

# ĐIỀU KHIỂN LƯU LƯỢNG TRONG MẠNG ATM

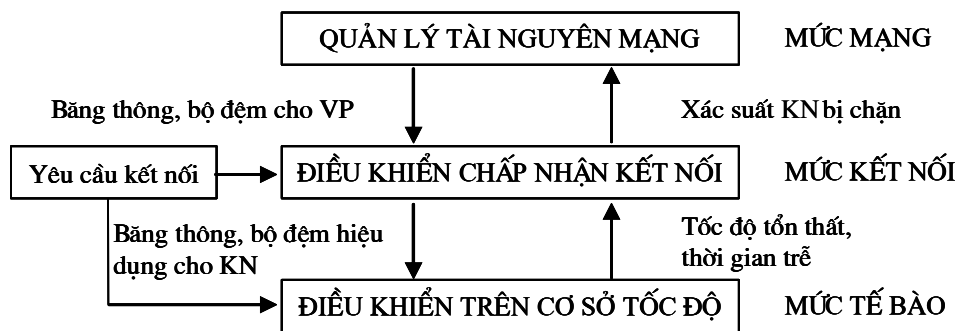
ĐINH MINH SƠN

Học viện Công nghệ Bưu chính Viễn thông

**Abstract.** This paper presents the structure and methods of controlling the traffic in ATM (Asynchronous Transfer Mode) networks. The advantages and disadvantages are analysed and several improvements are also proposed.

**Tóm tắt.** Bài báo trình bày cấu trúc và phương pháp điều khiển lưu lượng trong mạng ATM. Các ưu, khuyết điểm được phân tích và một số cải tiến cũng được đề xuất.

## 1. CẤU TRÚC ĐIỀU KHIỂN LƯU LƯỢNG PHÂN LỚP



Hình 1. Cấu trúc điều khiển luồng phân lớp

Theo chức năng có thể chia điều khiển lưu lượng trong mạng ATM thành:

- + Quản lý tài nguyên mạng NRM (Network Resource Management).
- + Điều khiển chấp nhận kết nối/cuộc gọi CAC (Connection Admission/Call Control).
- + Điều khiển trên cơ sở tốc độ RBC (Rate Based Control).

Hoặc, theo quan điểm phân cấp trong quản lý thì điều khiển luồng được chia ra: mức mạng, mức kết nối/cuộc gọi và mức tế bào. Cấu trúc này cần thiết cho các hệ thống phức tạp, đa dạng về chủng loại thiết bị, trải ra trong một không gian rộng lớn, làm việc liên tục 24/24 giờ, thường xuyên chịu sự thăng giáng lớn về phụ tải và nhiễu như các hệ thống phục vụ công cộng điển hình là mạng viễn thông.

## 2. QUẢN LÝ TÀI NGUYÊN MẠNG

Gồm các chức năng chỉ định băng thông và bộ đệm cho các đường ảo. Trong nhiều tài liệu vấn đề này được giải quyết theo cách tiếp cận tối ưu minmax ở (11). Để nghiên cứu một cách hệ thống, ở đây chúng tôi thành lập và phân tích bài toán chỉ định băng thông tối ưu cơ bản.

**Bài toán chỉ định băng thông tối ưu**

Gọi  $\lambda_{ni}$ ,  $\mu_{ni}$  và  $h_{ni}$  là tốc độ đến, tốc độ phục vụ và tốc độ đình của dịch vụ lớp thứ  $i$  trong đường ảo  $VP$  thứ  $n$ , trong đó  $n = 1, 2, \dots, N_{vp}$ ;  $i = 1, 2, \dots, M_n$ .  $RNM$  cần cấp băng thông  $C_{vpn}$  cho các đường ảo  $VP_n$  để truyền tải lưu lượng của các dịch vụ sao cho tổng xác suất kết nối bị chặn là cực tiểu

