. Mitzel, D. Estrin, S. Shenker, and L. Zhang - An architectural comparison of ST-II and RSVP, Proceed. IEEE INFOCOM'94, 1994.
45. W. Willinger - Variable-bit-rate video trac and long-range dependence, IEEE Trans. Communications **12** (2) (1994) 233-238.
46. H. Hees and D. Lucantoni - A Markov modulated characterization of packetized voice and data traffic and related statistical multiplexer performance, IEEE Selected Areas in Communication **4** (1986) 856-868.
47. H.-W. Braun and K. Clay - Network analysis in support of Internet policy requirements, Proceed. INET '93, 1993, pp. FAC:1-11.
48. V. Paxson - Empirically-Derived Analytic Models of Wide Area TCP Connections, IEEE/ACM Trans. Networking **14** (5) (1994) 524-529.
49. V. Paxson - Empirically-Derived Analytic Models of Wide Area TCP Connections: Extended Report, Technical Report LBL-34086, 1993.
50. I. Wakeman, D. Lewis, and J. Crowcroft - Trac analysis of trans-Atlantic trac, Proceed. Inet '92, 1992, pp. 417-430.
51. T. Asaba, K. Clay, O. Nakamura and J. Murai - An analysis of international academic research network trac between Japan and other nations, Proceed. Inet '92, 1992, pp. 431-440.
52. B. Chinoy and H.-W. Braun - The National Science Foundation Network, Technical report GA-A21029, 1992.
53. B. Chinoy and P. Smith - Final version of Aborted T3, ANS Update, 1992.

54. Y. Rekhter - NSFNET backbone SPF-based Interior Gateway Protocol, Internet Request for Comments Series RFC 1074, 1992.
55. ANS - ARTS: ANSnet Router Trac Statistics software, 1992.
56. Management Information Base for network management of TCP/IP-based internets, MIB-II, Internet Request for Comments Series RFC 1213, March 1991.
57. K. Clay, G. C. Polyzos, and H.-W. Braun - Application of sampling methodologies to network traffic characterization, Proceed. ACM SIGCOMM '93, pp. 194-203, 1993.
58. N. K. Groschwitz and G. C. Polyzos - A time series model of long-term NSFNET backbone traffic, Proceed. IEEE Inter. Conf. Communications (ICC '94), 1994.
59. Network information services." Data available on nis.nsf.net:/nsfnet/statistics.
60. E. Gerich - Guidelines for management of IP address space, Internet Request for Comments Series RFC 1366 obsoleted by RFC 1466, 1992.

SUMMARY

DIFFERENT OBSERVATIONS OVER CHARACTERIZATION OF IP TRAFFIC FLOWS

In a distributed environment with multiple form applications and a limited system resources the question of ensuring quality of services (QoS) from service providers (SP) to users becomes a more difficult task. With respect to the mentioned question, network operators and administrators have to understand the system dynamics through measuring IP traffic for characterizing patterns of IP traffic flows so that the ranking services can be set up for priority control strategy.

Starting from actual needs, a research topics entitled "An IP traffic flows characterization" has been being carried out with the aim of establishing a system of constrained conditions on the basics of each IP traffic characteristics. The system of constrained conditions would be exploited for network security and information surveillance, quality of services (QoS) assurance as well.

In this paper, a report in two paraghaphs is made on different important observations over characteristics of Internet traffic flows, excepting a brief introduction on scope of the work in the first paragraph and discussions on further research in the fourth one. In the second paragraph, a brief on the characterization of IP traffic flows is reviewed from internet providers viewpoint. In the third one, important characteristics of different IP trafics in scattering medium are sketched out from the QoS assurance point of view.

Different suggestions for carrying out further research from the point of network security, and information surveillance are discussed in the fourth paragraph. With respect to these points, it is important to address that a multi-dimensional characterizing IP traffic flows is required to be adopted.

Liên hệ với tác giả:

Email: *Anhdt.hue@gmail.com*
hoangminh@ptit.edu.vn