$$\min_{C_{vpn}} \left\{ \sum_{n=1}^{N_{vp}} B_{pn} \right\} \tag{1}$$

với ràng buộc

$$\sum_{n=1}^{N_{vp}} C_{vpn} \leq C_l \tag{2}$$

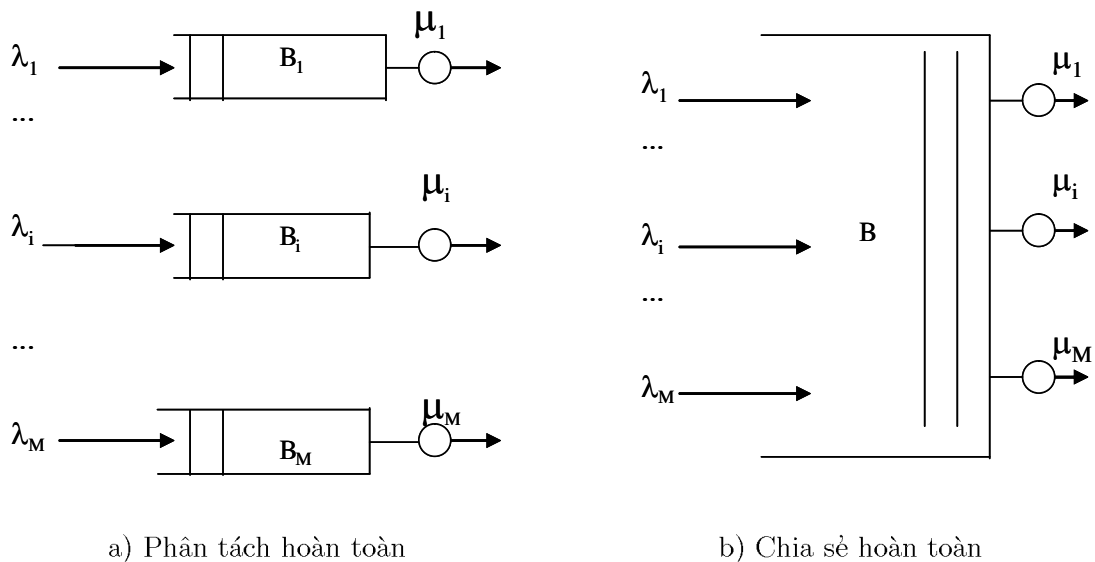
trong đó  $C_l$  là dung lượng truyền của tuyến (link)  $l$ ,  $B_{pn}$  là xác suất kết nối bị chặn của các lớp dịch vụ  $i$  trong đường ảo  $VP_n$

$$B_{pn} = \sum_{i=1}^{M_n} B_{pni} \tag{3}$$

Đây là bài toán tối ưu phi tuyến với các ràng buộc tuyến tính. Nếu  $B_{pn}$  là hàm khả vi của  $C_{vpn}$  thì ta có thể áp dụng phương pháp nhân tử Lagrange để tìm một lời giải giải tích tường minh.

**Các công thức**

Quan hệ giữa xác suất bị chặn  $B_{pni}$  với băng thông được cấp  $C_{vpn}$  hay tốc độ phục vụ  $\mu_{ni}$  phụ thuộc vào các tham số sử dụng, cấu trúc bộ đệm, mô hình hàng đợi,... Trên hình 2 là 2 kiểu bộ đệm: phân tách hoàn toàn và chia sẻ hoàn toàn. Dưới đây ta xét trường hợp thứ hai và ký hiệu  $B_{pni} = B_{pncsi}$ , bỏ qua chỉ số  $n$  của đường ảo nên các công thức có giá trị cho một  $VP$  và cả tuyến. Phương pháp ở đây dựa vào các phương trình cân bằng luồng vào và luồng ra của từng lớp lưu lượng ([3]).



Hình 2. Cấu trúc bộ đệm

$$\begin{aligned}
\mu_1 P_r(k_1 + 1, \dots, k_i, \dots, k_M) &= \lambda_1 P_r(k_1, \dots, k_i, \dots, k_M) \\
&\dots\dots\dots \\
\mu_i P_r(k_1, \dots, k_i + 1, \dots, k_M) &= \lambda_i P_r(k_1, \dots, k_i, \dots, k_M) \\
&\dots\dots\dots \\
\mu_M P_r(k_1, \dots, k_i, \dots, k_M + 1) &= \lambda_M P_r(k_1, \dots, k_i, \dots, k_M)
\end{aligned} \tag{4}$$

trong đó  $P_r(\cdot)$  là xác suất hệ thống ở trạng thái  $\mathbf{k} = \{k_1, \dots, k_i, \dots, k_M\}$ ,  $k_i$  là số dịch vụ của lớp dịch vụ thứ  $i$ . Từ đó ta có công thức

$$P_r(k_1, \dots, k_i, \dots, k_M) = G(S, M)^{-1} \prod_{i=1}^M \frac{\rho_i^{k_i}}{k_i!} \tag{5}$$

trong đó  $\rho_i = \lambda_i / \mu_i$  là độ sử dụng,  $G$  là hằng số chuẩn hoá

$$G(S, M) = \sum_{\mathbf{k} \in S(C, M)} \prod_{i=1}^M \frac{\rho_i^{k_i}}{k_i!} \tag{6}$$

$S(C, M)$  là không gian trạng thái được định nghĩa như sau

$$S(C, M) = \{k_1, \dots, k_i, \dots, k_M : \sum_{i=1}^M k_i h_i \leq C\} \tag{7}$$

trong đó  $C$  là dung lượng truyền  $C_{vpn}$  của  $VP_n$  hoặc  $C_l$  của tuyến. Cho trường hợp ghép kênh theo tốc độ định, ta có xác suất bị chặn của cửa dịch vụ lớp  $i$

$$B_{pcsi} = \frac{\sum_{j=0}^{h_i-1} G(C-j, M)}{G(S, M)} \tag{8}$$

trong đó

$$G(S, M) = \sum_{j=0}^C G(j, M). \tag{9}$$

Ta có công thức hồi quy

$$G(j, M) = \begin{cases} 1 & \text{cho } j = 0 \\ \frac{1}{j} \sum_{i=0}^M \rho_i h_i G(j - h_i, M) & \text{cho } j = 1, \dots, C \end{cases} \tag{10}$$

**Phương pháp giải.** Từ các công thức trên ta thấy, một lời giải giải tích cho bài toán tối ưu (1), (2), (3) là rất phức tạp. Nhưng, nếu biết được  $\mu_{ni}$  thì việc tính  $B_{pcsn}$  dễ dàng hơn rất nhiều. Do đó, người ta đi đến giải bài toán sau đây ([4]).

**Tiêu chuẩn minmax**

$$\min_{C_{vpn}} \{ \max B_{pcsn} \} \tag{11}$$

với ràng buộc

$$\sum_{n=1}^{N_{vp}} C_{vpn} \leq C_l \tag{12}$$

trong đó

$$B_{pcsn} = \sum_{i=1}^{M_n} B_{pcsn_i} \quad (13)$$

Đây cũng là bài toán quy hoạch phi tuyến, nhưng dễ giải hơn nhờ thuật toán sau ([4]).

**Thuật toán 1.** Chỉ định bằng thông tốc độ đình.

*Bước 1.* Khởi đầu bằng cách phân đều hay theo tỷ lệ chia bằng thông cho các  $VP_n$ . Theo (8) - (10), ta tính xác suất kết nối bị chặn  $B_{pcsn}$  và ký hiệu bằng  $B_{pcsn}(0)$ ,  $n = 1, 2, \dots, N_{vp}$ . Chọn số gia bằng thông  $\Delta C$ , số  $\varepsilon < 1$  và đặt  $k = 1$ .

*Bước 2.* Tìm  $B_{pmin}(k - 1) = \min_n \{B_{pcsn}(k - 1)\}$ , gọi  $VP_n$  tương ứng là  $VP_i$ . Tìm  $B_{pmax}(k - 1) = \max_n \{B_{pcsn}(k - 1)\}$ , gọi  $VP_n$  tương ứng là  $VP_j$ .

Nếu  $(B_{pmax}(k - 1) - B_{pmin}(k - 1)) / (B_{pmax}(k - 1) + B_{pmin}(k - 1)) \leq \varepsilon$  thì dừng.

*Bước 3.* Tăng (giảm)  $\Delta C$  cho bằng thông của  $VP_j (VP_i)$  là đường có xác suất kết nối bị chặn lớn nhất (nhỏ nhất). Tính  $B_{pcsj}(k - 1)$ ,  $B_{pcsi}(k - 1)$ .

*Bước 4.* Nếu  $(B_{pcsj}(k - 1) - B_{pcsi}(k - 1)) / (B_{pcsj}(k - 1) + B_{pcsi}(k - 1)) \leq \varepsilon$  thì nhảy về bước 2. Nếu khác đi thì đặt  $k = k + 1$  và nhảy về bước 3.

Thuật toán chỉ định bằng thông trên đây được coi là phù hợp với khả năng tính toán của mạng. Lời giải minmax trong (11) đảm bảo các đường ảo có xác suất bị chặn gần bằng nhau, tức là có khả năng thông giống nhau. Tuy nhiên, ta chỉ đạt được điểm tối ưu cục bộ trong lân cận của điểm bắt đầu. Bài toán tối ưu cơ bản (1) có ý nghĩa sâu sắc hơn, song cần phải có phương pháp tính gần đúng phương trình (8) hoặc phải dựa vào mô hình lưu lượng mới, ví dụ theo cách tiếp cận phương trình vi phân ma trận để tìm lời giải tiệm cận trên cơ sở nghiệm trội ([2]).

**Bài toán cấp phát bộ đệm cho các đường ảo.** Có thể giải tương tự như trên theo hàm mục tiêu:

$$\min_{B_n} \{ \max B_{pcsn} \} \quad (14)$$

với ràng buộc

$$\sum_{n=1}^{N_{vp}} B_n \leq B_t \quad (15)$$

trong đó  $B_n$  và  $B_t$  là dung lượng bộ đệm của  $VP_n$  và cả tuyến. Ta có các công thức

$$B_{pcsn} = \frac{Q(B_n)}{G(B_n)} \quad (16)$$

$$Q(B_n) = \sum_{i=1}^{M_n} A_i \rho_i^{B_n} \quad (17)$$

$$G(B_n) = \sum_{i=1}^{M_n} A_i \frac{1 - \rho_i^{B_n+1}}{1 - \rho_i} \quad (18)$$

$$A_i = \frac{1}{1 - \frac{1}{\rho_i}} \prod_{j \neq i} \frac{1}{1 - \frac{\rho_j}{\rho_i}} \quad (19)$$

### 3. ĐIỀU KHIỂN CHẤP NHẬN KẾT NỐI

**Quá trình CAC.** CAC là một tập các hành động mà mạng phải thực hiện để quyết định chấp nhận hay từ chối yêu cầu kết nối mới trên cơ sở tính toán tài nguyên mạng nhằm đảm bảo chất lượng dịch vụ (QoS) yêu cầu. Khi một yêu cầu kết nối mới đến, CAC khởi động chương trình định tuyến nhằm chọn một đường giữa nút nguồn và nút đích “tốt nhất” theo chính sách định tuyến nào đó. Giả sử tại nút đang xét hiện có  $N_n^i$  dịch vụ từ các nguồn khác chiếm băng thông hiệu dụng  $C_{vpebn}(N_n^i)$  và bộ đệm hiệu dụng  $B_{en}(N_n^i)$ . CAC phải tính lại bằng thông  $C_{vpebn}(N_n^i + 1)$  và bộ đệm  $B_{en}(N_n^i + 1)$  đủ đảm bảo QoS yêu cầu cho tất cả các dịch vụ cũ và mới. Nếu  $C_{vpebn}(N_n^i + 1) \leq C_{vpn}$  và  $B_{en}(N_n^i + 1) \leq B_n$  thì kết nối mới được chấp nhận, ngược lại sẽ bị từ chối. Một khi chấp nhận, CAC ấn định nhận dạng đường ảo  $VPI$  và nhận dạng kênh ảo  $VCI$  cho kết nối và gửi băng thông hiệu dụng  $s_n^i = C_{vpebn}(N_n^i + 1) - C_{vpebn}(N_n^i)$  tới bộ điều khiển gáo rò để đặt tốc độ sinh thê. Nếu kết nối bị từ chối ở một nút nào đó, CAC có thể định tuyến lại cho đến khi hết mọi khả năng mà vẫn không được thì kết nối bị chặn.

Trên đây là điều khiển kết nối trên cơ sở băng thông hiệu dụng. Tiếp theo, chúng tôi trình bày một phương pháp tính  $C_{vpebn}(N_n^i + 1)$ .

**Phương pháp tính băng thông hiệu dụng.** Băng thông cần thiết cho một dịch vụ lớp  $i$  được chấp nhận vào  $VP_n$  mà không ảnh hưởng đến QoS của các dịch vụ đang được phục vụ trong  $VP_n$  là ([3]):

$$h_{eni} = \frac{\left(\frac{1}{T_{onni}} + \frac{1}{T_{offni}} - A_n h_{ni}\right) - \sqrt{\left(\frac{1}{T_{onni}} + \frac{1}{T_{offni}} - A_n h_{ni}\right)^2 + \frac{4A_n h_{ni}}{T_{offni}}}}{-2A_n} \quad (20)$$

trong đó

$$A_n = \ln\left(\frac{1}{\lambda_{clreqn}}\right) \frac{1}{X_n}, \quad (21)$$

$$X_n = D_{reqn} C_{vpn}, \quad (22)$$

$T_{onni}$  là chu kỳ đóng (ON) của dịch vụ lớp  $i$ ,

$T_{offni}$  là chu kỳ mở (OFF) của dịch vụ lớp  $i$ ,

$h_{ni}$  là tốc độ đỉnh của dịch vụ lớp  $i$ ,

$\lambda_{clreqn}$  là yêu cầu tốc độ tổn thất tế bào,

$D_{reqn}$  là yêu cầu trễ tế bào,

$X_n$  là yêu cầu cỡ của bộ đệm.

Nếu thay  $h_i$  trong (8) bằng  $h_{eni}$  từ (20) ta nhận được xác suất bị chặn của dịch vụ lớp  $i$  vào  $VP_n$  trong trường hợp ghép kênh theo băng thông hiệu dụng. Do đó, có thể xây dựng thuật toán chỉ định băng thông hiệu dụng như sau.

**Thuật toán 2.** Chỉ định băng thông hiệu dụng.

*Bước 1.* Khởi đầu: sử dụng Thuật toán 1 để tính băng thông  $C_{vpebn}(0)$  cho các  $VP_n$ ,  $n = 1, 2, \dots, N_{vp}$ . Chọn  $\varepsilon'$  và đặt  $k = 1$ .

*Bước 2.* Tính  $X_n(k) = D_{reqn} \cdot C_{vpn}(k - 1)$ ,  $A_n = \ln(1/\lambda_{clreqn})/X_n(k)$  và  $h_{eni}(k)$  theo (20). Sử dụng Thuật toán 1 và  $h_{eni}(k)$  để tính  $C_{vpebn}(k)$  cho các  $VP_n$ ,  $n = 1, 2, \dots, N_{vp}$ .

*Bước 3.* Nếu  $(C_{vpebn}(k) - C_{vpebn}(k - 1))/C_{vpebn}(k) \leq \varepsilon'$  thì dừng. Khác đi thì đặt  $k = k + 1$  và trở về bước 2.

Ta thấy thuật toán chỉ định băng thông hiệu dụng khá phức tạp trong khi CAC đòi hỏi tương đối khắt khe về tính toán theo thời gian thực. Do đó cần thiết xây dựng mô hình lưu lượng mới, tìm lời giải gần đúng để rút ngắn thời gian tính. CAC đảm bảo được QoS. Song, có thể nâng cao hiệu năng của mạng bằng cách bổ sung cơ chế cho phép thâm nhập bộ đệm theo mức ưu tiên về tổn thất và thời gian trễ của dịch vụ với trợ giúp của bộ đệm nhiều ngưỡng ([2]).

#### 4. ĐIỀU KHIỂN TRÊN CƠ SỞ TỐC ĐỘ

Như trình bày ở trên, RNM chia băng thông và bộ đệm của nút cho các đường ảo, còn CAC căn cứ vào các yêu cầu QoS tính toán tài nguyên mạng để quyết định chấp nhận kết nối. Từ quan điểm điều khiển tự động thì lưu lượng được điều khiển theo kiểu chương trình hay điều khiển vòng hở, thiếu thông tin phản hồi từ đường truyền và nút đích nên không cho độ chính xác cao, kém khả năng chống nhiễu. Do đó, RBC cần áp dụng một số cơ chế, đặc biệt là nguyên lý phản hồi nhằm nâng cao chất lượng điều khiển. Các phương pháp chính của RBC là kỹ thuật cửa sổ, bộ điều khiển gáo rò, điều khiển tham số sử dụng/mạng, điều khiển nghẽn...

**Kỹ thuật cửa sổ.** Gồm cửa sổ trượt và cửa sổ bước. Trong kỹ thuật cửa sổ bước, cỡ cửa sổ cố định  $W$  được ấn định cho một kết nối đường ảo giữa nút nguồn và nút đích và bộ đếm bước PC được đặt bằng  $W$ . PC giảm đi 1 mỗi khi nút nguồn gửi 1 gói tin. Khi nội dung PC bằng 0 thì không có gói tin nào được gửi đi nữa. Nút đích nhận các gói tin và gửi tín hiệu trả lời. Nếu trả lời là tích cực (nút đích nhận đủ) thì nút nguồn đặt lại cỡ  $W$  cho PC. Ngược lại, chẳng hạn do nghẽn thì nút nguồn phải truyền lại hoặc giảm tốc độ.

**Bộ điều khiển gáo rò.** Do I. Cidon và I. S. Gopal sáng chế và áp dụng cho mạng thử nghiệm tích hợp tốc độ cao của IBM (1988), hiện được dùng rất rộng rãi và đã qua một số lần cải tiến. Nguyên tắc làm việc giống như điều khiển thẻ trong mạng chuyển mạch gói. Thẻ được sinh với tốc độ  $s$  bằng băng thông hiệu dụng đã tính cho dịch vụ hoặc theo hợp đồng giữa người sử dụng với mạng và sẽ ngừng khi thùng đựng thẻ đầy. Tế bào đến với tốc độ  $\lambda$  nếu thùng còn thẻ nó sẽ được vào mạng và làm giảm số thẻ trong thùng đi 1, còn nếu thùng rỗng tế bào bị gọi là “vi phạm”. Trong thiết kế ban đầu không có cơ chế đánh dấu, các tế bào vi phạm bị rơi. Do tính bùng phát của lưu lượng, điều này vẫn có thể xảy ra ngay cả khi bộ đệm của nút nguồn còn trống và băng thông của tuyến còn thừa. Để tận dụng tài nguyên mạng, ở những phiên bản sau các tế bào vi phạm vẫn được truyền nhưng bị đánh dấu để khi mạng nghẽn chúng bị rơi trước nhằm đảm bảo QoS cho các tế bào khác. Một số cải tiến khác cũng đã được thực hiện, như thêm vào bộ đệm tế bào, cơ chế tính lại số thẻ thích nghi với tình trạng của mạng v.v...

**Điều khiển nghẽn.** Theo phân loại của diễn đàn ATM, ta có các lớp dịch vụ:

- + Tốc độ bit không đổi CBR (Constant Bit Rate),
- + Tốc độ bit thay đổi thời gian thực RT VBR (Real-time Variable Bit Rate),
- + Tốc độ bit thay đổi không thời gian thực NRT VBR (Non Real-time Variable Bit Rate),
- + Tốc độ bit khả dụng ABR ( Available Bit Rate),
- + Tốc độ bit không xác định UBR (Unspecified Bit Rate).

Trong đó chỉ có lớp ABR được trang bị bộ điều khiển luồng làm việc theo nguyên lý phản hồi, nhằm thay đổi tốc độ của nguồn phù hợp với trạng thái mạng. Cứ truyền 32 tế bào số liệu, nút nguồn lại gửi 1 tế bào quản lý tài nguyên RM (Resource Management) để thu thập thông tin, cụ thể là khoảng trống nhỏ nhất của các nút trên suốt đường kết nối ảo. Căn cứ vào số liệu này, nút nguồn hiệu chỉnh lại tốc độ phát của dịch vụ ABR sao cho mạng không bị nghẽn. Do đó mà còn có tên “điều khiển nghẽn”. Bởi vì dịch vụ được cấp phát băng thông

và bộ đệm chỉ 1 lần và không thay đổi trong suốt quá trình phục vụ, nên điều khiển luồng còn có tác dụng điều hoà, tận dụng tài nguyên mạng (nhất là cho ghép kênh tốc độ đỉnh, các dịch vụ có tính bùng phát cao như thoại, hệ thống thường xuyên chịu nhiễu ngẫu nhiên ...). Do tầm quan trọng của nó, kỹ thuật điều khiển luồng đã được cải tiến không ngừng, từ đơn giản như tăng, giảm theo mức đến áp dụng những phương pháp tiên tiến của điều khiển tự động hiện đại ([5, 6]).

## 5. KẾT LUẬN

Mạng ATM là một thử nghiệm đã đạt được những kết quả bước đầu trong việc chuyển từ các mạng chuyên dụng sang một mạng tích hợp đa dịch vụ trong viễn thông hiện đại. Ưu điểm lớn của mạng ATM, ví dụ so với Internet (IPv4) là đã xây dựng được cơ chế đảm bảo QoS. Song, mạng ATM cũng bộc lộ những khuyết điểm của một mạng hướng kết nối, làm giảm lợi thế (so với IPv6) trong cuộc đua trở thành nền tảng của mạng đa dịch vụ tương lai. Trên đây chúng tôi đã trình bày tổng quát về điều khiển lưu lượng trong mạng ATM, phân tích đánh giá những ưu, khuyết điểm đồng thời nêu một số nhận xét, đề xuất góp phần cải tiến, hoàn chỉnh kỹ thuật điều khiển cho mạng đa dịch vụ.

## TÀI LIỆU THAM KHẢO

- [1] Z. Dziong, *ATM Network Resource Management*, McGraw-Hill, 1997.
- [2] A. Elwalid, D. Mitra, Analysis, approximation and admission control of a multiservice multiplexing systems with priorities, *IEEE INFOCOM*, 463–472.
- [3] X. Gu, K. Sohraby, D. Vaman, *Control and Performance in Packet, Circuit and ATM Networks*, Kluwer Academic Publishers, 1995.
- [4] M. Logothetis, S. Shioda, G. Kokkinakis, Optimal virtual path bandwidth management assuring network reliability, *Proceeding of ICC' 1993*.
- [5] S. Mascolo, Congestion control in high-speed communication networks using the Smith principle, *Automatica* **35** (1999) 1921–1935.
- [6] T. Ndousse, Fuzzy neural control of voice cell in ATM networks, *IEEE J. Select. Areas Commun.* **12** (9) (1994) 1488–1494.

Nhận bài ngày 25 - 7 - 2003